DOCUMENT RESUME

ED 229 210 RC 014 052

AUTHOR Powers, Stephen; Rossman, Mark H.
TITLE Evidence of the Impact of Bilingual Education: A Meta-analysis.
PUB DATE Mar 83
NOTE 13p.; Paper at the Annual Arizona Bilingual Education Conference (Tucson, AZ, March 17-19, 1983).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Bilingual Education Programs; Elementary Education; Limited English Speaking; *Longitudinal Studies; *Mathematics Achievement; *Program Effectiveness; *Reading Achievement
IDENTIFIERS Meta Analysis; *New Mexico (Santa Fe)

ABSTRACT

C.F. Leyba conducted a longitudinal study (1972-77) of the Title VII Bilingual Education Program in the Santa Fe Public Schools involving 1,257 students in grades 1-6. Three groups were studied: a longitudinal treatment group (LTG) which received uninterrupted bilingual education for a span of years; a nonlongitudinal treatment group which received bilingual education two or more years but only intermittently, due to dropping out and later returning; and control groups which were randomly selected each year from Title VII participating schools. The present study re-examined the Leyba data for the LTG and controls employing meta-analysis to obtain an overall estimate of the effect size of the program in reading and mathematics and to test the effect size for significance. The study used a weighted, unbiased estimate of effect size. Leyba reported a total of 16 comparisons of the LTG and controls in reading and 16 comparisons in mathematics. Data were analyzed using a computer program. Analysis showed an overall effect of the program on the students' mathematics achievement. The overall effect of the program did not reach statistical significance in reading. Results suggested gains for bilingual education students which clearly do not support a hypothesis of deleterious effects of bilingual education. (NQA)

EVIDENCE OF THE IMPACT OF BILINGUAL EDUCATION: A META-ANALYSIS[1]

Stephen Powers
Tucson Unified School District

Mark H. Rossman
Arizona State University

[1] Paper presented at the Annual Arizona Bilingual Education Conference (in the session: De la Garza, Jesus, Sanchez, Virginia, and Powers, Stephen: Language Assessment Parental Involvement and Evaluation: Current Trends in Bilingual Education), Tucson, Arizona, March 17-19, 1983.

Leyba (1978) conducted a longitudinal study of the Title VII Bilingual Education Program in the Santa Fe Public Schools 1972 - 1977 involving 1257 students in Grades 1 - 6. This program was identified as one of seven exemplary bilingual education programs by the American Institutes of Research (Leyba 1978). A description of the program is contained in the Leyba study (Leyba 1978).

Students in schools eligible for Title VII funding came from a bilingual home and had limited English speaking proficiency. However, to be a participant in the bilingual program, parents had to volunteer their child for the program. Each year control groups were randomly selected at each school. Among other measures, students were pre-tested (fall) and post-tested(spring) with the Metropolitan Achievement Test, Total Reading and Total Mathematics subtests.

Leyba studied three groups (1) a longitudinal treatment group (LTG) which received uninterrupted bilingual education for a span of years; (2) a nonlongitudinal treatment group (NLTG) which received bilingual education two or more years but only intermittently, due to dropping out and later returning, and (3) control groups which were randomly selected each year from Title VII participating schools. Four longitudinal groups were identified: Group A consisted of students enrolled in the bilingual program for three years beginning with Grade 1; Group B consisted of students enrolled in the bilingual program for four years beginning with Grade 1; Group C included students enrolled in the bilingual education program for five years beginning with Grade 2; and Group D consisted of students enrolled in the bilingual program for five years beginning with Grade 2.

Leyba computed the post-test minus pre-test for each student in reading and mathematics for the LTG, NLTG, and controls in Groups A - D. Independent t-tests were used to compare the mean difference scores of the LTG with the NLTG, the LTG with the controls, and the NLTG with the controls. Leyba (1978, p. ii) concluded from the comparisons of the NTG, NLTG, and controls that "The Title VII students over time in the majority of cases outperformed the non-Title VII students in reading and mathematics.

Several limitations surround Leybas (1978) conclusions:

1. Only 21% of the reported comparisons of the LTG and NLTG with controls yielded statistically significant results favoring LTG or NLTG groups. A total of 60% of the comparisons yielded positive mean differences which favored bilingual program students. These suggested a positive effect of the bilingual program. It is unclear how Leyba (1978) reached his conclusions about the favorable impact of the bilingual program with so few statistically significant results.

2. Leyba performed multiple comparisons of the LTG, NLTG, and control groups at the .05 level of significance. Miller (1981) has shown that such comparisons will change the experiment-wise error rate from .05 to .14 which will increase the number of untrue significant findings (Type I errors).

The purpose of the present study is to re-examine the Leyba data for the LTG and controls employing meta-analysis (1) in order to obtain an overall estimate of the effect size of the Santa Fe Bilingual Education Program in reading and mathematics and (2) to test the effect size for significance.

METHOD

The effect size has been described by Glass (1976) as the difference between the experimental group mean and the control group mean divided by the standard deviation of the control group. Hedges (1981) has shown that Glass's effect size has a small bias which can be corrected. He provides methods for computing an unbiased effect size for a series of experiments as well as a z-test for the significance of the effect size. Further, Hedges' (1981) procedures allow for pooled variances to be used. The present study used a weighted, unbiased estimate of effect size because weights give "more weight to experiments which contribute more precise information to the overall effect size estimate" (Hedges, 1981, p. 125). Since Leyba did not provide standard deviations in his report, effect sizes were obtained with the following formula:

$$ES = t \sqrt{(1/n_1 + 1/n_2)}$$

Refer to Appendix A for the data of each cohort.

PROCEDURE

The present analysis concerned only the LTG and control groups. Because the LTG group received uninterrupted bilingual education, it appeared that the LTG group would most clearly reflect the impact of the bilingual education program. Leyba (1978) reports a total of 16 comparisons of the LTG and controls in Reading and 16 comparisons in mathematics. A computer program was written (Powers, 1983) to analyze the data.

FINDINGS AND DISCUSSION

Effect sizes for reading ranged from .92 to -.39, and in mathematics from 1.27 to -.43. Refer to Appendixes B and C for computational details. The overall average effect size in reading was .12 ($z = 1.48$, $p < .14$) and for mathematics was .29 ($z = 3.63$, $p < .01$). The effect size in reading was positive but it did not reach significance whereas in mathematics the effect size was significant beyond the .01 level. Pearson correlations between grade level and effect sizes reveal a negative correlation of -.52 ($p < .05$) in reading and -.53 ($p < .05$) in mathematics. The negative correlation indicates that initial impact was greater in earlier grades.

The present analysis has shown an overall effect of the Santa Fe bilingual Education program on the mathematics achievement of students. The overall effect of the bilingual program did not reach statistical significance in reading. However, the results suggested gains for bilingual education students, which clearly do not support a hypothesis of deleterious effects of bilingual education.

FOOTNOTE

[1]The authors thank Darrell L. Sabers, Department of Educational Psychology, University of Arizona, for his comments on an earlier version of this paper. Copies of this paper may be obtained from Stephen Powers, Ph.D, Legal and Research Services, Tucson Unified School District, P. O. Box 40400, Tucson, Arizona 85717.

REFERENCES

Glass, G. V. Primary, Secondary, and Meta-analysis of Research. Educational Researcher, 1976, 5, 3-8.

Hedges, L. V. Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. Journal of Educational Statistics, 1981, 6, 107-128.

Leyba, C. F. Longitudinal Study Title VII Bilingual program: Santa Fe public schools. National Dissemination and Assessment Center, California State University, Los Angeles, 1978.

Miller, R. G. Simultaneous Statistical Inference (2nd ed.). New York: Springer-Verlag, 1981.

Powers, S. A Computer Program to Calculate the Modified Glass Estimator. Educational Psychological Measurement, 1983, in press.

APPENDIX A

Sample Sizes and T-Values Comparing Bilingual and Control Groups as Reported in Leyba (1978)

Cohort	Grade 1			Grade 2			Grade 3			Grade 4			Grade 5			Grade 6		
	N_1	N_2	t	N_1	N_2	t	N_1	N_2	t	N_1	N_2	t	N_1	N_2	t	N_1	N_2	t
Reading																		
A	36	24	.56[a]	32	20	.85	32	19	.10									
B	23	18	2.97*	23	17	2.21*	23	19	.59	23	19	-.20						
C							26	23	-.12	26	22	-.34	26	19	-1.30	26	23	.75
D				16	9	-.14	16	19	-.16	16	23	.00	16	20	.18	16	22	.15
Mathematics																		
A	36	24	.28[a]	32	20	2.82*	32	19	.71									
B	23	18	4.11*	23	17	2.51*	23	19	-.49	23	19	-.39						
C							26	23	.86	26	22	-1.50	26	19	.38	26	23	-.31
D				16	9	2.99*	16	19	1.68	16	23	.11	16	20	1.58	16	22	.82

*$p < .05$

[a]These values were incorrectly reported in Leyba (1978) as -.56 and -.28.

READING

THE MODIFIED GLASS EFFECT SIZE ESTIMATOR

	N1	N2	DF	T-VALUE	CORRECTION FACTOR	WEIGHTS	EFFECT SIZE	VARIANCE
1	36	24	58	0.5600	0.9870	0.0879	0.1457	0.0696
2	23	18	39	2.9700	0.9806	0.0591	0.9166	0.1098
3	32	20	50	0.8500	0.9849	0.0758	0.2386	0.0818
4	23	17	38	2.2100	0.9801	0.0576	0.6928	0.1086
5	16	9	23	-0.1400	0.9670	0.0348	-0.0564	0.1737
6	32	19	49	0.1000	0.9846	0.0742	0.0285	0.0839
7	23	19	40	0.5900	0.9811	0.0606	0.1795	0.0965
8	26	23	47	-0.1200	0.9840	0.0712	-0.0338	0.0820
9	16	19	33	-0.1600	0.9771	0.0500	-0.0530	0.1152
10	23	19	40	-0.2000	0.9811	0.0606	-0.0608	0.0962
11	26	27	46	-0.3400	0.9836	0.0697	-0.0969	0.0840
12	16	23	37	0.0000	0.9796	0.0561	0.0000	0.1060
13	26	19	43	-1.3000	0.9825	0.0652	-0.3855	0.0918
14	16	20	34	0.1800	0.9778	0.0515	0.0590	0.1126
15	25	23	47	0.7500	0.9840	0.0712	0.2112	0.0824
16	16	22	36	0.1500	0.9790	0.0545	0.0483	0.1080

Appendix B

EFFECT SIZE ESTIMATOR FOR A SERIES OF EXPERIMENTS

NUMBER OF EXPERIMENTS = 16

WEIGHTED UNBIASED ESTIMATOR = 0.1161 STANDARD DEVIATION= 0.0787

VARIANCE = 0.0062

95% CONFIDENCE INTERVAL: -0.038 , 0.270 Z-TEST = 1.4750

MATHEMATICS

THE MODIFIED GLASS EFFECT SIZE ESTIMATOR

	N1	N2	DF	T-VALUE	CORRECTION FACTOR	WEIGHTS	EFFECT SIZE	VARIANCE
1	36	24	58	0,2800	0,9870	0,0879	0,0728	0,0695
2	23	18	39	4,1100	0,9806	0,0591	1,2684	0,1197
3	32	20	50	2,8200	0,9849	0,0758	0,7917	0,0875
4	23	17	38	2,5100	0,9801	0,0576	0,7869	0,1104
5	16	9	23	2,9900	0,9670	0,0348	1,2048	0,2052
6	32	19	49	0,7100	0,9846	0,0742	0,2025	0,0843
7	23	19	40	-0,4900	0,9811	0,0606	-0,1490	0,0964
8	26	23	47	0,8600	0,9840	0,0712	0,2422	0,0826
9	16	19	33	1,6800	0,9771	0,0500	0,5570	0,1198
10	23	19	40	-0,3900	0,9811	0,0606	-0,1166	0,0963
11	26	22	46	-1,5000	0,9836	0,0697	-0,4274	0,0859
12	16	23	37	0,1100	0,9796	0,0561	0,0351	0,1060
13	26	19	43	0,3800	0,9825	0,0652	0,1127	0,0912
14	16	20	34	1,5800	0,9778	0,0515	0,5182	0,1164
15	26	23	47	-0,3100	0,9840	0,0712	-0,0873	0,0820
16	15	22	36	0,8200	0,9790	0,0545	0,2638	0,1089

EFFECT SIZE ESTIMATOR FOR A SERIES OF EXPERIMENTS

NUMBER OF EXPERIMENTS = 16

WEIGHTED UNBIASED ESTIMATOR = 0,2869 STANDARD DEVIATION= 0,0790

VARIANCE = 0,0062

95% CONFIDENCE INTERVAL: 0,132 , 0,442 Z-TEST = 3,6304

Appendix C