DOCUMENT RESUME

ED 076 708 TM 002 717

AUTHOR Veldman, Donald J.; Brophy, Jere E.
TITLE Measuring Teacher Effects on Pupil Achievement.
PUB DATE 73
NOTE 7p.; Paper presented at a symposium of annual meeting of American Educational Research Association (New Orleans, Louisiana, February 25-March 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; *Achievement Gains; Achievement Tests; Grade 2; Grade 3; Observation; Self Evaluation; Tables (Data); *Teacher Evaluation; *Teacher Influence; Technical Reports

ABSTRACT

Four successive years of Metropolitan Achievement Test scores for all second- and third-grade pupils in a large Southwestern urban school system were used to compute mean residual gain scores for 170 teachers who taught during the three focal years. Alternate regression models were compared for precision in estimating the magnitude and consistency of teacher influence on pupil achievement. Consistency was examined across MAT subtests, pupil sex, and the three focal school years. Thirty-one of the most consistent teachers were intensively studied with self-report instruments and classroom observations. (Author)

MEASURING TEACHER EFFECTS ON PUPIL ACHIEVEMENT

Donald J. Veldman and Jere E. Brophy

The University of Texas at Austin

Although few would argue that the amount of gain pupils show on standard achievement tests is the only or even the best measure of teaching effectiveness, it is one which is increasingly used for this purpose. The present paper concerns methodological considerations involved in obtaining unbiased estimates of teacher influence on pupil achievement.

Samples and Measures

All second and third grade teachers in a large Southwestern urban school system were considered for inclusion in a comprehensive investigation of teacher effectiveness, classroom behavior, and personal characteristics. All teachers were female. From the full sample teachers were selected who (1) had at least five years of teaching experience, (2) had taught the same grade level during the three focal years (1967-1969), and (3) had at least 14 children with available data each of these years. The teachers represented 15 Title I (poverty area) and 35 Non-Title I schools. The four samples resulting from this selection were:

21 Grade 2, Title I teachers (1210 pupils)

35 Grade 2, Non-Title I teachers (2168 pupils)

20 Grade 3, Title I teachers (1216 pupils)

39 Grade 3, Non-Title I teachers (2744 pupils)

Pupil records were retrieved from school files for each of four successive years of regular fall achievement testing. Grade equivalent scores were obtained for the Metropolitan Achievement Test (MAT) subscales.

Different forms of the Metropolitan battery were used with each of the four samples, necessitating separate statistical analyses.

In the tables reported later in this paper, the following abbreviations are used for the subtests of the MAT:

WK = Work Knowledge

WD = Word Discrimination

R = Reading

AC = Arithmetic Computation

AR = Arithmetic Reasoning

V = Verbal Subtest Average

Q = Quantitative Subtest Average

T = Average of V and Q

Influences on Predictive Efficiency

It is now generally accepted that residual gain scores are superior to simple pre-post difference scores as measures of teacher influence. What is not clear, however, is the importance of residualizing with more than the simple pre-test variable. The first series of analyses were designed to explore this problem.

A series of regression models were compared to one containing (1) pre-test, (2) squared pre-test, (3) pupil sex, (4) year of testing, and (5) teacher. In each comparison, one of these influences was omitted to determine its contribution to prediction of the post-test criterion.

Table 1 contains the results of these comparisons, expressed as percentages of criterion variance associated with each influence. In summary, we drew the following tentative conclusions from these data:

(1) Inclusion of a squared-score variable to permit curvilinear regression adds little to the precision of the estimates.

(2) Inclusion of pupil sex in the model adds very little to predictive efficiency.

(3) Systematic differences among the three years of testing are trivial.

(4) Variation among teachers is substantial.

(5) The influences of sex, year, and teacher appear to be stronger in Title I than in Non-Title I schools.

(6) Teacher impact appears to be stronger on verbal skills than on quantitative skills in Title I schools, and vice versa in Non-Title I schools.

(7) Teacher impact is stronger in grade 3 than in grade 2 in Title I schools, but about equal in Non-Title I schools.

(8) Predictability of post-test scores of pupils is generally greater in grade 3 than in grade 2, in Non-Title I than in Title I schools, and on verbal than on quantitative measures.

Consistency of Teacher Impact

The next step in the analysis addressed the question of the degree to which teachers' influence on child gain was consistent across three successive years, and hence, classes of pupils.

Residual gain scores for all pupils were obtained using only simple pre-test scores as covariates. These were then averaged for each teacher for each of her three classes. These average residual gains were then used to compute intraclass correlations among the three years for each of the four samples of teachers. The results of this analysis are shown in Table 2. With the exception of the second grade, Title I sample where none of the coefficients were statistically significant, it is

apparent that three-year averages provide a reasonably high level of reliability as estimates of teacher impact on student learning.

Summary

A few final comments may be in order. First of all, our data suggest that reasonably stable estimates of teacher influence can be obtained from standardized achievement measures of pupil performance. The increasing use of the team approach in elementary schools, however, lessens the practical interest of such measures.

The differences between Title I and Non-Title I schools are consistent with the theoretical position that the school is relatively more important, compared to the home, in determining the achievement levels of underprivileged pupils, than it is for advantaged pupils. This in turn suggests that the quality of teaching is more crucial in such settings than in advantaged schools.

Table 1a

Influences in Grade 2, Title I Sample

MAT Subtest	Means		Influences (%)				
	Pre	Post	Pre	Pre^2	Sex	Year	Teacher
WK	1.74	2.41	32.65	.14	.23	.20	6.14
WD	1.83	2.81	42.43	.14	.30	1.71	3.55
R	1.77	2.52	22.34	.07	1.43	.29	6.90
AC	1.87	2.79	27.80	.57	.00	1.04	4.50
V	1.78	2.58	46.77	.13	.47	1.10	5.30
T	1.83	2.68	42.85	.09	.34	1.56	5.75

Table 1b

Influences in Grade 2, Non-Title I Sample

MAT Subtest	Means		Influences (%)				
	Pre	Post	Pre	Pre^2	Sex	Year	Teacher
WK	2.50	3.71	61.01	.30	.03	.03	2.69
WD	2.82	3.70	58.11	.41	.31	.04	2.95
R	2.51	3.62	53.17	.01	.01	.04	2.97
AC	2.52	3.18	29.93	.92	.16	.22	5.66
AR	2.52	3.25	40.12	.00	.00	.07	4.22
V	2.61	3.67	71.72	.24	.00	.03	2.28
Q	2.52	3.22	42.30	.24	.03	.05	5.02
T	2.57	3.44	70.38	.00	.02	.00	3.04

Table 1c

Influences in Grade 3, Title I Sample

MAT Subtest	Means		Influences (%)				
	Pre	Post	Pre	Pre^2	Sex	Year	Teacher
WK	2.51	3.27	37.15	.90	.10	.33	17.64
WD	2.82	3.34	44.33	.35	.65	1.85	4.96
R	2.62	3.25	34.36	.36	.21	.16	10.86
AC	2.76	3.44	29.89	.03	.41	1.18	9.34
AR	2.76	3.05	26.63	.86	.36	.36	4.04
V	2.65	3.29	51.28	1.00	.16	.94	12.44
Q	2.76	3.25	34.40	.14	.47	.87	6.31
T	2.70	3.27	50.08	1.16	.41	1.44	9.83

Table 1d

Influences in Grade 3, Non-Title I Sample

MAT Subtest	Means		Influences (%)				
	Pre	Post	Pre	Pre	Sex	Year	Teacher
WK	3.67	4.85	64.98	.67	.00	.02	2.16
WD	3.66	4.62	67.08	.99	.03	.09	1.29
R	3.52	4.65	57.09	1.47	.00	.15	1.41
AC	3.14	4.13	33.84	1.03	.07	.07	6.94
AR	3.23	4.23	50.55	.00	.02	.12	4.28
V	3.62	4.71	76.04	1.16	.00	.08	1.08
Q	3.19	4.18	54.89	.14	.00	.12	5.23
T	3.40	4.44	75.92	.39	.00	.12	2.16

Table 2

Intraclass Correlations Across Three Years

Subtest	G2TI	G2NTI	G3TI	G3NTI
WK	.43	.66*	.78*	.63*
WD	.36	.74*	.26	.49*
R	.24	.66*	.50*	.23
AC	.00	.48*	.63*	.80*
AR	--	.61*	.27	.64*
V	.35	.71*	.65*	.38*
Q	--	.59*	.50*	.75*
T	.19	.69*	.54*	.65*

*p<.05.