DOCUMENT RESUME

ED 038 331 SP 003 630

AUTHOR Michelson, Stephan
TITLE The Association of Teacher Resourceness with Children's Characteristics.
PUB DATE 4 Feb 70
NOTE 75p.; Paper for the Bureau of Educational Personnel Development Conference, "How Do Teachers Make a Difference?"

EDRS PRICE EDRS Price MF-$0.50 HC-$3.85
DESCRIPTORS Evaluation Criteria, Exceptional Child Education, Research Needs, School Administration, Statistical Analysis, Student Teacher Relationship, *Teacher Characteristics, Teacher Distribution, *Teacher Evaluation, *Teacher Influence
IDENTIFIERS Equal Educational Opportunity Survey

ABSTRACT

A statistical analysis of data from the Equal Educational Opportunity Survey (EOS), conducted by the U.S. Office of Education in 1965, leads to the concept that teacher influence on or "resourceness" for a child differs by the type of child. This concept is called "teacher specificity." Single linear regression analysis and a 3-equation system with simultaneous estimation were applied to data from the EOS sixth grade questionnaire, teacher questionnaire, and principal questionnaire to develop a correlation between school resources and variations in students' raw test scores for two populations--white students and black students. The difference in correlation coefficients for the two populations is the basis for the idea that these resources have different value "resourceness" for whites and blacks. Extension of this idea suggests that teacher evaluation might be based on the specific situation rather than a general set of standards. Greater local control, specifically by school principals, would be a method of implementing this type of evaluation. An analogy is drawn between the teacher-student relationship for normal children and for exceptional children. (RT)

THE ASSOCIATION OF TEACHER RESOURCENESS WITH CHILDREN'S CHARACTERISTICS

Stephan Michelson

February 4, 1970

Prepared for the Office of Education, Bureau of Educational Personnel Development, Conference "How Do Teachers Make A Difference?"

The author is a Research Associate at the Center for Educational Policy Research, and Lecturer at the Graduate School of Education, Harvard University. My colleagues at CEPR, David K. Cohen, Herbert Gintis, Christopher S. Jencks, Martin Katzman, and Marshall S. Smith, have all contributed to the production of this paper. In addition, the influence of Gordon Gillies, Mildred Howe, and Carol Stewart should be noted. Much of this work should bear the joint authorship of Henry M. Levin, from whose initiative the study was undertaken, and in conjunction with whom it has continued. Extraordinary research assistance was provided by Polly Harold.

If we can arbitrarily, and without precise distinction, consider that schooling might affect skills, values, and personalities, there is a difference of opinion about which of these actually occurs:

> The school, then, is an organizational embodiment of a major social institution whose prime function is to bring about developmental changes in individuals. It is an agency of socialization whose task is to effect psychological changes that enable persons to make transitions among other institutions; that is, to develop capacities necessary for appropriate conduct in social settings that make different kinds of demands on them and pose different kinds of opportunities.* [(9), p. 3.]

> As social scientists, we maintain a skeptical view concerning the efficacy of formal schooling for the teaching of values. To the social scientist a view of formal education as an omnipotent socializing agent shows an exaggerated regard for education. The social scientist is not convinced that institutions of formal education are capable of accomplishing all the mammoth tasks that some apparently expect of them. The classroom may well be a place where formal skills are learned; it may also contribute to the transition from the family to the larger society. Finally, it may contribute somewhat to the maintenance of a core culture or the creation of a cultural synthesis. But whether formal education really has much influence on either cultural values or social behavior is not evident.** [(14), p. 7.]

The recent rapid entry of model-oriented social scientists, sociologists and economists particularly, into educational research has brought an unfortunate emphasis on the latter point of view. Skills, being more measurable, are taken to be the outcome of schooling in most statistical studies. An empirical approach not relying on statistical analyses, led Dreeben to his conclusion: He observed the structure of schools, asked what that structure could produce. With Callahan's work [5] as additional evidence, one could conclude that the major outcome of schools has not historically been meant to be cognitive skills. And for purposes of generating income, the work of Gintis [12] and Berg [2] indicates that cognitive skills are not necessarily the most useful outcomes of schooling.

Nonetheless, recent investigations of school outcomes and the school characteristics that affect them (or do not affect them) have centered on these skills which schools may not have been intended to produce, are not structured to produce, and would not necessarily benefit people if they did produce. Studies continue, this one no exception, to ask questions about the relationship between inputs and outputs despite the fundamental lack of knowledge of what outputs are desired, possible, and efficacious.

The ideas set out here, the kind of research described, therefore must not be taken as evidence for one kind of school structure as against another. It is too facile--and too common--to investigate one area of school production, ignoring the consequences in other areas. It could certainly be that a technique, say tracking, did successfully increase

cognitive skill acquisition at all levels, and yet was entirely unacceptable as a method of school organization.* Thus I will discuss the question of the specificity of teacher characteristics in producing outputs such as reading scores, or even student attitudes, without meaning to imply that if certain types of children respond better to different types of teachers, then the schools should be organized to match them. This will be one argument that some such organization might be desirable, but for many reasons it may not be. I will conclude the paper with a suggestion about a school authority structure which might better accommodate my findings and general theory. But this is meant to be tentative and suggestive, not persuasive. That is, there are two kinds of arguments against my findings: First, one could argue that they are incorrect or at least inconclusive. This is a technical kind of discussion which would hopefully result in the design of a test which would confirm or deny the results reported here. But second, one could accept my results, but reject their implications because the school policies they imply are unacceptable. I hope only to set the tone, and, I pray, a trend, that one cannot advocate school policy on the basis of a very limited set of school outcomes, say, on the basis of skill production, absent any knowledge of the personality or value system effects of that policy.

---

*I am not sure this possibility is actually as likely as, in warning against it, I must assume it is. If the social outcomes are disastrous, the test scores are likely to be poor also. In fact, to assume that students could be both extremely alienated and maximum performers is absurd. But since skills as measured by test scores and other social outcomes are not perfectly correlated, the warning is still in order. And the question of deciding on a method when it helps some people but not others, and yet must be imposed on none or all--which is the nature of tracking--points out the inadequacy of correlation as a substitute for value judgments.

The Paper in Outline

With this brief caveat, I will here outline the intended progress of this paper. The next section begins with a limited discussion of school production, and discusses some characteristics which I deem important to an ex post cross section investigation of the effects of schooling. This discussion is intended to begin to clear the air about different conclusions which have been reached regarding the association of school and teacher characteristics with student test scores. The way to determine which study has reached correct statistical conclusions is to investigate the properties of the investigations: the samples, definitions of variables, statistical techniques employed. These must be justified, and the results of a study must be weighted by the appropriateness of the techniques.

Following this exposition, ordinary least square estimates of the relationships between test scores and school inputs are presented and discussed. The interpretation of statistical results is a separate issue from their correctness, and my claims for my interpretation will be far more cautious than my claims for my findings. There, however, some basic points of this paper will begin to emerge. A brief exposition of a simultaneous equations system will add fuel to the fire.

In the third section, the implications which might be drawn from the statistical presentation are examined. Concepts such as "resourceness" and "specificity" will be defined in terms of the regression results. However, the inferences are tentative, and some ways in which they might be altered are suggested. I will conclude the paper, then, with a brief fourth section about the implications of this work and its tentative interpretation for school administration. A possible modification of the

present structure is offered--as is the whole paper--as suggestive, not definitive.

I include, as an Appendix, a review of some material from the field of teaching "exceptional children," especially the blind, deaf, and mentally retarded. The emphasis will be on the acceptance, in these cases, of the concept of teacher specialization by type of child, as opposed to specialization by subject matter. My major effort in the text of this paper is merely to extend that already accepted notion to a broader view of the need to consider the characteristics of the pupils in making teacher assignment, and in teacher training.

STATISTICAL INVESTIGATION OF TEACHER RESOURCENESS

The exposition here will not be abstract theory, but the theory which leads directly to use of the Equal Educational Opportunity Survey (1965) data to investigate the association of school and teacher characteristics with student outcomes. The exposition will discuss the following, in order: the data sample, the observations, the variables, and the statistical technique. Especially in the last section, I will assume that the reader is familiar with the paper by Henry M. Levin [27] prepared for this conference. The sample and variables used here are identical to those used by Levin, and the simultaneous model is similar.*

The Sample

The data used in this study came from the Equal Educational Opportunity Survey (hereinafter referred to as EOS), conducted by the U. S. Office of Education in 1965, and reported in 1966 as Equality of Educational Opportunity, (hereinafter referred to as EEO), often called "The Coleman Report" after its major author [7]. Many people have investigated the EOS data, arriving at different conclusions about the association of school characteristics with achievement.** I believe most of the differences, be-

---

*The data and models have been derived jointly by Levin and myself both concurrently at Stanford and Harvard, and in summer work together at Stanford. Randall D. Weiss has also contributed to the formulation and estimation of the simultaneous equations model. The first person singular is used on the following pages to assign responsibility, not credit, to myself.

**See [1], [7], [16], [18], [28], [37].

sides those in statistical technique, can be attributed to the choice of sample. The question must be: what sample of the population should we look at to determine the extent of this association?

The basic constant which must be assumed in these studies is that all schools observed must be trying to maximize the same thing, hopefully our output measure, though that is not strictly necessary.* And they must be acting this way for all children in the school, or else we must observe only those children for whom this is true. Figure 1 shows the case in which two outputs, A and B, are related by the "production frontier" as indicated. This is merely the locus of possible outcomes with the resources at hand.** Schools A_1 and B_1 tend to produce A and B respectively, as do A_2 and B_2, which are endowed with more resources. The more resources of B_2 do not produce more of A than A_1, nor does A_2 produce as much B as B_1. We can find statistically that resources do not affect either A or B, when in fact they affect both regardless of which is preferred by the school.***

Within a school district there is a variation of social class among schools which might lead to variation in aims of programs. There is also variation of class within schools which might induce differential program

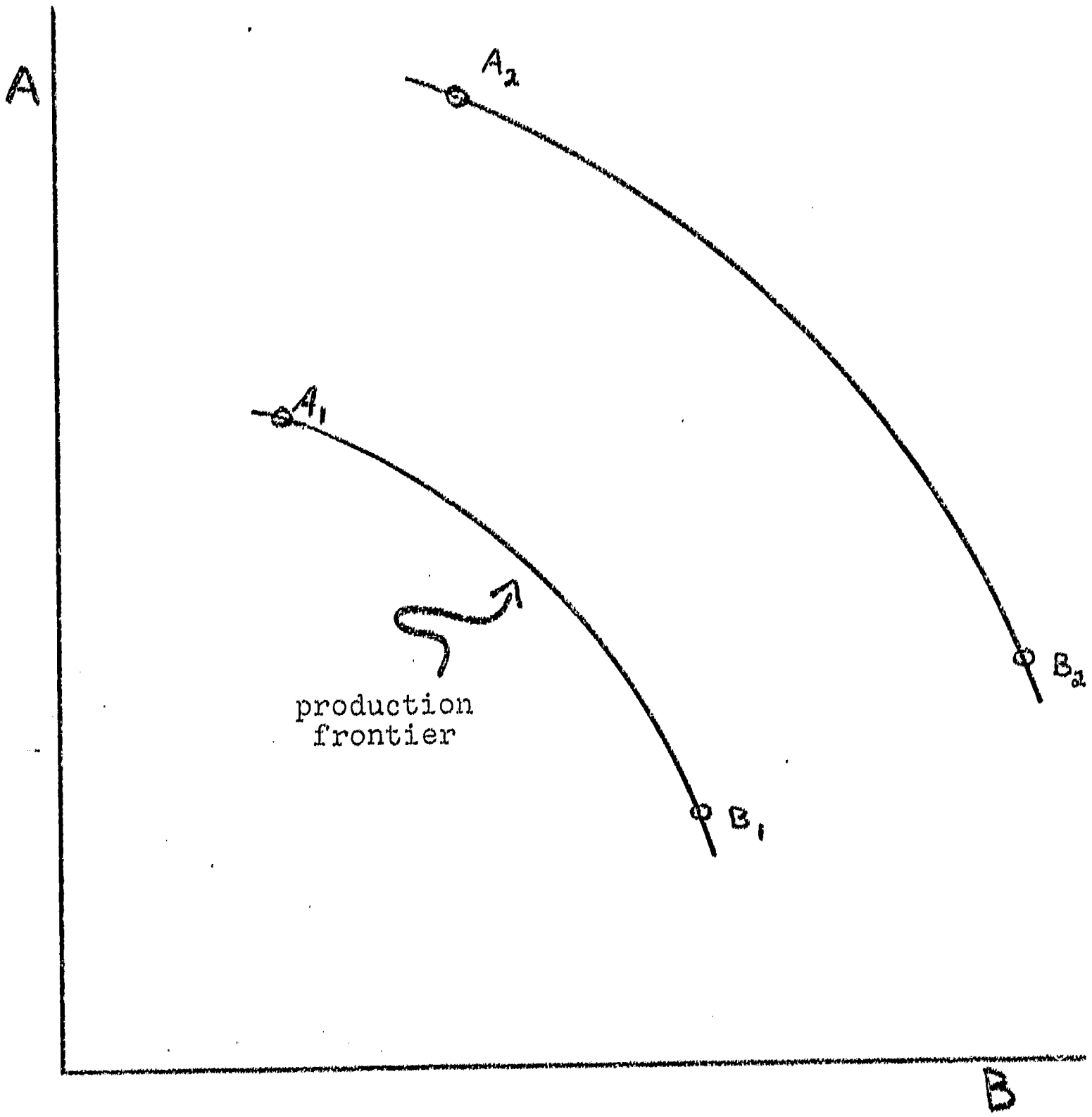
---

*Maximization of a complement of our output measure would suffice, if the complementarity were strictly linear.

**One might object that if schools tried to maximize different things, they would not do so with the same kinds of inputs, but would employ those best for the output. For example, trade schools do not hire verbally proficient, but manually proficient teachers. However, elementary schools are equipped by tradition more than by rational management, the maximization of various outputs taking place on location, not overtly on central direction.

***Variations in inputs do not correlate with either output when the other output is not accounted for.

figure 1



Note: A and B represent outcomes of schooling.

Each production frontier represents the locus of possible outcomes from the school resources (2 indicating more resources than 1), depending on to what ends they are used.

aims for different children. The same kind of variation in aim occurs among districts, but I think less of this variation occurs within than between districts. Many overt and covert policies of school boards which indicate differences in their aims can be controlled: the factory town which in general produces workers for the plant, the prestige suburb which produces college graduates, the central city which produces a spectrum and, like New York or Boston, allows its citizens to be chosen "fairly" (that is, by exam) into the prestige high schools. The aims of the school board, the environment of the city (air pollution, garbage collection, etc., all of which could have education consequences; even the mean temperature)--all of these variables are controlled by choosing one large city with several schools. This sample is not perfect: the dilemma of Figure 1 has not been solved.* But I believe it is considerably reduced. To the extent that this problem still occurs, the observed association between school characteristics and children's achievement is reduced below the actual association.

In addition, previous studies have included children in the sample who had not been in the same school in preceding years. They were identified with their correct home variables, but incorrect school variables. In many cases, this is probably not serious: some children transfer among

---

*Katzman [23] shows, for example, that the outputs of different elementary schools in Boston are quite different. I infer from his findings that the aims of these schools differ, though Katzman does not agree that this inference should be drawn. Different goals of schools, and the different goals which the school has for different children, is a vital problem in this type of analysis.

very similar schools. By choosing a central city, the upwardly mobile children who have recently moved to the suburb are eliminated. Those who will soon move out may remain unidentified. However, the resulting bias in the association of school variables with output is toward zero, while not affecting the home variables. Although this bias is not unacceptable, it is not necessary. I have eliminated from the sample those children who had not been in the school in question since the first grade.*

The sample, then, comprises those children in a large eastern city, "Eastmet," who had attended only one elementary school. This sample was divided into whites, blacks, and others, only the white and black samples being utilized for this study.**

Observations

Debate among researchers has been endless about whether one ought to observe individual children or school means in this type of study. The question is often based on argument about the number of degrees of freedom when individual children are used: is the number of schools, or the number of children the base? I will surely not answer this question to the satisfaction of people who think differently, but explanation of my procedures follows.

Most of the variation in test scores occurs within schools. Children within schools differ more from each other than schools as groups do

---

*Strictly speaking, I need only have eliminated those who had not been there since the fourth grade, since I used only the later grade teachers. However, the questions in EOS did not allow this distinction.

**Those pupils who said they were black and something else (Puerto Rican or Mexican) were included due to a coding error.

from each other. This is an interesting finding. It has been used to show that schools are relatively ineffective, for better schools should produce better students. However, since there is grouping within as well as between schools, there is no reason to believe that schools are ineffective on these grounds. We are back to the Figure 1 problem: if each school chooses some students on whom to stress the outputs we measure, others to stress other outputs, then schools could be totally effective, produce all variations, and yet there would be more variation within than between schools. Furthermore, if the selection were made by social class, then the social class variables would be associated with output differences.

To see this, consider several schools which are formed by random selection of students from a community. Within each school, children are grouped by their behavior, which is correlated with their social class. The more cooperative, passive students are put in the high "track," which stresses academic output. The lower tracks stress behavioral outputs more and more. By grade six, the upper track has been reduced in relative size by elimination of those who, though behaviorally adept, do not succeed academically. Lower track academic successes, however, do not move up.* The mean social class and mean test scores will be equal among schools. Within schools social class and test score will correlate highly. If one were bound to interpret "social class" as necessarily indicating home influence, and observed school means, he would conclude that schools had no effect.

---

*This description of the school is essentially adopted from Mackler (30).

By construction, however, this conclusion would be incorrect. In fact schools are not alike by social class or achievement. Some inter-school variation is observed, and it correlates with social class more than with school characteristics. However, consider the other polar case: the tracking I described above now occurs among schools. School #1 is initially selected by social class, though by grade six some upper class children have been moved to schools #2 and lower. The interaction of high social class and reasonably high ability would perfectly predict placement in school #1, and therefore test score. By linear regression where only social class is entered, that variable would predict quite well. Since school resources in this case would be allocated by function--academic resources to the academic school, etc.--school variables would also predict outcome.

The facts seem to lie somewhere in the middle: schools are relatively homogeneous by social class, as in the first polar case, but not completely so. Since abilities vary within social class, and social behavior varies within each school, each school can have its academic, each its non-academic group. The variation between schools, which would be greater if schools were treated as in the second polar case, is reduced by intra-school grouping. But some between school output variation still occurs, and it is associated with the mean social class of the school. The interpretation that it is therefore "due to" the social class of the school is correct, but the interpretation that this operates through home life of the students is incorrect. Similarly, when one finds that a lower class child does better academically in an upper class school, one need not con-

clude that this is due to the direct influence of his classmates on him. It may be that the school he is in stresses academic outputs more than the schools of his social class equals. There is simply no reason to believe, from the correlation between social class and academic success, either by school mean or by individuals, that the cause of this association is the home life of the children.*

This argument, then, speaks to the issue of whether to observe school means or individuals in this sense: By the models just presented, the association is between the child and his output. To what extent this association is found between schools depends on the school structure, i.e., on to what extent grouping occurs within or between schools. This extent may vary from city to city, and even within cities. It seems wiser, then, to observe children directly.

There are other arguments: Children are of more interest than schools. I don't know what to make of the fact that mean school resources do not correlate with mean school output. The resources going to a child might still be very important. Since the variables labeled "school characteristics" do not vary within schools, obviously I cannot determine the effect of within school variation in these characteristics with these variables. But I can still pick up their effect to the extent that I can identify the individual characteristics by which these inputs are allocated. The problem is partly one of interpretation, and partly that the correlation between individual characteristics which we measure (which exclude, for

---

*There may be great reason to believe this, and it may be true, but no direct statistical inference of this nature can be made.

example, direct behavioral measures) and the allocation of school inputs may not be perfect.

The variation which we want to explain, then, is variation in student scores, not variation in school scores. The fact that this variation occurs mostly within schools, that the percent of this variation which we can explain with the variables we have is small (about 47 percent of verbal score variation, 36 percent of reading score, for whites), is a fact not to be covered up by observing the relatively invariant school means.

The argument about degrees of freedom, in this context, is nonsense. We observe children in situations. There are not as many situations as children. But similarly there are, for example, only two sexes, nine categories of possessions, 50 possible scores on the verbal test. These numbers have nothing to do with degrees of freedom. When two children in the same school receive different test scores, then the association between the school characteristics and those scores is reduced. That is an accurate portrayal of the situation: knowledge of aggregate resources does not predict individual success. It is like observing the difference in behavior between married men and bachelors. If a thousand observations are taken, then the degrees of freedom calculation begins with 999 on taking the mean, and is reduced from that figure by adding independent variables. It is not two. To the extent that variations of behavior within the categories "married" and "bachelor" may occur, they indicate that this variable is not a good explainer of that variation. But the degrees of freedom are not affected by this consideration.

Suppose everyone who is married lives in a private home, and all

bachelors live in apartments. Then entering type of living quarters would be redundant if marital status is already included. Similarly, if there are only 34 schools with whites in Eastmet, no more than 33 school variables can be entered into a regression equation. From the 34th on, each variable can be expressed as a linear combination of the others. But this does not limit the degrees of freedom when some small number of school variables are entered, any more than one would argue that there are only two degrees of freedom in an equation which contains only marital status, despite the fact that marital status and type of dwelling cannot appear in the same equation. In the white equations, 597 children are observed in situations in which the ordering of schoc variables is restricted. All children in school A receive all the inputs in school A, and those in B receive B. Not all possible interactions are directly observed in the data. This is typical of regression data--it is why regression analysis is used. The statistical degrees of freedom does not depend on the many possible (and redundant) variables which are <u>not</u> entered into the equation, but on the number of observations, less 1 plus the number of independent variables.

The argument, then, is that it is reasonable, preferable, and statistically valid to consider children as observations. It is reasonable and preferable because the object of the investigation is to determine the effects of variables on children, not on schools. It is valid because school variables act like any situation variables, and do not restrict the degrees of freedom of the equation.

The Variables

Data are from the sixth grade questionnaire, the teacher questionnaire, and the principal questionnaire, all of which are reprinted at the end of Volume I of EEO. I selected those teachers who were in the third through fifth grades, because the test was given in September of the sixth grade.* The teacher responses were averaged over the school, and the average was applied to each pupil in the school.

This procedure implies that each student moves randomly among teachers through the grades. For future researchers, a suggestion from Marshall Smith** is to weight each teacher by the percent white which he reports relative to the percent white in the school, and apply this weighted figure to white students, and apply the complementary weights to the teachers for black students. This seems to be a better approximation than mine to the data we all desire, but no one has: the correspondence of particular teachers with pupils through several grades. In either case, errors of association should bias significance tests, and possibly (if assignment is non-random) even the statistical relations between teacher characteristics and student outcome towards zero.

A recent study notes that "the evidence suggests that the quality of

---

*In practice, this distinction is of little importance. Sixth grade teachers are not different from fifth grade teachers. Teachers in the fourth grade who were not in that school when the children were in the fourth grade were not eliminated, implying the assumption that they replaced teachers like themselves. The extent and direction to which this is biased is unknown, though replacement of likes seems more probable in high turnover schools, less probable in low turnover schools, where the replacement may be considerably younger than the person replaced.

**From conversation, I understand that Christopher Jencks is experimenting with this weighting scheme.

the principal and staff has a profound influence on [student] improvement." [(33), p. 1] Though in EOS there was evidence on the principal's degree, major, and experience, there was no direct measure of the principal's performance (such as the 30 question test taken by teachers), or attitudes (such as teacher preferences for other school, for different race or "ability" of pupils). I therefore used only his answers to questions about the school, and not about himself.*

Individual student questions were sometimes combined, sometimes divided by possible answers, usually according to my judgment or interest, sometimes according to preliminary findings. For example, I started with a linear age variable, which associated negatively with output: the older the child, the lower the achievement score, controlling for other factors. But there was really no significant difference between a 10 year old and an 11 year old--and in fact, 9 year olds (children who reported that they were nine) were below average. Thus I created binary coded variables for 12 or older, and 9 or younger. On the other hand, I combined nine home items into an index of possessions, not being ready to believe that the possessions of any one provided the information I was seeking.** The names attached to these variables should indicate how they were created.

---

*Preliminary investigation indicated little success with principals' personal variables anyway.

**Unfortunately, however, 80 percent of the total sample had 6 or more of the items, the median being between 7 and 8. For the samples actually employed in the regression analysis, 85 percent of the whites (though only 36 percent of the blacks) had 8 or 9 items. Thus the index does not necessarily contain the precision implied by nine questions. If that item which the children with only 8 do not have is the same item for most of these children, the index merely measures the presence or absence of that item.

For some of the equations to be presented, some interaction variables were also created. These were formed by visual inspection of school summary data. School resources and average student characteristics were looked at, where "resources" were average teacher test score and experience, and the pupil characteristics were possessions and a socio-economic index.* At least four schools had to meet criteria of "low," "mid," or "high" socio-economic status of the students ("peers"), or three categories of resources to qualify as a variable. Three categories of schools were selected this way: high resources but low peer, low resources but mid peer, high resources and mid peer. The effect of each of these categories was not assumed homogeneous, but was made into a separate variable for above median and below median SES for each child. The interaction effect of being a high SES child in a low SES school, or a low SES child in a low SES school could be accounted for separately. These interaction variables were not included in the simultaneous equation system.

The outputs considered are raw test scores of students. A verbal test was the basis of most findings previously reported. This test, and in addition a reading and a mathematics test, are used in the single equation study. In the simultaneous model, only verbal score is used as an academic output. An index of student attitude and his grade aspirations, are also outputs in the model. Grade aspiration means how far the student says he wants to go in school. However, 87 percent of the blacks in the final sample, and 93 percent of the whites had the highest two values among five

---

*The SES index was created by weighting the listed father's occupation by the mean income for his occupation and presumed race (from the race of the child) from the 1960 Census of Population reports for the area of the sample.

possible values. The student attitude question on the other hand, was very evenly distributed. Of eleven possible values, between 10 percent and 20 percent of the blacks in the final sample had each of five values, and 10-20 percent of the whites had each of six values. It seems trivial to assume, but nonetheless important to mention, that high values of grade aspiration indicate "expected" or "socialized" response. The attitude questions, such as "If I could change, I would be someone different from myself" (answers "yes," "no," "not sure"), are not those ordinarily asked of a sixth grade pupil, and so elicit less socialized, more spontaneous responses.

Finally, I will touch here a little on interpretation of variables. The authors of EEO sagely warned about "the danger of unconsidered surrogates," which "can lead to seriously misleading conclusions." They give an example:

> Let us suppose that community attitudes toward the importance and quality of education have substantial effects on the development of student achievement. What would we expect about the apparent relation between achievement and teacher characteristics? Surely we would expect that communities more concerned with education and educational quality would--(1) be more selective in hiring teachers, and (2) pay higher salaries, thus attracting better candidates. As a consequence we might expect an apparent relationship between development of achievement and measurable teacher characteristics to be generated as a surrogate for an underlying relationship between development of achievement and community regard for education, even if teacher characteristics themselves had no effect on achievement. [All quotes, EEO, (7), p. 327.]

This warning is perfectly in order. The example, of course, does not apply in the present case, where one city only is being studied. Strangely, nowhere in EEO is the suggestion made that surrogates can work the other way round: that home items can be surrogates for access to school facilities. Take, for example, the problem of student assignment to teachers, mentioned above. Though there is some meaning to the average teacher characteristic in setting the atmosphere of the school, the deviation from that average which is each child's history may have a regular pattern. I have been told, for example, of a very aware teacher in a Boston suburb who takes her low-tracked class through the school corridors, looking into other classrooms. The students one by one mark, from visual observation through a window in the door only, which track each class is in.* Their estimates correlated well with the actual tracking, the identification coming, says my informant, from the dress of the children in each room. If teacher assignment among tracks is biased, and if the characteristic by which teachers are assigned to higher track students is truly effective, that effect will show as a student variable. It may be in the possessions index, size of family, father's education, mother has a job, etc., whatever correlates with type of dress.

In fact, in assessing the probable direction of surrogates, the side taken by EEO seems perverse. Only student characteristics vary within schools. We know that school facilities are not distributed randomly within schools, and any student variable which is associated with a bias in re-

---

*As this is just an anecdote, not much analysis is required. But I did ask if the class knew either the children or the teachers, i.e., knew the track of each class from some external information. I was assured this was not the case.

source allocation may be a surrogate for the effect of that resource. There is no such striking argument on the other side, especially in a one city sample. One must assume that individual student items are more likely to be surrogates for school effects than vice versa.

There is no way to add the possible biases together to come with a resultant. However I have attempted to bias all estimates away from finding that school resources are associated with the outputs. Other studies have been similarly biased, but they have either not recognized or not stressed this bias.

In interpreting the variables, the prime rule will be a priori to suspect the label of the variable. All schools probably track, so what the "tracking" variable indicates is something about the form of the tracking, the nature of the principal who decides which way to answer the question, a student body so homogeneous that tracking is not feasible, or something else. The teacher test, often taken by the teachers together, never under professional supervision, may indicate degree of cooperation among teachers. The number of library volumes is presumably an estimate from purchases or the card file, and not an indication of the actual number available for students, nor of course of their quality, the physical ease of taking them out, the extent to which students are introduced to the library, encouraged to use it, etc. Each item has the same interpretation problem.*

---

*Even sex: 1 percent of the pupils in the SMSA sample from which our data is drawn gave no sex. I am not sure that all children who did not know their sex--or, more likely, could not read the question--did not mark it. There might be another 1 percent who randomly marked, and therefore 1/2 percent who are incorrectly coded by sex. This is not enough error, surely, to cause mistrust of that variable, but it is an example of how even the simplest item contains some error.

Statistical Techniques

The common technique applied to EOS and similar data is the single linear regression. A dependent variable is made a function of a set of independent variables, and fitted to the data to accord to the form:

$$Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n$$

The fit is made according to the principal of least squares, which minimizes the sum of the squares of the distance (in the Y direction) of the observations (data points) from the fitted n dimensional plane, where n is the number of independent variables. I assume that the reader is somewhat familiar with this technique. I will mention here that by minimizing the sum of squares, distant points receive a weight greater than the researcher would perhaps like to give them. They may be due to some different relationship--such as the desired output of the school, as discussed above--and should not be allowed to affect the estimates.

In using time series data, or other data with a limited number of observations, one often performs a residual analysis. War years, depression years (in time series), Alaska and Hawaii (in state observations), and other such identifiable anomalies from common patterns often cause the outlying points. Sometimes they are entered into the equation by creating special variables, sometimes they are excluded. In the case at hand, however, even if we did find one school or two with observations far from the rest, we would not know why this was so. If we did know, it would be because we had a variable describing those schools which had different values for

them, in which case inclusion of these variables should solve the problem. Hawaii and Alaska are often different from the other states because the meaning of "nonwhite" in, say, generating income, is different in these states from those in the 48 other states. Rationing in war years made the notion of "price" different from ordinary years, and the composition of output, demand for labor, etc., were unusual. A dummy variable in these cases corrects, from information external to the data, for a variable in the data which has different meanings over different observations.

Not knowing which Eastmet schools are which, not having any information about them individually outside of the data, a dummy variable for certain schools would only be a measure of ignorance in an effort to improve R^2 or other measures of goodness of fit. It might be an interesting investigating device, but not an explanatory device. On the other hand, as explained above, I did pick out some combinations which could lead to extreme observations, and defined variables accordingly as "interaction" variables. Their purpose is to bring extreme points into the general scatter, to reduce their influence on the resulting coefficients. The coefficients of these variables themselves are not interesting in this context.

There are a number of basic problems with the single linear regression. One is in its use: It does not, and cannot in simple application be a description of the production process within schools. A process should be described before being estimated, and I cannot believe that anyone would describe the schooling process as linear additive. Surely there are many interactions, many non-linear effects. One might be able to estimate them by linear regression on a reduced form model, deduced from a series of equa-

tions describing student preparedness, teacher ability, desire, etc. I have not seen such an attempt made.

What a linear regression on the variables might do is give coefficients which describe in some average way the effect of the independent variables on the dependent variables. The production function must be correct on the margin: it should predict what an increment of X_i will do to Y, holding the other X's constant.

The linear equations presented here and elsewhere in the education literature should not pretend to do this. They perform, rather, an averaging function. They designate what the linearly isolated effect of a particular variable seems to be; at least, what the linearly isolated association of an independent variable with the dependent variable is over a large number of observations. If there is a large coefficient for an inexpensive variable, the linear regression does not imply that more of that variable should be purchased. On the margin, that variable may have little effect.

A regression estimate fits the scatter of observations such that it is the variations in the observations which creates the hyperplane, not their levels. One problem in interpreting the results of average equations is in determining the effect of variation in inputs relative to their base. Explanations of variations in scores are not explanations in levels. Most students in our total sample scored 30 or better out of 50, and all students scored 20 or better in the verbal test. Most of the questions had five possible answers, so pure guessing would have produced a mean of at most 10

correct answers.* The worst student did twice as well as that, and the average student did three times as well. This does not indicate that schools, as opposed to home life, produced this level of achievement, but it is possible that, at least for some children, schools did perform this function. The variation in school resources may produce little of the variation in outcome, but the existence of schools might produce most of the test score level--or none. That is still an open question.

The single equation linear variable cannot account for the effect of attitudes on achievement, if attitudes are also the result of achievement. Simultaneous determination of attitudes and achievement requires a simultaneous equations model. The three equation model presented here is a variant of that employed by Henry M. Levin [27] in his paper for this conference, and I will not go into detail about it here. Student's grade aspiration and "fate control" attitude are assumed functions of the same variables as his achievement, and also a function of the achievement itself. Achievement is also a function of these attitudes. Three equations containing arguments which are dependent variables elsewhere in the system, must be estimated by two-stage least squares. The model is overdetermined a priori.

---

*The median might have been lower under guessing, since the random selection distribution is skewed about the expected value. The median was in fact higher than the mean. The expected mean under guessing would be below 10 if some students did not finish the test.

THE EQUATIONS AND THEIR IMPLICATIONS

In this part of the paper I will present regression equations derived from the Eastmet samples. In the first section, the ordinary least squares "average effect" equations will be presented and briefly discussed. Hazards of interpretation will be stressed. In the next section the equations for blacks and whites will be compared with each other to see if the same equations describe the average effect of the variables on different children. The concepts "resource" and "resourceness" will be defined in this section. In the third section, equations for whites will be compared by social class. Finally, a simultaneous equations system is presented and compared by race.

Average Effect Equations

The average effect equations, as explained above, are regression estimates of the average relationship between the dependent variables (verbal, reading, and math scores)---one at a time---and student background, school and teacher variables, with some attempt to account for points far from the resulting hyperplane. They are not attempts to describe the production process where the independent variables are "inputs," the dependent variables "outputs."* I do not feel constrained to choose a "best" equation for each output, but will present alternatives when no clear choice can be made.

With this kind of data, the crude measurements, the many possible in-

---

*Such independent determinations would violate the very concepts of joint production which they are supposed to estimate. In determining average effects, the production of other outputs is not accounted for, as it would be in joint production estimation, nor is an index of the joint product assumed to be maximized.

terpretations for any variable, this freedom is advantageous. For the white sample, two equations with verbal score as the dependent variable, three with reading score, and two with math score are presented in Table 1. For blacks, two verbal, two reading and one math equation are presented. The fewer black equations is a manifestation of the common finding that black behavior and outcomes are not as associated with typically measured variables as white behavior and outcomes. This is because we measure the wrong variables for blacks, their behavior is erratic with respect to the variables, or society's behavior is erratic with respect to the variables when dealing with blacks. All of these reasons doubtless hold. By measures of goodness of fit also, the black equations do not explain as much of the variation in scores as do the white equations.*

The different specifications of equations generally contain the same student variables, substitutions being made among teacher and school variables. Sex and age were included a priori, and possessions and size of family, the most significant variables in almost every equation,** were included essentially automatically. The other variables were experimented with, but the bias in selection was to include as many student variables as

---

*The black equations have similar standard errors to the white equations, but the black dependent variables have smaller variances. In terms of standard errors, then, the black equations are just as "good" as the white equations, and the difference in R^2 might be considered a difference in the data, not in the equations.

**I refer to the variable as "family size," though the question asked for number of people living in the same home, which may include non-family. Because of the lack of variation in the possessions index, as noted above, a great deal of social class variation is left to be accounted for by other variables.

Table 1

AVERAGE EFFECT EQUATIONS

WHITES

Independent Variable	$Verbal_1$	$Verbal_2$	$Reading_1$	$Reading_2$	$Reading_3$	$Math_1$	$Math_2$
Constant	17.5 (5.4)	-16.9 (3.1)	.6 (.2)	-5.8 (.8)	10.1 (1.9)	5.8 (3.8)	10.5 (2.9)
Background:							
Sex	.8 (1.3)	.8 (1.4)	1.8 (3.9)	1.8 (3.7)	1.8 (3.8)	-.7 (2.0)	-.8 (2.3)
Age 12+	-7.5 (5.0)	-7.4 (4.9)	-4.6 (4.0)	-4.5 (3.9)	-4.7 (4.1)	-3.0 (3.4)	-3.0 (3.5)
People at home	-.5 (3.3)	-.5 (3.4)	-.5 (3.8)	-.4 (3.7)	-.5 (3.9)	-.3 (2.3)	-.2 (2.3)
Possessions	1.4 (5.3)	1.3 (4.8)	1.0 (5.2)	1.0 (5.1)	1.0 (5.0)	.3 (2.2)	.3 (2.0)
Father's Education	.3 (2.7)	.3 (2.9)	.3 (2.9)	.3 (2.9)	.2 (2.6)	.3 (4.6)	.3 (4.4)
Kindergarten	2.1 (2.3)	2.2 (2.5)				1.3 (2.5)	1.2 (2.3)
Mother's I.D.						-.3 (1.5)	
Teacher:							
Test		.8 (3.2)		.5 (3.0)			
Experience	.4 (5.1)	.6 (8.2)	.4 (6.5)	.4 (6.4)	.1 (1.4)		
Tenure					-2.0 (1.2)		-2.6 (2.5)
Race Discrepancy	-2.7 (5.2)						-.3 (1.1)
Race Preference		1.5 (4.5)	1.1 (4.6)				
College Major							2.6 (2.0)

(Table 1 -- cont.)

astrict Central City — Average Effect Equations-whites-cont.

	$Verbal_1$	$Verbal_2$	$Reading_1$	$Reading_2$	$Reading_3$	$Math_1$	$Math_2$
chool							
Tracking				-.6 (1.9)	-.5 (1.4)	-.7 (2.7)	-.5 (1.9)
Library [1,000's]	.3 (2.1)					-.6 (2.5)	
Aud.-Cef.-Gym	.7 (3.2)			.3 (2.2)	.3 (1.9)	.3 (2.8)	.2 (2.0)
Acres						.6 (3.2)	
Upper Quars. x 10^2					6.6 (3.2)	5.9 (4.7)	5.9 (3.4)
nteractions:							
HiWh.-LoNW	-3.6 (2.2)					1.8 (2.2)	
HiSES-LORes-MidPr	2.7 (1.5)	5.8 (2.9)	2.3 (1.7)	4.5 (2.9)	1.8 (1.2)		
LoSES-LORes-MidPr	-3.6 (1.7)		-2.8 (1.7)		-3.5 (2.0)		
loSES-HiRes-LoPr	-7.4 (2.5)	-6.1 (2.1)					
^{2} [corrected]	.476	.470	.361	.353	.358	.333	.327
.E.	7.305	7.347	5.666	5.699	5.678	4.181	4.200
eter.	.140	.222	.492	.246	.033	.0178	.0666

T statistic below coefficient refers to coefficient = 0

(Table 1 -- cont.)

AVERAGE EFFECT EQUATIONS

BLACKS

independent variable	$Verbal_1$	$Verbal_2$	$Reading_1$	$Reading_2$	$Math_1$
Constant	1.4	2.7	2.8	5.1	7.2
	(.2)	(.3)	(.7)	(1.2)	(1.4)
Background:					
Sex	.7	.8	1.2	1.2	- .6
	(.9)	(1.0)	(2.0)	(2.1)	(1.5)
Age 12+	-3.7	-4.0	-2.8	-2.8	-2.0
	(2.6)	(2.8)	(2.8)	(2.8)	(3.0)
Age 9-	-4.2	-4.5			
	(1.2)	(1.3)			
Possessions	1.0	.9	.2	.3	.2
	(4.0)	(3.9)	(1.3)	(1.5)	(1.8)
People at home	- .4	- .4	- .3	- .3	
	(2.4)	(2.5)	(2.4)	(2.3)	
Father's ed.			.3		.2
			(2.8)		(2.5)
Father's occ.			.4	.4	.3
			(1.6)	(1.7)	(2.1)
Mother's ed.	.5	.5			
	(2.9)	(2.9)			
Mother's I.D.			- .5	- .5	
			(2.4)	(2.3)	
Kindergarten			1.2	1.2	
			(2.0)	(2.0)	
Teacher:					
Test score			.2		
			(1.1)		
Race				-2.0	
				(1.5)	
Parents' ed.	.8	.5	.3	.4	.2
	(3.1)	(1.9)	(1.8)	(2.8)	(1.9)
*Years of school					-1.5
					(1.5)
*Academic major			-7.1	-6.3	
			(2.1)	(2.1)	
Tenure	-1.1				
	(2.2)				

(Table 1 -- cont.)

Eastmet, Average Effect, Blacks, cont.

	$Verbal_1$	$Verbal_2$	$Reading_1$	$Reading_2$	$Math_1$
School:					
Adequate texts	2.5 (2.0)	2.1 (1.7)			
Tracking	-3.3 (1.9)	-1.6 (2.8)			
Building age		- .05 (1.8)			
Library [1000's]		.6 (1.6)			
*Assignment					.8 (2.1)
Interactions:					
*Hi SES-Hi Res-Mid Pr			7.6 (2.4)	6.7 (2.1)	
R^2 [corrected]	.152	.155	.132	.134	.074
S. E. of estimate	8.77	8.76	6.10	6.10	4.12
Determinant	.638	.343	.436	.638	.780

* indicates no prior hypothesis about the sign of the coefficient

possible. There is therefore a bias against the inclusion of school and teacher variables, so that there is no question about their appropriateness in these equations.

An example of the distinction between the average effect equations as presented here and production estimations can be drawn from the "kindergarten" variable, which appears positively wherever it is included. This does not indicate that sending a child to kindergarten will raise his sixth grade verbal score by over two points (if he is white). It indicates that white children who went to kindergarten scored, on the average, two points higher on this test than other white children with otherwise similar characteristics. The kindergarten may or may not have played a role in this higher score; it may indicate the concern of his parents, or the neighborhood in which the family lived, or their social milieu (in which it was understood that children went to kindergarten before elementary school). More importantly, no claim is made that the marginal effect of sending a child to kindergarten would be to add two points to his score.

The same distinction must be made for the teacher and school variables. For example, in the white verbal equations, the average discrepancy (per school) between the teacher's reported percentage of white students and desired percentage of white students is strongly associated with the score of the children if the teachers' average test scores are not in the equation. When we account for the test score, then not the discrepancy, but the absolute preference for whites has a strong effect. $Verbal_2$ surely does not mean that we should take teachers with mean test scores and consider those with strong preferences for white students to be

the better teachers. If we did, we might then send them to schools where there are many blacks, where their discrepancy is high, and where they are consequently bad teachers.* Or we might find that these characteristics alone make no difference at all, on the margin.

What these coefficients probably mean is one of two things: (1) Teachers are found to move towards their preferences, and white children who score higher tend to move toward whiter schools, so that teachers with strong preferences for whites tend to reduce their racial discrepancy and be associated with better students. (2) Some teacher attitude, which may find some expression in racial preference, affects their teaching.

No policy conclusion follows from either interpretation, though the latter indicates that an area of investigation might be revealing: the effect of teacher attitudes on student performance. Some work on this question is being done, as is well known.** Whether the attitudes involved are trainable or selectable, whether they can be applied to all children in a classroom or by definition select within a classroom; to these questions I have no answers. And of course, whether these equations imply an effect of these attitudes on children or on teacher location is also open to investigation.

Comparing Equations by Race

It is not clear why, if the school variables are to be interpreted as

---

*The correlation between teacher racial preference and discrepancy is -.60 in the white sample. For blacks, the correlation is only -.06. Teachers of whites, then, apparently are more free to follow their preferences in regard to race of their students than teachers of blacks.

**Rosenthal and Jacobson [34], but see their critics, for example Thorndike [41].

social class phenomena, the black equations look so different from the white equations. The teachers' parents' education is an important variable in the black equations, but does not enter the white equations. Academic majors (as opposed to education or physical education majors) are negatively associated with black reading scores, but positively associated with white math scores. Teacher experience does not help black children--at least not experience in the teachers blacks have--and the race variable in the black $Reading_2$ equation substitutes for the test score in the $Reading_1$ equation, whereas neither variable appears in two of the three white reading equations. This is a serious question, to which there are several possible answers.

Blacks, it might be argued, are not able to gain resources by improving their social class status. [See Michelson (31)]. The phenomenon of the teacher associating himself with better students does not occur among blacks, possibly because housing discrimination is so strong that upper class blacks do not have access to upper class schools. Thus the association of quality teachers with quality students, which is the explanation behind the equations--this argument continues--does not apply to blacks, and the school and teacher variables which appear in the white equations have no chance of appearing in the black equations.

This argument is more incorrect than correct, though it probably has some of both elements. In my recent publication cited above, I presented resource indexes derived from some of the equations of Table 1. "Resources" were defined as those school and teacher items which appeared in the equations. Black resources were therefore different from white resources, and black

resources were not distributed to blacks over social class, but white resources were so distributed among whites. However, whites' resources are also distributed by social class among blacks. There is an association between the average characteristics of schools and social class, when these characteristics are the variables entered in the white equations, whether white or black students are considered. These variables could have been associated with scores of blacks, which are also associated with social class (though not as strongly as white scores). But they were not. Instead, different variables appeared to be associated with black scores, and these variables were not distributed among blacks (or whites!) according to social class.*

A different argument, which accords with the allocation of these items, is that different things affect blacks and whites. That is, a characteristic of a teacher may be a resource for a white child--i.e., would increment his score--but not a resource for a black child. "Resource" then is not just anything which appears in a school, but an input which has an effect. What is a resource to whom is an empirical question. That question is not answered here, as I hope I have made clear. But it is raised here. It implies that the equations indicate some sort of causal relationship between something measured by some of the variables, and academic achievement. We do not know what that something is, because the variables are simply not that precise. But if there is any implication of causality in these equations, the implication should be stretched to in-

---

*More detail about these indexes will appear in future publications.

clude differential causality: different things affecting white and black children.*

At this point I have indicated that blacks and whites seem to respond to school variables differentially--i.e., that different variables have different resourceness to blacks than to whites. To indicate that this difference is statistically significant, I estimated the coefficients which blacks have for the white specifications, and the coefficients for whites with the black specifications. I then tested to see if their responses were the same. This is equivalent to asking if, with respect to these equations, blacks and whites could be said to be drawn from the same population.

In Table 2, the F test, degrees of freedom, and significance level are given for all of the average effect equations. The conception behind this statistical measure is simple. The regression equation is estimated to minimize the sum of squared residuals: if Y is the test score and $\hat{Y}$ is the equation's estimate of the test score, then define $e = Y - \hat{Y}$. Minimizing Σe^2 is the same as minimizing $\frac{\Sigma e^2}{k}$ where k is any constant. If k is the number of observations (actually the number of degrees of freedom), then this expression is essentially the average value of a residual. If the average squared residual value is lower for separate regressions on subsamples than for the sample as a whole--it can never be higher--then the equations which generated these average squared residuals must be different. This will almost always be true to some extent, but since the

---

*It should be pointed out that 33 of the 35 Eastmet city schools had both white and black pupils. The weighting of resources, but not access to some of the resources, varied by race.

TABLE 2

AVERAGE EFFECT EQUATIONS

F Test of Black-White Differences

	F	d.f.	Sig. %
WHITE EQUATIONS			
$Verbal_1$	4.07	14,1027	1%
$Verbal_2$	5.54	11,1033	1%
$Reading_1$	6.00	9,1037	1%
$Reading_2$	3.68	11,1033	1%
$Reading_3$	2.59	12,1031	5%
$Math_1$	3.99	13,1029	1%
$Math_2$	2.71	12,1031	5%
BLACK EQUATIONS			
$Verbal_1$	9.81	10,1035	1%
$Verbal_2$	9.69	11,1033	1%
$Reading_1$	4.73	12,1031	1%
$Reading_2$	4.14	12,1031	1%
$Math_1$	9.18	8,1039	1%

average residuals from samples from the same population follow a known probability distribution (Chi^2), so does their ratio (F), and we can calculate if the reduction in average residual squared is statistically significant, i.e., highly improbable under the assumption that the samples were truly from the same population.

There should be no question that the blacks and the whites form two distinctly different samples. In fact, since most of the variables are the same in black and white equations--the background variables--this is a somewhat weak test. Further investigation of the individual school coefficients verified that they are different for blacks and whites under similar equation specifications. The educational impact of the statistical difference in resourceness cannot be so easily tested. This will be discussed below. But the point should be clear: the school variables which seem to be resources are different for blacks and whites.

Social Class Differences in Resourceness

Whites were split into bottom quartile and the rest, and the same test was performed. The results appear in Table 3. Here, however, a few more words on the regression sample should be offered. In quartiling the sample by social class, the entire SMSA sample was included. Though I doubt the representativeness of the suburb sample, together with the city sample I had a much more representative picture of class variation. In selecting the central city to study, a bias towards lower classes was produced. That is, more than one fourth of the city sample is in the bottom quartile. However, in selecting the sample of children who had been in one

TABLE 3

AVERAGE EFFECT EQUATIONS

F Test of Bottom vs. Top Three Quartiles - Whites

	F	D.F.	Significance Level
$Verbal_1$	.97	14,569	n.s.
$Verbal_2$	1.20	11,575	n.s.
$Reading_1$	2.49	9,579	10%
$Reading_2$	2.16	11,575	10%
$Reading_3$	2.34	12,573	5%
$Math_1$	.048	13,571	n.s.
$Math_2$	.73	12,573	n.s.

school only, the opposite bias was produced. I had no a priori expectations as to the result, but in fact only 32 of the 597 whites in the regression sample (5.4%) were in the bottom quartile sample. They therefore could not represent the entire spectrum of schools, though bottom quartile children are probably not in every school anyway.

In interpreting the results of Table 3, the sample problem must be kept in mind. Difference in equations could be due to nonlinearities in the relationships, not differences in the sample, if the 32 children here represent extreme observations.

The Reading equations are apparently different. The coefficients were strikingly different for the bottom quartile regressions, including reversed signs for racial discrepancy and preference variables in all four equations in which these variables appear.

I partitioned the white sample again at the midpoint of the second to bottom quartile, creating a new lower sample with 88 (14.7 percent) observations. This adds more children to the bottom sample than were originally in it--and also undoubtedly adds more schools. Three of the four above-mentioned reversed signs reverted back to the signs from the total sample regressions. The R^2, which had been extremely high in the bottom quartile sample (above .7) went down (though were still high compared to the total sample R^2), and not one F test for difference proved significant. Once again, this could be a function of the particular schools involved. But it could also indicate that the bottom 5 percent of the regression sample children are very different in their reactions to school (and background) variables from the rest of the population,

whereas the bottom 15 percent are not. Whether this means the bottom quartile of the entire sample is different, I do not know, and cannot determine from this data. None of these results can do more than suggest what may be true. But I think this kind of result is striking in educational possibilities, if not in statistical definitiveness.

The Simultaneous Equations System

The schooling process is not as simple as a single linear regression would indicate. One way in which to conceive of it is as a system which simultaneously determines several outputs which affect each other. As long as each output has determinants which are unique to it, such a system can be estimated. I propose a three equation model in which verbal score, student attitude (control over his life), and his grade aspiration are three outputs.* His attitude and his grade aspiration are functions of his score, in that they give him a sense of reality about himself.** Neither his attitude nor his grade aspiration influences the other directly, though they both influence the verbal score, hence each other indirectly.

Most of the background variables are assumed to influence all three outcomes, though whether the parents are "real at home" or something else (say, an uncle or aunt for father or mother) is assumed to have no direct effect on verbal score. Of the school variables, the teacher attitude ques-

---

*The reader is reminded to refer to Levin (27) for details on simultaneous equation systems.

**The process by which this works is not clear, especially if grades do not correlate well with test scores, which often seems the case. If I had data on grades, the information system could be specified and the model would be greatly improved.

tion (preference for another school) is assumed to affect only attitudes and grade aspirations. Attitudes are affected by teacher turnover (principal's response to the question "What percentage of your teachers quit last year?") in that teachers in a school with high turnover might not pay as much attention to an individual as teachers in a low turnover school. Disruptions from turnover, and the other teacher and school characteristics (except teacher preference) all affect verbal score directly. The teacher's undergraduate institution was assumed to influence grade aspiration, though in this case (and this case only) the sign of the coefficient in the equation for whites was other than expected.

This three equation system looks like this:

$$V = b_1 A + c_1 G + \sum_i d_{1i} X_i$$

$$A = a_2 V + \sum_i d_{2i} X_i$$

$$G = a_3 V + \sum_i d_{3i} X_i$$

where V is verbal score, A is attitude, G is grade aspiration, X_i are the exogenous variables, and there is at least one $d_{1k} = 0$, $d_{2h} = 0$, and $d_{3j} = 0$, where $k \neq h \neq j$. In vector form, where Y is the output vector and X is the vector of exogenous variables,

$$Y = MY + NX$$

In this system, M is a 3 x 3 matrix, N is a 3 x 17 matrix, and Y and X are vectors with three and 17 cells. The solution is:

$$Y = (I - M)^{-1} NX$$

The structural equations are estimated by two-stage least squares, and are given in Table 4 for whites and blacks. The solution, or reduced form equations, is given in Table 5.

I have not performed any statistical tests on these equations. Nonetheless, looking at the differences by race, the impression is strong that these are not the same systems. The number of different signs is striking.

The specification was par ly a priori, partly experimental. It was, however, perfected on the white sample.* Thus I could have derived an optimal black system, and asked what the coefficients for whites were like in that system, analogous to the work in the previous section. For the purposes of this exposition, the work presented here should suffice.

---

*For this reason, T statistics are not given for the black coefficients.

TABLE 4a

STRUCTURAL EQUATIONS, WHITES

N = 597

	T S L S		
	Verbal	Student Attitude	Grade Aspiration
Verbal	---	.054 (1.97)	.067 (3.34)
Student's Attitude	2.391 (1.62)	---	---
Grade Aspiration	1.622 (1.63)	---	---
Background			
Sex	-.467 (.42)	.550 (3.08)	-.125 (.94)
Age - 12+	-5.026 (2.61)	.122 (.25)	-.284 (.79)
Family Size	-.080 (.29)	-.129 (2.49)	-.048 (1.27)
Possessions	.630 (1.41)	.151 (1.57)	.021 (.29)
Kindergarten	.969 (.77)	-.116 (.41)	.579 (2.78)
Mother ID	---	-.021 (.13)	-.219 (2.45)
Father ID	---	-.091 (1.34)	-.051 (1.01)
Father's Education	.066 (.33)	.084 (2.13)	.017 (.59)
Mother has job	---	-.293 (1.45)	.305 (2.04)
School			
Teacher Test Score	.246 (.96)	---	---
Teacher's Undergraduate Institution	6.457 (2.27)	---	-.349 (.80)

TABLE 4a

STRUCTURAL EQUATIONS, WHITES (Continued)

	T S L S		
	Verbal	Student Attitude	Grade Aspiration
Teacher's Experience	.637 (5.10)	---	---
Teacher's Preference for another school	---	-.147 (.37)	.701 (2.42)
Teacher Turnover	-.023 (.19)	-.048 (2.74)	---
Volumes per student	.380 (1.08)	---	---
Constant	-33.55	5.514	8.774
R^2	.364	.184	.254
S.E. of Estimate	8.144	2.163	1.603

TABLE 4b

STRUCTURAL EQUATIONS, BLACKS, WHITE SPECIFICATION

N = 458

	T S L S		
	Verbal	Student Attitude	Grade Aspiration
Verbal		.072	.059
Student's Attitude	3.33		
Grade Aspiration	.048#		
Background			
Sex	-.481	.199#	.551*
Age - 12+	-2.160#	-.210*	-.421
Family Size	-.395#	.032*	.019*
Possessions	.947	-.022*	.067#
Kindergarten	.253#	.017*	.793
Mother ID	---	-.089#	-.034#
Father ID	---	.050*	.085*
Father's Education	-.084*	.097	.098#
Mother has job	---	.001*	-.077*
School			
Teacher Test Score	.254	---	---
Teacher's Undergraduate Institution	-1.463*	---	.675*
Teacher's Experience	-.179*	---	---
Teacher's Preference for another school	---	-.136	.960
Teacher Turnover	-.016	-.025	---
Volumes per student	.076#	---	---
Constant	-8.578	5.326	5.833
R^2	.146	.082	.194
S.E. of Estimate	10.36	2.179	1.992

*Black and white coefficients differ in sign

#Value of black coefficient more than twice or less than one half

TABLE 5a

REDUCED FORM EQUATIONS, WHITES

	REDUCED FORM		
	Verbal	Student Attitude	Grade Aspiration
Background			
Sex	.846	.595	-.068
Age - 12+	-6.806	-.243	-.739
Family Size	-.613	-.162	-.089
Possessions	1.344	.223	.110
Kindergarten	2.135	-.002	.721
Mother ID	-.532	-.050	-.254
Father ID	-.395	-.112	-.078
Father's Education	.385	.104	-.043
Mother has job	-.270	-.308	.287
School			
Teacher Test Score	.323	.017	.022
Teacher's Undergraduate Institution	7.718	.414	.167
Teacher's Experience	.835	.045	.056
Teacher's Preference for another school	1.030	-.092	.770
Teacher Turnover	-.181	-.058	-.012
Volumes per student	.498	.027	.033
Constant	-8.030	5.084	8.237

TABLE 5b

REDUCED FORM EQUATIONS, BLACKS

	Verbal	Student Attitude	Grade Aspiration
Background			
Sex	.277#	.219#	.568*
Age - 12+	-3.808	-.485#	-.647
Family Size	- .382	.004*	-.003#
Possessions	1.159	.062#	.136
Kindergarten	.461#	.050*	.820
Mother ID	-.395	-.118#	-.057#
Father ID	.227*	.067*	.099*
Father's Education	.322	.120	.117#
Mother has job	-.002#	.0005*	-.077*
School			
Teacher Test Score	.336	.024	.020
Teacher's Undergraduate Institution	-1.891*	-.136*	.563#
Teacher's Experience	-.237*	-.017*	-.014*
Teacher's Preference for another school	-.540*	-.175	.928
Teacher turnover	-.133	-.035	-.008
Volumes per student	.101#	.007#	.006#
Constant	12.497	6.228	6.573

*Black and white coefficients differ in sign

#Value of black coefficient more than twice or less than one half of white coefficient.

INTERPRETATION OF STATISTICS AND BEYOND

Some school inputs might be resources to some children, not to others. But this "all or nothing" approach to resources probably does not describe most of the things which affect children. Nor, of course, does it adequately account for the output problem: that what is an important resource for one output may be less of a resource for another, and may even have a negative effect on some objectives of schooling.* It seems easy to me to use the word "resourceness" to indicate that children respond to an input, realizing that some inputs have more resourceness (for some outputs) than others. Those inputs which have no resourceness are not resources, just as materials vary in their fluidity and those which have none are not fluids.

There are a number of ways to determine how much is "a lot" in terms of resourceness. Those items which have no statistically significant resourceness were generally excluded from the equations.** Besides statistical significance, one should consider the concept of educational significance. For example, the teacher test score for the one black equation in which it appears, $Reading_1$, has a coefficient of .2. We could ask: how many points would a teacher have to gain on his test score

---

*Resources which induce discipline might stifle curiosity or inventiveness, for example.

**In the ordinary single regressions, large coefficients in the meaning given in the text below were considered if the T values were 1 or greater, even though not significant by conventional standards.

to raise the reading score one point, or one standard deviation.* Obviously 5 teacher points are required, on the average, to produce a point of reading score. The mean teacher test score for blacks is 22 points, and the highest possible is 30 points. Thus, as far as we can discriminate by this test, the best teacher would produce, on the average, 1.6 points more than the current average teacher. The difference between the average black and the average white reading score for the sample is 5.7 points.** Thus the experiment of putting the "best" teachers with the blacks reduces the black-white gap by 28 percent. On the other hand, calculating the black score if they had teachers with average test score equivalent to that of teachers of white children, 8.8 percent of the student score gap is closed. Both of these seem to be educationally significant.

On the other side, one might care more that these increases are 24 percent and 7.5 percent of a standard deviation, respectively, which might seem less significant. Another way to look at it is by asking how many whites score above the black mean, and how many whites would the black mean surpass under various assumptions. If the scores are normally distributed, then in the case where the means were equal, 50 percent of the

---

*I am not concerned with observed variation in teacher test score, because the observed variation may not represent the potential variation. However, this exercise comes dangerously close to using the equation for purposes it cannot perform, estimation of marginal effect.

**I am using here means of the samples containing 1599 blacks and 1727 whites. This is a reduction from 4505 students in Eastmet after elimination of those reporting no sex, those neither black nor white, and those with incomplete records (students but no teacher, for example). This sample includes the suburbs of Eastmet, which gives a broader range of scores than the city sample alone.

whites would score above the black (= white) mean. Taking the white standard deviation and maintaining the normality assumption, then 78.5 percent of the whites score above the average black. Under the most favorable assumption, teachers who score 30 points assigned to blacks, but white teachers staying as they are, then 71.4 percent of the whites would still be above the black mean. With equal teachers, 76.3 percent of the whites would still be above the black mean. That is, for each 1000 whites, 785 now score above the black reading mean (as opposed to 500 if blacks and whites were equal), and with "equal" teachers, the black mean would surpass only 22/whites; with the best teachers, the average black would surpass 71 more whites (or 49 more than with equal teachers). One might consider these numbers educationally insignificant.

more

I see no unique measure of educational significance. Much of the question about the effect of variables is, like many other educational questions, a social problem, not a scientific one. Do blacks care more about their mean score relative to whites, or the number of whites who score better? I do not pretend to know.

Implications for Teacher Training

To this point, no inferences have been drawn from the statistical study to questions of policy. Two major areas of concern here are: teacher training and resource allocation. For this conference, the stress will be on teacher training.

The equations do not indicate that "resourceness" is a trainable phenomenon. Nor, assuming that to some extent it is, are the implications for training clear in terms of the content of any program. I have often

thought that the Peace Corps and VISTA were excellent training for teaching, and several school districts have begun to think the same thing in the past few years. It does not seem to me to be necessarily true that school is a good place to train teachers.

Whatever the outputs desired, whatever the ways to train teachers to induce these outputs in children, what the foregoing does imply is that the structure of the training must respond to differences in the children who will be under the teacher's care. The concept that teacher resourceness differs by type of child I call "teacher specificity." Since different students will respond differently to different styles, attitudes, activities, language, strictness, etc., these properties of teacher activities should be investigated and directed to teachers who need them.

The concept of teacher resource being a function of the children being taught might lead one to conclude that segregated teaching was a preferred school structure. If this were so, one could still reject it, as I indicated at the beginning. But it leads no such place. There are two obvious reasons why teacher specificity does not imply segregation.

First, other children may well be resources in addition to teachers. Teacher resourceness is not the only item in the entire resource package. Again, we don't know to what extent other children influence any particular child--nor do we know which other children influence any one.* But in this ignorance, to structure the schools by teacher resourceness would be to assume that other children have no effect. Even if this were true, the fact of

---

*We do know that some children are generally recognized as class leaders, but that "outgroups" sometimes have their own leaders. We do not know the extent to which this leadership influences outcomes of schooling.

separation (and the inevitable invidious comparison) is believed to have a detrimental effect on some of the children. Thus ignorance of the resource effect of children on children should, if anything, lead to more heterogeneous classes.

Secondly, teacher specialization itself need not lead to separation of children because that specialization may be different for different outputs. By and large, some teachers are probably better with under-privileged children, others better with over-privileged children. To that extent, they may go to schools which are also characterized as under- or over-privileged. But some combination of resources may work best in a heterogeneous setting. That is, the specialization of some resources might be directed more at "mixed" children, whereas other resources might better be directed at one group or the other.

All of this is a land of mystery. Some teachers' talents are clearly in bringing diverse groups together, and other teachers are incapable of that. Some teach better with strict discipline, others with more freedom. Some have a conceptual approach to mathematics, some a mechanical approach. Some teachers will interpret Hamlet as weak, some will stress that he was tormented. Some are verbally oriented, communicate by words. Others prefer to play physical games, construct things. Some want to direct the class according to plan, some want to develop the sense of planning and conclusion seeking in children. Too much the search has been to differentiate between these characteristics in a search for the "right" ones. It seems strikingly obvious to me that the right teacher or method for some children may be wrong

for others.* Even for the same children, different approaches may work at different times. Teachers should be more prepared to specify their styles to the situations at hand, and administrators should be more prepared to select teachers for the students they will have.** This means we should learn more about appropriate ways to deal with children starting from a knowledge and acceptance of their present receptivity.

On Statistical Inference

Perhaps more mileage has been implied from the crude statistical estimation than can legitimately be claimed. The F test for sameness of regression coefficients is sensitive to the range of the observations and the linearity assumption of the regression. I explicitly stated that I do not assume that linearity holds, though one could define an "average" effect which is the linear fit. By stratifying on social class variables, then including correlates of social class in the equation, the likelihood of the fit being subject to nonlinearities is particularly severe.*** For

---

*Levin [25] gives an example which makes this point so clearly that conventional standards and measures appear ridiculous: "If black schools and white schools have the same number of teachers with the same preparation and experience, the two sets of schools are considered to be equal according to conventional criteria. Now, what if all of the teachers have white racist views?" Such views might not hinder, say, mathematics teaching in the white schools; but they might make serious teaching in black schools impossible.

**In the current school organization one could say this is done already: the better teachers, who might be able to adapt to the poorer students, nonetheless get the better students. The plea that administrators optimally assign teachers is empty within the current incentive structure. Optimum for whom?

***Note that I did not stratify by variables explicitly entered into the equation. Social class was defined by father's occupation, which is not used as a variable. However, social class is so highly correlated with possessions, family size, and father's occupation, that following a technically correct procedure is no salvation.

example, picture a circle of radius 10, centered at (0,0) on conventional Cartesian coordinates. Consider the upper half of the circle as the shape of the relationship being investigated. Suppose the data for the entire sample runs from -10 to 4. Then we will find a positive slope coefficient for the range of the observations. Suppose we split the sample: from -10 to 0, and from 0 to 4. Then we will have a negative slope for the upper sample, a positive slope for the lower sample, and a positive (but lower) slope for the pooled sample. The test might say that these were samples from different populations. The truth is that the calculated average effect in the first place was a function of the range of observation (for the slope would have been 0 if -10 to +10 had been observed), that the population fitted the true relationship perfectly, but the F test says these are most likely two different populations being sampled.

This sounds harsh, but it is important to demystify the notion that involved statistical models can, of themselves, confirm or deny hypotheses. That whole procedure is involved with the nature of the data, the range of the observations, the amount of knowledge external to the data, the complexities of the relationships and the simplicity of the equations, etc. I will propose here how the tests conducted above might be amplified upon. I plan to investigate another city in the EEO data. I will code that city's data the same way, and test whether the middle class whites in that city and in Eastmet can be said to derive from the same population.* If the two white populations react the same way to school variables, but the black populations do not; if the middle classes do, but the lower and possibly

---

*Actually, there are more powerful tests than the F test on squared residuals with which I intend to ask these questions.

upper classes do not; then the case will be quite a bit stronger. If all groups are unlike each other, then the test will say nothing, and one should feel dubious about the conclusions I am now drawing.

TEACHER SPECIFICITY AND SCHOOLS

Briefly, the argument of this paper has proceeded in this manner: Two methods of associating school resources with variation in cognitive outcomes (verbal, reading and mathematics tests) were presented. Single linear regression estimates were derived for a single city, Eastmet, on observations of children who had not changed elementary schools, stratified by race. A three equation system with simultaneous estimation was also offered on this sample. The equations were compared between the races, and the associations between school variables and outcomes were found to be different. Some difference also was suggested between bottom quartile whites and the rest of the whites. An interpretation offered was that those school characteristics which affect whites, particularly middle class whites, are different from those characteristics which affect blacks and lower class whites. This was not the only possible interpretation, and indication was given of research in process on this question.

Characteristics which are associated with outcome are called "resources," the amount of their "resourceness" to the different populations being indicated by the relative size of their coefficients. Teacher "specificity" then refers to the theory that certain characteristics have more resourceness for some children than for others. Since this concept is commonly accepted in the area of teaching exceptional children, an Appendix reviews some of the special education literature (that dealing with integrating exceptional children into the normal classroom).

I argued that these concepts could be applied to situations in which not "normalcy," but simply differences among children in response to simi-

lar characteristics was the issue. Unfortunately, the literature on special education is not convincing about the nature of the characteristics of special teachers. "Empirical proof of the validity of special preparation does not exist. . . . Proof must be forthcoming that there is more special about special education than the children assigned to these classes." [(38), pp. 245, 246] Nor, in comparative studies, were the characteristics of either the teachers or the students in the special and the regular classes examined. Conflicting findings indicate to me that there might be some powerful variables at work which need to be investigated.*

One such type of variable might be a trainable teacher characteristic. If the evidence that there are teacher characteristics which affect output is considered weak, then the argument for specificity of this effect is equally weak, and the implication that such a characteristic is trainable is weaker still. Thus this paper is a tentative dip of the foot into the pond. The temperature feels right, but I would prefer to know about the temperament of the fish before actually advocating that we swim.

I am nonetheless willing to ask what swimming in this pond would be like, if the fish proved friendly. For that reason, I suggested that teacher specificity did not necessarily lead to segregated education, although most elementary education is segregated, and teacher training and hiring might therefore take note of those characteristics which are most

---

*Yamamoto [44], for example, found creativity measures differing in teachers, and unrelated to their backgrounds. He could not associate this difference to output differences, but he noted that his output measures might not have been appropriate.

useful for the particular children which the teacher will have. Teacher certification by one set of standards is perverse if teacher specificity has any validity at all. A highly verbal teacher might be such a resource that he might not need to fulfill other requirements, such as college graduation. Or perhaps some children need more attention paid to them than a single teacher can produce in a day: several part time teachers might man one classroom. Perhaps some children learn best from "call and response" techniques, in which case a teacher with strong vocal chords and a room with sound proofing are resources.

These are just ideas. Some are being tested, others should be. Meanwhile, how ought schools to be structured? In the absence of answers, what do we do?

Inertia or Control?

The history of education, as any other public institution, is one of inertia. In the absence of information--though usually the impetus is a belief which may or may not hold true--a bureaucracy tends to make minimally disruptive decisions. And bureaucracy is the name of the education game. It takes an aroused public to stir the system, and the evidence presented here is not the kind to kindle the public spirit. I do not envision an enraged mob storming the educational portals, demanding "teacher specificity for all!"

Despairing of a revolution of the masses, I still plea for changes in the structure of decision making (a revolution by another name).* Spe-

---

*I will assume a public school system basically of the type we now have. Such schemes as "voucher plans," in which people purchase education from private schools, require more consideration than I can give them here.

cifically, at first, for principal-power. I would like to see each principal given a budget from which he could purchase resources, instead of being sent inputs (which may not be resources) from the central board. For example, some schools ordinarily cannot get substitutes. Under the present structure, they do not get the salary of the substitute spent in their school unless it is spent on a teacher. The principal, in effect, has a coupon from the board of education which is redeemable only in teacher services. No teacher, no redemption. All I am advocating is that the nature of this coupon be expanded: it should be able to purchase any educational service. A television set, perhaps; but that is not very imaginative, and given the nature of most television programs, not very educational. Perhaps art materials with which the students could decorate the teacher-less room.

I can lose the point by being too specific. The possibilities should not be limited to my imagination and inexperience. Nor should they be limited by our notion of principals as they are now. If most principals, unable to cope with such new responsibility, would make essentially the same decisions--hire the same teachers, purchase the same other inputs--as they do now, then what is lost? If some principals struck out into new forms of school organization, then what possible gains! Most importantly, the principal with the power to decide how his own school would operate would have to respond to the community, including the teachers. This has both the dangers of faddism and the possibilities of relevance about which we are all aware. At the moment, I am more impressed with the possibilities.*

---

*For a lucid discussion of the extreme public, the extreme private, and intermediate forms of school organization, see Levin [26]. Levin urges that "The time is ripe to experiment with at least one of these plans. . ." (p. 37). But why just one? A really daring experimental approach would outline the major differences in plans, and experiment with several so their outcomes could be directly compared.

cifically, at first, for principal-power. I would like to see each principal given a budget from which he could purchase resources, instead of being sent inputs (which may not be resources) from the central board. For example, some schools ordinarily cannot get substitutes. Under the present structure, they do not get the salary of the substitute spent in their school unless it is spent on a teacher. The principal, in effect, has a coupon from the board of education which is redeemable only in teacher services. No teacher, no redemption. All I am advocating is that the nature of this coupon be expanded: it should be able to purchase any educational service. A television set, perhaps; but that is not very imaginative, and given the nature of most television programs, not very educational. Perhaps art materials with which the students could decorate the teacher-less room.

I can lose the point by being too specific. The possibilities should not be limited to my imagination and inexperience. Nor should they be limited by our notion of principals as they are now. If most principals, unable to cope with such new responsibility, would make essentially the same decisions--hire the same teachers, purchase the same other inputs--as they do now, then what is lost? If some principals struck out into new forms of school organization, then what possible gains! Most importantly, the principal with the power to decide how his own school would operate would have to respond to the community, including the teachers. This has both the dangers of faddism and the possibilities of relevance about which we are all aware. At the moment, I am more impressed with the possibilities.*

---

*For a lucid discussion of the extreme public, the extreme private, and intermediate forms of school organization, see Levin [26]. Levin urges that "The time is ripe to experiment with at least one of these plans. . ." (p. 37). But why just one? A really daring experimental approach would outline the major differences in plans, and experiment with several so their outcomes could be directly compared.

Not just the ratio of teachers to other resources, but the type of teacher, should somehow be more a matter for local control, relating to the students. A principal might want to have one very expensive (but charismatic) teacher, and several community aides who are underpaid volunteers. Or he might want a teacher who is not acceptable to the school board, because that teacher has the specific talents needed in the school, but not the nominal qualifications. A principal might be restricted by his community from hiring unconventional teachers. But now he is restricted by his school board. And "unconventional" teachers is exactly what "teacher specificity" must mean. Eventually, if teachers appropriate to the situation are induced into schools, the conventions will change. Conventions are what schools of education transmit. So I contend that the place to start change is the public school, and the way to start is with principal control of his budget. Experimentation could take place within this context, and teacher specificity investigated. Then, with an idea of what kinds of things produce results for different kinds of children, teacher training can attempt to "produce" the kinds of teachers being called for.

Obviously such an idea as principal-power needs more exposition, more defense.* But so does the concept of teacher specificity. The two are somewhat tied together, though, in that the allocation decisions implied by

---

*This idea goes farther than that proposed by the First National City Bank, that "Title I and other monies could be put to good advantage by providing principals of disadvantaged schools with discretionary funds to use for whatever special purposes they themselves deem necessary to improve achievement levels in their schools" [(33), p. 33]. My suggestion is in regard to the entire school budget (except the capital budget, though there, too, the principal and his community should have a great deal more authority).

teacher specificity seem too difficult for large central control. A central board might act as a referral agency, taking "want ads" from principals, and "personals" from prospective teachers. But such decision making as I envision, based on the school needs, must be local. The point of this ending, then, is merely to indicate some of the implications of such a seemingly technical idea as the association of teacher resourceness with children's characteristics. If that concept seems reasonable, then perhaps the places it leads will seem more reasonable now than they once did. That would be a happy outcome of a long article, one as difficult for me to write, I assure you, as it has been for you to read.

APPENDIX

THE EXCEPTIONAL CHILD ANALOGY

Given the concept of the "normal child," to whom public schools address their attention, there must be the "exceptional child" who falls outside the range of ability described by "normal." Mackie estimates that 10 percent of the school age children are exceptional on the low end, and 2 percent on the high end. "A total of 35 percent of all exceptional children were enrolled in special education classes in 1966." [(29), p. 5.] But the distribution of aid to exceptional children is not uniform by type of exception. Thus 50 percent of the blind and deaf, 80 percent of the mentally retarded, but 12 percent of the emotionally disturbed and socially maladjusted are in special classes.

I cannot here go into detail about the problems of diagnosis of exception, or even the concept of "normal" itself--the dimensions of normality which may be missed by standard measures. In fact, the whole effort of this paper might be seen as directed against the concept of "normal" children. I will devote some space to outlining the literature about integrating exceptional children into normal classrooms. Teachers are trained in one of two ways: specialists who see only the exceptional child and his teacher, and ordinary teachers who accept exceptional children into their classrooms with some training on how to handle the situation.* The

---

*Stephens and Birch [39] outline three organization plans for dealing with special education of partially seeing students which probably applies to most exceptional children: full time special class, resource teacher, itinerant teacher. In the latter two cases, however, the child is placed in a regular classroom for most of his instruction. See also Fouracre [11] for school organizations.

point of this Appendix is to investigate the extent to which teacher specificity and integrated classrooms are in conflict. The analogy between the situation of the physically handicapped child and the variations which I find in the "normal" category is not exact, but may lead to some insight into the question.

Those resources which enable a blind or deaf child to be integrated into the classroom are presumably not directly applicable to the ordinary child. But the presence of the exceptional child may benefit the others, as well as himself.

> It has been found that the sighted children in the school not only gain some insight into the abilities of one blind person but that some less enthusiastic pupils are motivated to better achievement while learning with a blind companion. [(21), p. 133]

Though we might accept such a "finding" with skepticism, the process which could create it is obvious, and its verity is possible. Not the <u>presence</u> of exceptional children, but their <u>success and acceptance</u> by the teacher could produce such reactions.

> Because these [exceptional] children will eventually be required to achieve a satisfactory adjustment within a predominantly normal society, the experiences they have as children with this society are invaluable to them. Furthermore, normal children should be given an opportunity to understand, accept and adjust to children with exceptionalities. [(17), p. 3]

A resource to the exceptional child could produce a resource to the other children in the same simultaneous sense that a resource to grade aspiration produces verbal or reading score, though it is not directly associated with verbal or reading score, in the system presented above. The possibility that teachers can be trained to handle the special problems of the poor and culturally deprived is taken as a premise for most of this discussion, though there is no direct evidence supporting it.

Academic Achievement

The research on the success of integration of handicapped children is inconsistent. One study reports success, another, failure. O'Connor and Connor [32] report that children in special classes for the very hard of hearing (losses above 60 db) performed better than those integrated into regular classes, even after special preparation. Jones [21] found that visually handicapped children could be integrated; Fouracre [11] has investigated ways in which regular teachers could be trained to help the visually handicapped; and Leshin [24] and Berry [3] have separately stressed that such training must be given, because there are not enough specialists available. Edgerton implies that efforts to integrate mentally retarded may be misplaced:

> What I am suggesting is this: there is unquestionably some intellectual minimum below which no one can fall and yet claim competent membership in any society. We would all agree, I think, that no one whose IQ is 20 or 30 or 40 could become fully competent in any society. I am suggesting that the threshold between incompetence and competence in any society is actually closer to 60 or 70. [(10), p. 86]

Johnson's position [19] is much the same.

Sparks and Blackman, on the other hand, report for the educable mentally retarded (usually IQ 75-90), "children in regular classes almost invariably demonstrate academic achievement superior to that of special class children." [(38), p. 243.] However, they also report that most studies are characterized by a "lack of control of the teaching in the experimentation." [(38), p. 244.] Vacc [43] reports achievement gains for emotionally disturbed children were greater from special classes than integrated classes.

The parallel between teaching these specialized cases and teaching the disadvantaged has been made before. Tannenbaum notes that it is "entirely appropriate to canvass specialists in special education for some points of relevance between their unique expertise and the needs of the socially disadvantaged." [(40), p. 2.] Jordan, however, warns against such facile comparisons. He defines the concept "Disadvantaged Group," referring to "a particular, discernible physiological defect," [(22), p. 314] and offers several arguments why the problems of the Disadvantaged Group are different from those of the "disadvantaged."

Far be it from me to try to draw strong conclusions from such a

literature. But whether in special classes or in ordinary classes, "Teachers of atypical children require special training above that required for normal children." [(36), p. 81.] And if more children were seen as "atypical," then more special training would be necessary. Edmund W. Gordon [(13), p. 15] suggests that the failure of EEO to find association between teacher characteristics and student output might be due to the teachers' failure "to plan learning experiences that outweigh home influences." He suggests that one could train teachers toward that goal, but he offers no evidence that this is possible.

The EEO findings, of course, can be faulted on statistical grounds, but Gordon's point is still important.* He reviews the literature on differences between lower class and upper class children, concentrating on their motivation. He concludes that the values of the children are the same, but the feedback to middle and upper class children is more direct. They do not learn delayed gratification, in essence, but have immediate gratification. Perhaps teachers have to learn how to offer important rewards to lower class children, but do not have to do that for other children.** Whatever the answer, if little can be said about school organization from the literature on special education, at least this much seems true of teacher training: we do not know what differential skills are re-

---

*See, for example, Bowles and Levin [4], or the work in the body of this paper.

**I am frightened, however, by the report that monetary incentives have been taken into the classroom. Hamblin et al. [15] report that a market is set up in class, tokens given for "good" behavior, which are redeemable for candy. What is frightening is the idea that markets are a socially desirable method of accomplishing production. Markets may be efficient in terms of physical resources, but there is no evidence that they have anything but deleterious effects on attitudes towards other people.

quired to produce academic achievement in different types of children. And this ignorance must produce failure.

Social Outcomes

What can special education in integrated setting do for socialization? Thurstone's [42] 1959 study is most often cited as evidence that the educable mentally retarded tend to have more friends if they are in special classes than in integrated classes. Sparks and Blackman, who reported achievement gains for these children from integrated classes, report social gains from the special classes. Carroll, however, claims the opposite. "The current investigation supported the hypothesis that EMR children in a segregated setting would show less improvement in self concept than would EMR children in a partially integrated setting over a period of one academic year." [(6), p. 97.] Darrah reports that special classes for educable mentally retarded "do not produce more potentially constructive members of society." [(8), p. 523.]

Johnson and Kirk, studying social segregation, found mentally deficient children rejected by their classmates, but not directly "because they did not learn as fast as other children, because they did not read, or because they could not achieve in the academic areas. They rejected the mentally handicapped child because of his behaviorisms," such as teasing, cheating in games, and physical aggression. "These . . . can be interpreted as compensations for frustrations resulting from failure in school situations in which they cannot compete." [(20), p. 87.] Vacc found that emotionally disturbed children also tended to be rejected by their classmates, but he did not ask why [43]. He found that behavior gains (Behavior Rating Scale) were

greater for emotionally disturbed children (in matched samples) who had spent a year in special classes than those who had been in integrated classes. But no mention was made of the amount (or lack) of teacher training in the integrated classrooms. That is, this finding is consistent with my position that there is a teacher characteristic which is more a resource for emotionally disturbed children than for normal children. Presumably the teacher of the special classes in the study reported by Vacc had more of this resource, whether it be an attitude or training or whatever. If it is training, then his achievement and behavior results need not hold in the situation where the integrated class teacher has special training.

Rucker, Howe and Snider confirm that mentally retarded children are less acceptable socially to their classmates than normal children, this time in a junior high school sample. [35] They also test whether the social ratings of the retarded children would be higher in a non-academic class than in an academic. The differences, stratified by sex, actually went the other way. However, again the question "why?" was not asked. Since the "non-academic" class chosen for this test was physical education, the hypothesis of Johnson and Kirk that academic frustration leads the retarded child to physical aggression could easily explain the finding: where better than in physical education class can one be physically aggressive?

The Analogy Reconsidered

The literature on the retarded and disturbed child is even

less clear about the benefits of integration than that on the blind or deaf child. But several things do seem important. First, there seems to be a teacher characteristic which is a resource to these children in producing both affective and cognitive outcomes. Second, it is conceivable that the failure of integration is due to the failure of the teacher of the integrated class to have this resource. If this is true, and if, as in the case of the physically disabled child, integration seemed preferable to separation (except for some special classes), then whatever of this analogy is acceptable points clearly to more evaluation of what characteristics of teachers are necessary to integrate various children into one class. On the other hand, the basis of the analogy is just that only in special education is differential teacher training by type of child recognized. It is not clear that anything more can be drawn from such an analogy to the problem of different backgrounds among "normal" students. But it is an area worth investigating.

REFERENCES

1. Armor, David J. "School and Family Effects on Black and White Achievement: A Re-Examination of the USOE Data," prepared for On Equality of Educational Opportunity, edited by Frederick Mosteller and Daniel P. Moynihan, Random House: 1970, (forthcoming).

2. Berg, Ivan. "Rich Man's Qualifications for Poor Man's Jobs," Transaction, Vol. 6, March, 1969.

3. Berry, Charles Scott. "The Exceptional Child in Regular Classes," Journal of Exceptional Children, Vol. 3, 1936-37.

4. Bowles, Samuel S., and Levin, Henry M. "The Determinants of Scholastic Achievement: An Appraisal of Some Recent Findings," Journal of Human Resources, Vol. III, No. 1, Winter, 1968.

5. Callahan, Raymond E. Education and the Cult of Efficiency. Chicago: University of Chicago Press, 1962.

6. Carroll, Anne Welch. "The Effects of Segregated and Partially Integrated School Programs on Self Concept and Academic Achievement of Educable Mental Retardates." Journal for Exceptional Children, Vol. 34, October, 1967.

7. Coleman, James S., et al. Equality of Educational Opportunity. Washington, D. C.: U. S. Government Printing Office, 1966.

8. Darrah, Joan. "Diagnostic Practices and Special Classes for the Educable Mentally Retarded--A Layman's Critical View," Journal for Exceptional Children, Vol. 33, April, 1967.

9. Dreeben, Robert. On What Is Learned in School. Reading, Massachusetts: Addison-Wesley Publishing Company, 1968.

10. Edgarton, Robert B. "Anthropology and Mental Retardation: A Plea for the Comparative Study of Incompetence." Behavioral Research in Mental Retardation. Herbert J. Prehm, Leo A. Hamerlynck, James E. Crosson, Editors. Eugene, Oregon: School of Education, 1968.

11. Fouracre, Maurice H. Helping the Visually Handicapped Child in a Regular Class, Teachers College, Columbia University, 1957.

12. Gintis, Herbert. Alienation and Power: Towards A Radical Welfare Economics. Ph. D. Dissertation (Economics), Harvard University, 1969.

13. Gordon, Edmund W. "A View of the Target Population," in Special Education and Programs for Disadvantaged Children and Youth, Tannenbaum, Abraham (editor), Washington, D. C.: Council for Exceptional Children, NEA, 1968.

14. Greeley, Andrew M. and Peter H. Rossi. The Education of Catholic Americans, Chicago: Aldine, 1966.

15. Hamblin, Robert L., David Buckholdt, Donald Bushell, Desmond Ellis, Daniel Ferritor. "Changing the Game from 'Get the Teacher' to 'Learn'", Transaction, January, 1969.

16. Hanuskek, Eric A. The Education of Negroes and Whites, unpublished doctoral dissertation, Department of Economics, M.I.T., 1968.

17. Haring, Norris G., George G. Stern, and William M. Cruickshank. Attitudes of Educators toward Exceptional Children. Syracuse: Syracuse University Press, 1958.

18. Jencks, Christopher. "The Coleman Report and the Conventional Wisdom", prepared for On Equality of Educational Opportunity, edited by Frederick Mosteller and Daniel P. Moynihan and published by Random House, 1970 (forthcoming).

19. Johnson, G. Orville. "The Mentally Handicapped," Appendix B to Attitudes of Educators Toward Exceptional Children, by Haring, Stern, Cruickshank. Syracuse: Syracuse University Press, 1958.

20. _____, and Kirk, Samuel A. "Are Mentally-Handicapped Children Segregated in the Regular Grades?" Journal of Exceptional Children, Vol. 16-17, June, 1951.

21. Jones, John W. "Developments in Oregon's Program for Educating Blind Children," Journal of Exceptional Children, Vol. 19, 1952-53.

22. Jordan, Sidney. "The Disadvantaged Group: A Concept Applicable to the Handicapped." The Journal of Psychology, Vol. 55, April, 1963.

23. Katzman, Martin T. "Distribution and Production in a Big City Elementary School System," Yale Economic Essays, Vol. 8, No. 1, Spring, 1968.

24. Leshin, George. The Exceptional Child in the Regular Classroom, University of Arizona, 1967.

25. Levin, Henry M. "The Case for Community Control of the Schools," 1969, paper delivered at University of California School of Education.

26. _____. "The Failure of Public Schools and the Free Market Remedy," Urban Review, Vol. 2, No. 7, June 1968.

27. _____. "New Models of School Effectiveness." 1970, prepared for Office of Education, Bureau of Educational Personnel Development, Conference "How Do Teachers Make a Difference?"

28. _____. Recruiting Teachers, Charles E. Merrill (forthcoming, 1970).

29. Mackie, Romaine. Special Education in the U. S.: Statistics 1948-1966 (New York: Teachers College Press, Columbia University, 1969).

30. Mackler, Bernard. "Grouping in the Ghetto," Education and Urban Society, Vol. 2, No. 1, November, 1969, pp. 80-96.

31. Michelson, Stephan. "Resource Allocation: Reflections On the Law and the Data," Inequality in Education, Cambridge, Mass.: Harvard Center for Law and Education, 1969.

32. O'Connor, Clarence D., and Connor, Leo E. "Deaf Children in Regular Classrooms," Journal of Exceptional Children, Vol. 27, May, 1961.

33. Public Education in New York City, First National City Bank of New York, November, 1969.

34. Rosenthal, Robert, and Jacobson, Lenore. Pygmalion in the Classroom, New York: Holt, Rinehart and Winston, 1968.

35. Rucker, Chauncy N., Clifford E. Howe, Bill Snider. "The Participation of Retarded Children in Junior High Academic and Regular Classes," Journal of Exceptional Children, Vol. 35, No. 8, April, 1969.

36. Sengstock, Wayne L. "Contributions of Programs for the Mentally Retarded," in Special Education and Programs for Disadvantaged Children and Youth (Abraham J. Tannenbaum, editor). Washington, D. C.: Council for Exceptional Children, NEA, 1968.

37. Smith, Marshall S. "The Coleman Report: The Basic Findings Reconsidered," January 1, 1970, prepared for On Equality of Educational Opportunity, edited by Frederick Mosteller and Daniel P. Moynihan and published by Random House, 1970 (forthcoming).

38. Sparks, Howard L. and Blackman, Leonard S. "What Is Special About Special Education Revisited: the Mentally Retarded." Journal for Exceptional Children, January, 1965.

39. Stephens, Thomas M. and Birch, Jack W. "Merits of Special Class, Resource, and Itinerant Plans for Teaching Partially Seeing Children," Journal of Exceptional Children, Vol. 35, February, 1969.

40. Tannenbaum, Abraham J. (editor). Special Education and Programs for Disadvantaged Children and Youth. Washington, D. C.: Council for Exceptional Children, NEA, 1968.