DOCUMENT RESUME

ED 236 135 SP 023 158

AUTHOR Leiter, Jeffrey; Brown, James S.
TITLE Sources of Elementary School Grading.
PUB DATE 28 Aug 83
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; Elementary Education; *Elementary School Students; *Elementary School Teachers; Expectation; *Grading; Social Bias; *Student Behavior; *Student Evaluation; *Student Teacher Relationship; Teacher Attitudes; *Track System (Education)

ABSTRACT

The objective and subjective determinants that influence the way elementary school teachers grade students were explored. A longitudinal study was made of 213 students in 6 elementary schools as the students progressed from the first to the third grade. Possible determinants for assigning grades included student acquisition of valuable skills as reflected in standardized test scores, and classroom-specific achievement, based upon what the teacher considered important. Subjective factors in grading were seen to be teacher bias based upon individual reputations, the track level of the student, and the possibility that teacher expectations may be based on ethnic or gender prejudice. Student compliance with the teacher's preferred attitudes and behaviors was also considered. Factors shaping student achievement included the student's previous grades and general ability, track level, compliance and involvement, and race and gender. In analyzing the study findings, the question was asked: "Do teachers assign grades solely on the basis of merit or do other factors enter into grading decisions?" Student conformity to teacher preferences for certain attitudes and behaviors consistently appeared to be one of the strongest determinants of grading at this level. (JD)

Sources of Elementary School Grading*

Jeffrey Leiter and James S. Brown
Department of Sociology and Anthropology
North Carolina State University
Raleigh, North Carolina 27650

August 28, 1983

*We appreciate the suggestions of Carrie Knowles, Peter Marsden, Mary Haywood Metz, Joel Rosch, Robert Serow, and Matthew Zingraff. The data were gathered with funds from a Faculty Research and Professional Development Grant at North Carolina State University.

Sources of Elementary School Grading

This paper investigates the grades teachers give to their students. As the determinants of grades, it distinguishes student achievement from teacher misevaluation of achievement. Misevaluation may be due to the potentially biasing effects of student reputation, tracking, student deportment in class, and student race and gender. The paper further examines the impact of grades on achievement. To these ends, the analysis investigates the causal processes that relate reading and mathematics grades and achievement for second and third graders.

Grades are important because of the signals they convey to the student whom the grades describe. As such, grades may encourage the underconfident student, discourage the overconfident student, spur the underachieving student to extra effort, or lull the successful student into complacency (Davidson and Lang, 1960; Evans, 1976; Kirschenbaum et al., 1971). Beyond the signals grades convey to the student him/herself, they portray the student to others who may take the grades as valid evaluations. Among these consumers of grades are other students, other teachers (including those who will instruct the student in the future), school administrators (including those charged with admitting the student to selective schools and programs), and employers. Jencks et al. (1979) found that grades mediate a substantial part of the impact of academic ability on the number of years of schooling attained. This is likely due to a combination of the signalling and evaluative effects of grades.

DETERMINANTS OF GRADING

We conceive of grades as a teacher behavior or output and ask what the determinants of this teacher behavior are. Possibilities to be investigated include:

(1) Student achievement. Teachers sometimes claim they take student effort or progress into account in assigning grades, especially in elementary grades (Terwilliger, 1966). Most often, however, grades officially are meant to evaluate student performances relative to each other or to some ideal (Dreeben, 1968; Terwilliger, 1966; Waller, 1932). To the extent that the grades teachers assign reflect student achievement, they can be seen as meritocratic evaluations. Clifton (1981), Pedulla et al. (1980), and Williams (1976) found grading to be substantially meritocratic. Achievement in school has two components:

(a) Widely valued achievement. Most schools try to teach a body of basic "cognitive skills" that includes "the ability to manipulate words and numbers, assimilate information, [and] make logical inferences" (Jencks et al., 1972:53). These are the skills generally measured by standardized tests. While teachers may differ in the extent to which they value the development of such skills in their students, school systems impose requirements for instruction in these skills directly via curriculum guides and indirectly by standardized testing. Even teachers with their own classroom agendas feel encumbered to prepare their students for standardized tests (Kohl, 1967). Grades teachers give may reflect student acquisition of skills defined as valuable throughout the school system or in educational circles more generally.

(b) Classroom-specific achievement. A teacher may also award grades on the basis of achievements that the specific teacher considers particularly important, for example, creativity in writing, ingenuity in problem solving, neatness with pencil and paper, or fluid verbal delivery. These are not skills that standardized tests measure, but they represent achievements nonetheless and, as such, their assessment by grades can be considered meritocratic. (Of course, the student may be hard put at first to figure out "what the teacher wants.")

(2) Bias. Grades, however, may reflect factors other than the widely valued and classroom-specific achievements they claim to measure. The deviation of grades from an accurate assessment of achievement amounts to biased or non-meritocratic evaluation. Factors that influence grades after taking account of the effect of achievement on grades constitute determinants of such bias. Clifton (1981) and Williams found grades to be "affected markedly by the cognitive and normative expectations that teachers have for students" (1976:233). Expectations may distort the evaluation of homework and examinations (Finn, 1972) or even lead teachers to enter grades on report cards that are not justified by classroom performance. There are a number of possible sources of such expectations and, hence, of non-meritocratic evaluation:

(a) Individual Reputations. Informal conversation among teachers frequently spreads student reputations (Boocock, 1980). In addition, teachers are usually free to examine students' grades and aptitude and achievement test scores from previous years. Expectations are shaped to the extent that teachers take advantage of these opportunities and remember these previous grades and scores (Brophy and Good, 1974). Moreover, teachers may try to make their grades conform to previous grades for fear of needing to justify new evaluations (Schlechty, 1976).

(b) Track level. Beyond students' individual reputations, classrooms whose compositions are determined by alleged measures of previous achievement, including test scores, grades, or recommendations, may carry group reputations that become attached to students in the class. Furthermore, teachers may use grades to justify track placement post hoc. In a study of one high school, Rosenbaum (1978) found high grades to be harder to obtain in the lower tracks. In a study of eight high schools, however, Alexander et al. (1978) found little track effect on grades.

(c) Prejudice. Teachers may carry racial, ethnic, or gender prejudices that influence the grades they give. These prejudices may take the form of belief in differences in ability by race, ethnicity, or gender, either generally or for particular subjects. In studies of high school students, neither Williams (1976) nor Clifton (1981) found race or ethnicity effects on teacher expectations. They point out, however, that demographic variables may have greater direct effects on teacher expectations in younger grades, as Doyle et al. (1972) and Harvey and Slatin (1975) found in their studies of elementary school teachers. In older grades, race, ethnicity, and gender may make their effects felt indirectly via their effects on student performance and attitudes, which in turn may influence teacher expectations.

(d) Student compliance with teacher's preferred attitudes and behaviors. Teachers may give lower grades to students who challenge discipline standards, who question commonly held viewpoints, who do not appear to be interested or involved in activities the teacher organizes, or who through frequent absences seem to betray a lack of commitment to school (Bowles and Gintis 1976; Brophy and Good 1974; Gravenburg and Collins 1976). Here, teachers would be awarding grades on the basis of student attitudes or behaviors that do not bear on learning. Teachers probably differ in their preferences for such non-cognitive attitudes and behaviors. We are speaking, therefore, of classroom-specific non-cognitive attitudes and behaviors.

Determinants of Grading via Achievement

So far, we have specified sources of the grades teachers give to students. These grades have consequences, chiefly for subsequent achievement, that then shape later grades through meritocratic evaluation. A complete understanding of the determination of grading requires, therefore, some consideration of the determination of achievement, both widely valued and classroom-specific. Factors shaping achievement to be investigated here include:

(1) Previous achievement. Achievement may theoretically be traced back to ability, but the two can rarely be measured separately (Jencks et al., 1972). Previous achievement is a determinant of current achievement both because it subsumes ability in the subject and because of the cumulative character of much cognitive learning.

(2) General ability. Certain abilities or skills that contribute to current achievement in a specific subject can be distinguished from previous achievement in that subject. These are abilities that facilitate achievement in academic subjects more generally. Insofar as achievement is measured by tests, general test taking skills (Jencks et al., 1972) would fall into this category of contributors to achievement.

(3) Previous grades. As suggested earlier, grades give signals to students. Aside from ability and actual prior accomplishments, these signals may motivate or discourage effort and achievement (Bloom, 1976; Dreeben, 1968; Maehr, 1976; Salili et al., 1976). Moreover, to the extent that previous grades create teacher expectations, they may cause teachers to differentiate among students with regard to quantity of interaction, quality of reinforcements, difficulty of questions, and time allowed for answers to questions (Brophy and Good, 1974). Through these and other process variables, teacher expectations are realized as student performances.

(4) Track level. Assignment to a particular class on the basis of some evaluation or measurement of ability or previous achievement may influence current achievement for two reasons. First, such assignment signals, as do grades, the expectations held by others. Second, learning opportunities may

vary by track if effective teachers, classmates with special qualities, and instructional resources are differentially allocated, if instruction is differentially organized, or if the content of instruction differs by track (Sorenson, 1970). Several studies have found higher achievement in higher tracks, controlling for previous achievement (e.g., Alexander et al., 1978; Heyns, 1974; Leiter, 1983).

(5) Student compliance and involvement. Wiley (1976) and Stallings (1980) have found "time on task" to be an important determinant of achievement. To the extent that misbehavior, lack of concentration and involvement, and absenteeism diminish time on task, achievement may suffer.

(6) Race and gender. We do not expect large direct effects of race and gender on achievement. Any race and gender differences in ability (see, for example, Hunt, 1961; Jencks, 1980; Jensen, 1973; Maccoby and Jacklin, 1974) are subsumed in the earliest measure of achievement. Race and gender may affect subsequent achievement by influencing grades: race and gender may inject non-meritocratic expectations and bias into grades; the grades in turn may both shape teacher expectations, and, hence, learning opportunities for students and, also, signal students about teacher expectations, influencing their motivation to achieve. Any remaining direct effect of race or gender on achievement likely reflects differential allocations of resources for learning by administrators rather than by teachers.

DATA AND VARIABLES

The data for this paper describe one cohort of students in a rural consolidated district in North Carolina. This cohort is defined as students who passed from first grade in 1977-78, through second grade in 1978-79, into third grade in 1979-80. Students assigned to special education classes have been excluded. Consideration of only students with complete data on the variables to be analyzed left 213 students in the six elementary schools of the district for the analysis. The district strictly adheres to the practice of grouping its elementary school students by achievement into self-contained classrooms. Within school and grade, classroom assignments are based on reading achievement tests scores from the previous May.

All the data were taken from students' individual records maintained in the district office. The student's teacher (each class had one teacher for all its academic subjects) recorded one grade for reading and one for mathematics at the end of the school year without knowledge of results from the end of the year achievement tests. The single grade entered in the district's records summarized report card grades from the entire school year. The grading scale ranges along a fourteen point scale from F to A+. The widely valued portion of student achievement in reading and mathematics was measured by raw scores, grade equivalents, or mastery levels on achievement tests administered each May throughout the district.[1] General abilities not specific to one subject, including test taking skills are indicated by the test score in the other subject. For example, a significant effect of the reading test score on the mathematics test score suggests the contribution to widely valued mathematics achievement of widely valued skills not specific to mathematics. The track level in reading and in mathematics is the class mean for the achievement test scores in reading and mathematics, respectively, from the previous May's testing. Student's race, gender, and days absent were available from district records. Days absent serves as a partial indicator of time on task (time off task but in school is unmeasured). A direct measure is unavailable for the teacher's perceptions of student conformity to non-cognitive attitudinal and behavioral patterns the teacher prefers. On the central office records, the teachers recorded no grade for deportment, nor any indication of behavior problems. A proxy measure is available, however. The grade given by the same teacher in the other subject may be used for this purpose when its effect is estimated net of the effects of current and previous widely valued achievement, the previous grade, race, gender, track level, and absenteeism. So estimated, the grade given in the other subject stands for the teacher's generalized (i.e., non-subject specific) and non-cognitive (given controls for previous grade and previous and current test scores in the subject) assessment of the student.

The operationalization of achievement is particularly crucial to the analysis and subsequent interpretation. Widely valued achievement is measured by tests administered throughout the district. This is a narrower use than the common practice of trusting standardized achievements tests as measures of achievement per se and residual gains on such tests as measures of teacher and school effectiveness (Veldman and Brophy, 1974). The district in question uses the tests in just this way, taking them as the sole measure of achievement on the basis of which to group students into classes. Classroom-specific achievement is captured by the grades teachers award. These grades, however, embody widely valued achievement and any non-meritocratic evaluations, as well. Isolating the impact of classroom-specific achievement on grades, therefore, requires partialling out these other effects. The effect of widely valued achievement can easily be partialled by controlling for test score. Non-meritocratic effects are harder to control. The data specify race, gender, and track level, whose distinct effects on grading can be identified. Beyond these direct measures, distinguishing non-meritocratic effects from meritocratic effects of classroom-specific achievement depends on using the grade given in the other subject as a proxy for student conformity to the teacher's preferred attitude and behavior patterns. With controls for previous grade and current and previous test scores, this interpretation seems reasonable. Residual variation in grade received is, thus, likely due to classroom-specific, subject-specific achievement.

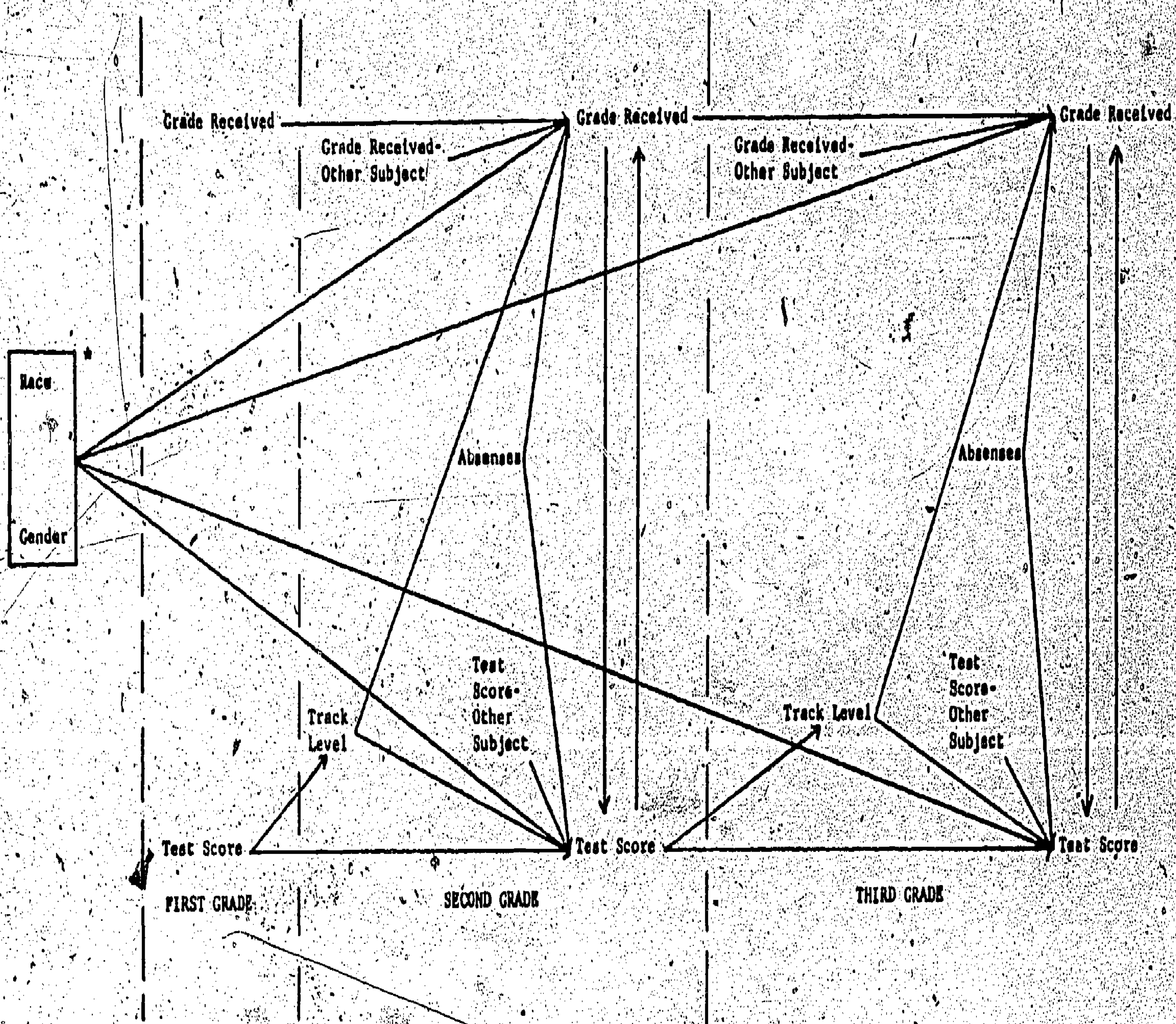
ANALYSIS PLAN

Figure 1 pictures the causal relations to be estimated. The model includes grades and test scores, their concurrent and lagged reciprocal relationships, and their determination by race, gender, track level, absenteeism, and grades and test scores in the other subject. While the causal expectations already explained need not be repeated, the reciprocal causal paths between grade and test score in the same year require further comment. The effect of test score on grade is simply the extent to which the grade embodies widely valued achievements, measured by the standardized test. The effect of grade on test score is more complex, actually two-fold. First, the test score embodies achievements, some of which the teacher has included in the grade. Beyond this, however, the grades students receive all year, which are then summarized for the office file, may have a signalling and/or motivating effect on the acquisition of widely valued skills, measured at the end of the year on the standardized test. Through the signalling/motivating effect of grades on widely valued achievement, measured by standardized tests, and the embodiment of the latter in grades, grades in year y may indirectly affect grades in year y. This seemingly anomalous possibility is provided for by the statistical treatment to be described shortly.

(Figure 1 about here.)

The longitudinal analysis suggested by Figure 1 is a decomposition of associations between variables into direct causal effects, indirect causal effects, and non-causal (i.e., unanalyzed or spurious) components. The sum of these three is the total relationship (i.e., zero-order slope) implied by the model as specified. This total can be compared to the observed zero-order

Figure 1. Causal Model



*Block presentation is for simplicity only. Each variable has effect designated by arrow from block.

relationship to assess the adequacy of the specification. The matrix approach outlined by Fox (1980) provides a convenient method for this decomposition.

Reciprocal causation makes ordinary least squares estimation inappropriate. We have used two-stage least squares estimation, instead. The grades and test scores in the other subject have been used as instrumental variables to overcome identification problems and to obtain predicted values for the reciprocally related endogenous variables. These predicted values were used in the second stage to estimate the causal paths suggested by Figure 1. Fox's (1980) method includes provisions for decomposing relationships in non-recursive models such as ours.

An instrumental variable must have a strong net relationship with the variable for which it is used to obtain a predicted value. As the tables will show, this requirement is well fulfilled here. Moreover, the instrumental variable may have no direct effect on the variable with which the variable for which it is an instrument is reciprocally related (Heise, 1975). Since model estimation depends on this assumption, this second requirement can only be justified through theory and logic (Asher, 1976). In the present study, we must argue that grades in one subject do not affect test scores in the other, net of other causes, and that test scores in one subject do not affect grades in the other, net of other causes. The first statement translates to the assertion that the teacher's perception of student conformity to preferred non-cognitive patterns does not have a net effect on the achievement of widely valued skills. A zero-order relationship is undeniable: the extent to which student and teacher personalities mesh almost certainly affeects learning of all sorts. The student, however, learns the teacher's evaluation of him or her through grades given in the subject in question. Hence the personality mesh (grade in other subject) has its effect on widely valued learning (test score) through the grade given in that subject. This fits the requirement for an instrumental vairable precisely.

The second statement that must be justified theoretically translates into the claim that general abilities for widely valued achievement have no direct effect on grades received. Again, the zero-order relationship is clear, because in all probability teachers use grades in part to evaluate widely valued achievement. This relationship, however, is captured in the effect of the test score in the subject in question on the grade given the same year in that subject. The effect of generalized abilities for widely valued achievement (test score in other subject) on the grade given is, thus, mediated by widely valued achievement (test score). This again fits the requirement of instrumental variables.

The decompositions are computed separately for the second and third ...es. Since the estimation yielded standardized coefficients to input into Fox's algorithm,[2] the effects can be interpreted and compared as one would path coefficients. Indirect effects sum all indirect paths from cause to outcome variable, but only those indirect paths whose separate components are all substantively significant will contribute notably to the overall indirect effect. This will make interpreting indirect effects much easier. Among the indirect effects are seemingly endless "reverberations" between reciprocally related variables. Even if the structural coefficients on these paths are fairly large, however, the reverberation quickly dies away after a single cycle.

ANALYSIS

Reading

Table 1 presents the decomposition of effects on students' reading grades and te (the parallel analysis for mathematics follows in Table 2).

(Table 1 about here.)

Table 1. Decomposition of Standardized Effects on Grade Two and Three Reading Grades and Test Scores (N=213).

Legend for Columns
(1) Direct effect
(2) Indirect effect
(3) Non-causal component
(4) Model implied zero-order standardized slope
(5) Observed zero-order standardized slope (i.e., zero-order correlation)

	GRADE TWO										GRADE THREE									
	Grade Received					Test Score					Grade Received					Test Score				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Ascriptive Sources																				
Race (1=nonwhite, 2=white)	.09	-.01	.32	.40	.39	.03	.01	.33	.37	.38	.00	.04	.24	.28	.29	.14*	.01	.29	.44	.45
Gender (1=male, 2=female)	.01	-.01	.17	.17	.17	.05	.00	.13	.18	.18	.04	.01	.13	.19	.19	.03	.03	.12	.18	.17*
Grade One Sources																				
Reading grade received	.26*	-.02	.42	.66	.66	.08	.03	.48	.61	.61		.03	.51	.54	.59		.01	.52	.53	.58
Reading test score	.35*	-.10	.43	.68	.67	.36*	.08	.32	.76	.76		.08	.51	.59	.60		.15	.50	.65	.71
Grade Two Sources																				
Reading grade received		-.02	1.03	1.01	1.00	.12	.00	.60	.72	.74	.10	-.01	.53	.62	.68	-.16*	.06	.60	.59	.61
Mathematics grade recd.	.51*	-.01	.27	.77	.77		.06	.56	.62	.64		.05	.57	.62	.60		-.05	.65	.61	.54
Reading test score	-.16	.00	.88	.72	.74		-.02	1.00	.98	1.00	.09	.06	.48	.61	.65	.39*	.04	.26	.70	.73
Mathematics test score		-.05	.64	.59	.58	.3[illegible]*	-.01	.37	.72	.73		.05	.51	.56	.57		.15	.51	.66	.66
Track level	-.04	-.01	.53	.48	.50	.07	-.01	.53	.59	.64		.01	.39	.40	.40		.03	.47	.50	.59
Absences	.03	-.01	-.07	-.05	-.06	.04	.00	-.05	-.01	.01		.01	-.09	-.08	-.02		.01	-.01	.00	-.03
Grade Three Sources																				
Reading grade received			.62	.62	.68			.61	.61	.65		.03	.93	.96	1.00	.19*	.01	.45	.64	.67
Mathematics grade recd.			.60	.60	.65			.57	.57	.59	.59*	.02	.19	.80	.81		.12	.50	.61	.61
Reading test score			.59	.59	.61			.70	.70	.73	.16	.01	.48	.64	.67		.03	.95	.98	1.00
Mathematics test score			.65	.65	.69			.67	.67	.75		.07	.58	.66	.69	.44*	.01	.32	.70	.78
Track level			.60	.60	.62			.85	.85	.88	.01	.00	.53	.54	.56	.00	.00	.61	.61	.55
Absences			.04	.04	-.03			.03	.03	-.05	.02	.00	-.09	-.06	-.07	.02	.00	.00	.03	.01
Variance Explained	.68*					.74*					.72*					.70*				

*Significant at the .05 level.

The most important of the many findings in Table 1 can be summarized as follows:

(1) In neither second nor third grade is widely valued achievement, as measured by end-of-the-year standardized test scores, an important determinant of the grades teachers give. Indeed, in second grade, the grade received is weakly negative related to the test score. Reading grades, thus, do not reflect this important element of merit net of other causes.

(2) The grade received in first grade shapes the grade assigned in the second grade. This may reflect a meritocratic continuity of classroom-specific achievements; a non-meritocratic contamination of grading by student reputations transmitted by the earlier grades; or the shaping of teacher expectations which in turn shape student achievement. The absence of a similar effect on second grade test scores argues against the third of these possibilities.

(3) The first grade test score has a fairly strong effect on the grade assigned in the second grade. Since this effect is net of the second grade test score, it can be explained with some confidence as a reputational or expectation effect. We cannot say to what extent the effect simply reflects teachers bringing grades in line with earlier test scores and to what extent it shows that teachers fit their expectations and the learning opportunities they offer students to these earlier test scores. In either case, the effect is much reduced the next year.

(4) By far the strongest force shaping grading in both years is the teacher's perception of student conformity with the teacher's preferred attitude and behavior patterns. This effect is represented by the coefficients for mathematics grades received. The strength of this non-meritocratic effect is particularly noteworthy given the many other variables

whose effects on grades assigned have already been considered before computing these coefficients.

(5) Non-meritocratic effects of race, gender, and track level on grading are absent. Likewise, time on task (absenses) does not affect the grades students are assigned.

(6) The determinants of widely valued achievement would be more important for grading if test scores had a larger effect on grade assignment. We should note these effects, however, for what they show about the test results. Notable among these determinants are:

(a) The previous year's test score, indicating stability and cumulation in widely valued achievement;

(b) The grade received the same year, but only weakly and less in the second than the third grade. This again shows that grades and standardized tests do not measure the same skills very much. The weak effect may even show how little report card grades given during the year motivate or demotivate student acquisition of widely valued skills; nor is the motivational effect of grades one year on test scores the next any larger, though in one case the direction is reversed;

(c) Race, but not gender nor track level, and then only weakly in the third grade.

(d) General, widely valued abilities and skills. Measured by mathematics test scores, this factor has at least as strong an effect on the widely valued part of reading achieement as any other determinant. It persists despite numerous controls.

Interpretation of these results for reading (the same will be true for mathematics) is eased by the absense of important indirect effects. We were able to concentrate on the direct effects. Note also that the model appears well specified to judge by the match between model implied and observed zero-order slopes. This, too, will be true for the mathematics results.

Mathematics

Table 2 presents the parallel analysis for the determination of mathematics grades. In many important aspects, the findings are quite similar to those for reading grades. Again, previous grades have a significant net effect on grades assigned, indicating some combination of continuity of classroom-specific skill attainment, reputational, and expectation effects. Unlike reading grades, this effect continues for mathematics grades through the third grade. Moreover, as with reading grades, mathematics grades are strongly influenced by student conformity to the teacher's preferred attitude and behavior patterns, as indicated by the large coefficients for same year reading grades. Such non-meritocratic effects on grading do not extend, however, to race and gender, which again have no effect on grading. Nor does time on task (absenses) influence the grades students receive.

(Table 2 about here.)

We should note three differences from the analysis of reading grades. First, mathematics more consistently than reading grades reflect the widely valued achievement measured by statndardized tests, an important meritocratic element of mathematics grading. The determinants of widely valued achievement, therefore, are more important for mathematics than for reading grades because of their indirect effects. These include: (1) general widely valued skills (reading test scores) in both grades; (2) motivational or signalling effects of the previous year's grade, especially for the third grade test score; (3) a weak continuity of widely valued mathematics achievement, but only from second to third grade; (4) for the third grade test score, the third grade grade, probably marking overlap in content of these two evaluations; and (5) race, but only weakly in the second grade.

Decomposition of Standardized Effects on Grades Two and Three Mathematics Grades and Test Scores (N=213).

r Columns
ect effect (3) Non-causal component (5) Observed zero-order standardized slope,
irect effect (4) Model implied zero-order standardized slope (i.e., zero-order correlation)

GRADE TWO

	Grade Received					Test Score				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
e Sources										
onwhite, 2=white)	.09	.03	.29	.40	.40	.15*	-.01	.27	.41	.40
=male, 2=female)	.04	-.01	.14	.17	.16	-.05	.00	.13	.08	.06
Sources										
ics grade recd.	.14*	.02	.43	.59	.58	.13	-.01	.38	.49	.50
ics test score	.08	.00	.45	.53	.53	.06	.00	.44	.51	.52
Sources										
ics grade recd.		-.02	1.05	1.03	1.00	-.09	.00	.81	.72	.60
grade received	.51*	-.01	.28	.78	.77		-.05	.64	.59	.58
ics test score	.18	.00	.54	.72	.68		-.02	1.03	1.02	1.00
test score		.12	.55	.67	.64	.67*	-.01	.11	.76	.73
vel	-.02	.00	.28	.26	.29	.02	.00	.26	.28	.41
	-.07	-.01	-.02	-.10	-.10	-.06	.01	.02	-.03	-.04
ee Sources										
ics grade recd.			.67	.67	.66			.54	.54	.57
grade received			.58	.58	.60			.52	.52	.57
ics test score			.73	.73	.70			.70	.70	.71
test score			.57	.57	.54			.60	.60	.66
vel			.55	.55	.54			.74	.74	.75
			-.03	-.03	.00			-.02	-.02	-.03
Explained	.67*					.58*				

GRADE THREE

	Grade Received					Test Score				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
e Sources										
onwhite, 2=white)	-.07	.01	.29	.23	.23	-.05	.04	.37	.37	.37
=male, 2=female)	-.03	.01	.18	.16	.15	.01	-.01	.17	.17	.16
Sources										
ics grade recd.		.05	.45	.49	.51		.07	-.47	.54	.55
ics test score		.02	.48	.50	.62		.04	.51	.55	.58
Sources										
ics grade recd.	.24*	.07	.35	.67	.66	.23*	.04	.45	.73	.70
grade received		.16	.51	.67	.65		.14	.54	.68	.69
ics test score	.02	-.01	.52	.54	.57	.17*	.07	.45	.70	.71
test score		.01	.61	.62	.59		.16	.58	.74	.75
vel		-.01	.27	.27	.41		.00	.17	.17	.29
		-.02	-.10	-.12	-.15		-.03	-.04	-.08	-.06
ee Sources										
ics grade recd.		.05	.97	1.02	1.00	.20*	.01	.50	.70	.69
grade received	.58*	.03	.21	.82	.81		.12	.57	.69	.69
ics test score	.25*	.01	.45	.70	.69		.05	.97	1.02	1.00
test score		.12	.52	.64	.61	.44*	.03	.34	.80	.78
vel	-.13*	.00	.52	.39	.44	.04	-.03	.55	.56	.64
	-.09	-.01	-.07	-.16	-.15	.00	-.02	-.03	-.05	-.04
Explained	.73*					.74*				

ficant at the .05 level.

teacher perceptions of student personality and deportment, the grade the teacher assigns to the student for the other subject. While a proxy, its interpretation, given the statistical controls is reasonable.

Overall, then, the results are not unequivocal. They point, however, to enough evidence of non-meritocratic inputs into grading to raise serious questions about the meritocratic pretensions of elementary school grading.

NOTES

1. The analysis was constrained to use achievement test scores from the batteries chosen by the district and in the scoring form recorded in cumulative records that year as follows:

Grade	Testing Date	Subject	Test	Form of Score
First	May, 1978	Reading	CAT* reading	raw score
First	May, 1978	Mathematics	DMI**	objectives achieved
Second	May, 1979	Reading	CAT reading	raw score
Second	May, 1979	Mathematics	DMI	objectives achieved
Third	May, 1980	Reading	CAT reading	grade equivalent
Third	May, 1980	Mathematics	CAT mathematics	grade equivalent

*California Achievement Test

**Diagnositc Mathematics Inventory

Research using the reading subtests of the norm-referenced California Achievement Test has found its language not to give special advantages to students of a particular race (Marwit and Neumann, 1974). Its strength is in comparing students' achievements, rather than in measuring performance relative to learning objectives (Smith, 1978). The strength of the criterion-referenced Diagnostic Mathematics Inventory is its utility for individualizing instruction. It suffers from overemphasis on computation and fact, rather than on mathematical reasoning (O'Brien, 1978).

2. Reciprocal causation calls for two stage least squares. Fox's (1980) algorithm requires the covariance matrix of the error terms. This is available in SAS's Proc Sysreg from three stage least squares. In order to use consistent inputs into the algorithm, we have used third stage structural coefficients. These may differ somewhat from those derived at the second stage.

REFERENCES

Alexander, K. E., M. Cook, and E. L. McDill

1978 "Curriculum tracking and educational stratification: some further evidence." American Sociological Review 43:47-66.

Asher, H. B.

1976 Causal Modeling. Beverly Hills, Calif.: Sage.

Bloom, B. S.

1976 Human Characteristics and School Learning. New York: McGraw-Hill.

Boocock, S. S.

1980 Sociology of Education: An Introduction. 2nd ed. Boston: Houghton Mifflin.

Bowles, S. and H. Gintis

1976 Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books.

Brophy, J. E. and T. L. Good

1974 Teacher-Student Relationships: Causes and Consequences. New York: Holt, Rinehart, and Winston.

Clifton, R.

1981 "Ethnicity, teachers' expectations, and the academic achievement process in Canada." Sociology of Education 54:291-301.

Davidson, H. H. and G. Lang

1960 "Children's perceptions of their teachers' feelings toward them related to self-perception, school achievement and behavior." Journal of Experimental Education 29:107-18.

Dreeben, R.

1968 On What is Learned in School. Reading, Mass.: Addison-Wesley.

Duncan, Otis Dudley

1968 "Inheritance of poverty or inheritance of race." Pp in D. P. Moynihan (ed.), On Understanding Poverty: Perspectives from the Social Sciences. New York: Basic Books.

Evans, F. B.

1976 "What research says about grading." Pp. 30-50 in Degrading the Grading Myths: Primer of Alternatives to Grades and Marks, edited by S. B. Simon and J. A. Bellanca. Washington, D. C.: Association for Supervision and Curriculum Development.

Finn, J.

1972 "Expectations and the educational environment." Review of Educational Research 42:387-410.

Fox, J.

1980 "Effect analysis in structural equation models: Extensions and simplified methods of computation." Sociological Methods and Research 9:3-28.

Gravenburg, O. and G. Collins

1976 "Grades: Just a measure of conformity." Humboldt Journal of Social Relations 3:58-62.

Harvey, D.G. and G.T. Slatin

1975 "The relationship between child's SES and teacher expectations: A test of the middle-class bias hypothesis." Social Forces 54:140-59.

Heise, D.R.

1976 Causal Analysis. New York: Wiley.

Heyns, B.

1974 "Social selection and stratification within schools." American Journal of Sociology 79:1434-1451.

Hunt, J. McV.

1961 Intelligence and Experience. New York: Ronald Press.

Kohl, H.

1967 36 Children. New York: New American Library.

Jencks, C.

1980 "Heredity, environment, and public policy reconsidered." American Sociological Review 45:723-36.

Jencks, C., S. Barlett, M. Corcoran, J. Crouse, D. Eaglesfield, G. Jackson, K. McClelland, P. Mueser, M. Olneck, J. Schwartz, S. Ward, and J. Williams

1979 Who Gets Ahead? The Determinants of Economic Success in America. New York: Basic Books.

Jencks, C., M. Smith, H. Acland, M.J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson

1972 Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books

Jensen, A.R.

1973 Educability and Group Differences. New York: Harper and Row.

Kirschenbaum, H., R. Napier, and S.B. Simon

1971 Wad-ja-get? The Grading Game in American Education. New York: Hart.

Leiter, J.

1982 "Classroom composition and achievement gains." Sociology of Education 56:126-32.

Maccoby, E. and C.N. Jacklin

1974 The Psychology of Sex Differences. Stanford, Calif.: Stanford University Press.

Maehr, M.L.

1976 "Continuing motivation: An analysis of a seldom considered educational outcome." Review of Educational Research 46:443-62.

Marwit, S.J. and G. Neumann

1974 "Black and white children's comprehension of standard and nonstandard English passages." Journal of Educational Psychology 66:329-32.

O'Brien, T. C.

1978 "Diagnostic Mathematics Inventory." Pp. 401-5 in O. K. Buros (ed.), The Eighth Mental Measurements Yearbook, volume 1. Highland Park, N.J.: Gryphon.

Pedulla, J. J., P. W. Airasian, and G. F. Madaus

1980 "Do teacher ratings and standardized test results of students yield the same information?" American Educational Resarch Journal 17:303-7.

Persell, C. H.

1977 Education and Inequality: A Theoretical and Empirical Synthesis. New York: Free Press.

Rosenbaum, J. E.

1976 Making Inequality: The Hidden Curriculum of High School Tracking. New York: Wiley.

1978 "The structure of opportunity in school." Social Forces 57:236-56.

Salili, F., M. L. Maehr, R. L. Sorensen, and L. J. Tyans Jr.

1976 "A further consideration of the effects of evaluation on motivation." American Educational Research Journal 13:85-102.

Schlechty, P. C.

1976 Teaching and Social Behavior: Toward an Organizational Theory of Instruction. Boston: Allyn and Bacon.

Smith, Kenneth

1978 "California Achievement Test: Reading, 1970 Edition." Pp. 1160-3 in O. K. Buros (ed.), The Eighth Mental Measurements Yearbook, volume 2. Highland Park, N. J.: Gryphon.

Sorenson, A. B.

1970 "Organizational differentiation of students and educational opportunity." Sociology of Education 43:355-76.

Stallings, J.

1980 "Allocated academic learning time revisited, or beyond time on task." Educational Researcher 9:11-16.

Terwilliger, J. S.

1966 "Self-reported marking practices and policies in public secondary schools." Bulletin of the National Assocation of Secondary School Principals 50:5-37.

Veldman, D. J. and J. E. Brophy

1974 "Measuring teacher effects on pupil achievement." Journal of Educational Psychology 66:319-24.

Waller, W.

1932 The Sociology of Teaching. New York: Wiley.

Wiley, D. E.

1976 "Another hour, another day: Quantity of schooling, a potent path for policy." Pp. 225-65 in W. H. Sewall. R. M. Hauser, and D. L. Featherman (eds.), Schooling and Achievement in American Society. New York: Academic Press.

Williams, T.

1976 "Teacher prophecies and the inheritance of inequality." Sociology of Education 49:223-36.

Appendix: Zero-order Correlations and Descriptive Statistics for Variables in Analyses (excludin correlations reported in Tables 1 and 2) (N=213)

Ascriptive Variables	(2)	(3)	(4)	(5)	(6)	(11)	(12)	(13)	(18)	(19)	(20)	Mean	SD
(1) Race (1=non-white, 2=white)	-.02*	.25	.31	.25	.27	.29	.16	.21	.37	.37	.20	1.48	.50
(2) Gender (1=male, 2=female)		.22	.18	.16	.11*	.16	.13*	.03*	.17	.13*	-.02*	1.54	.50
Grade One													
(3) Reading grade			.62	.76	.53	.45	.36	-.09*	.52	.46	.01*	10.10	2.74
(4) Reading test score				.51	.70	.74	.58	-.06*	.67	.58	.01*	41.70	11.18
(5) Mathematics grade					.47	.38	.26	-.03*	.48	.41	.90*	10.69	2.59
(6) Mathematics test score						.39	.49	-.13*	.43	.37	-.02*	28.79	5.55
Grade Two													
(7) Reading grade	.17	.66	.67	.58	.52	.50	.29	-.06*	.62	.55	-.03*	9.77	2.61
(8) Reading test score	.18	.61	.76	.52	.57	.64	.48	.01*	.88	.79	-.05*	53.81	13.90
(9) Mathematics grade	.16	.54	.57	.58	.53	.47	.29	-.10*	.57	.54	.00*	10.70	2.45
(10) Mathematics test score	.06*	.50	.59	.50	.52	.54	.41	-.04*	.69	.75	-.03*	39.42	9.20
(11) Reading track level	.16	.45	.74	.38	.39		.77	.06*	.69	.61	-.02*	41.62	8.22
(12) Mathematics track level	.13*	.36	.58	.26	.49			.04*	.52	.47	.00*	28.65	2.66
(13) Absences	-.03*	-.09*	-.06*	-.03*	-.13*				.04*	.01*	.63	5.73	5.70
Grade Three													
(14) Reading grade	.19	.59	.60	.51	.52	.40	.37	-.02*	.56	.51	-.07*	9.64	2.74
(15) Reading test score	.17	.58	.71	.48	.54	.55	.42	-.03*	.67	.68	.01*	39.92	13.89
(16) Mathematics grade	.15	.57	.59	.51	.62	.41	.41	-.15	.49	.44	-.15	9.78	3.01
(17) Mathematics test score	.16	.58	.68	.55	.58	.48	.29	-.06*	.68	.64	-.04*	40.86	8.87
(18) Reading track level	.17	.52	.67	.48	.43	.69	.52	.04*		.90	.01*	53.62	12.12
(19) Mathematics track level	.13*	.46	.58	.41	.37	.61	.47	.01*			.00*	39.07	7.21
(20) Absences	-.02*	.01*	.01*	.09*	-.02*	-.02*	.00*	.63				6.29	6.39

*NOT significant at the .05 level.