DOCUMENT RESUME

ED 065 585 TM 001 854

AUTHOR Norton, Daniel P.; Faunce, Richard W.
TITLE Sibling Methodology in Evaluative Research.
PUB DATE Apr 72
NOTE 3p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Analysis of Variance; *Comparative Analysis; *Evaluation Techniques; Kindergarten Children; *Reading Readiness Tests; *Siblings; Student Testing
IDENTIFIERS Metropolitan Readiness Test; Minneapolis; *Pyramid Reading Program

ABSTRACT

Evaluation of the Pyramid Reading Program, Title I, ESEA, of the Minneapolis Public Schools was carried out using sibling methodology. The reading readiness at entry of kindergarten students was compared with that of their older siblings at their time of first grade entry, and fourth grade reading achievements will also be compared. The first phase of data collection led to the identification of 343 first graders of 1970 whose older siblings had also been tested with the Metropolitan Readiness Test at first grade entry. The 1970 examinees were classified by sex and by exposure to program materials alone or to both materials and teachers. Two analyses were performed on the data. Results showed a significant difference in favor of target pupils over their older siblings. Neither sex nor program exposure was significant. (DB)

13.10

Daniel P. Norton
Educational Testing Service
960 Grove Street
Evanston, Illinois 60201

Sibling Methodology in Evaluative Research

by

Daniel P. Norton
Educational Testing Service

Richard W. Faunce
Minneapolis Public Schools

The development and implementation of innovative educational programs almost always proceed so the program evaluator cannot apply true experimental methodology to program evaluation. Frequently the performances of treatment groups cannot be meaningfully compared to those of any other group(s). Comparisons that are made often poorly account for highly relevant family, community and school variables. Evaluation reports that arise from such projects may simply reflect artifacts of program implementation. At best recommendations are often ambiguous and indecisive.

The Minneapolis Public Schools sought to deal with this problem in evaluating the Pyramid Reading Program, a Title I, ESEA, development project. The program was begun with the intention of responding to intensely expressed requests to help improve reading instruction. It called for the design, field testing and production of instructional materials and the education of teachers in their proper use. It was proposed that the target areas would fully adopt the new program as it was prepared. Individual teachers would be given training until, after several years, all would have had an opportunity to learn how to properly apply it. Gradual inclusion of students from different grade levels and buildings would, it was hoped, lead to full adoption if the program proved to be successful. Nothing resembling a true experiment was planned or possible within the circumstances of program initiation.

Program directors did not predict that reading achievement would immediately begin to show improvement. On the contrary, their best estimate was that decrements in reading achievement could occur over the first two years of program exposure. Only during a third or fourth year was it predicted that gains might be evident. This somewhat unusual projection for program impact made clear that mobility of both teachers and students had to be accounted for in evaluation design. Furthermore, unless unplanned diffusion of materials and ideas between children and teachers were effectively handled in design there would be slight chance of identifying program effects after four or more years of program operation.

Sibling methodology offered important advantages over other alternatives. During the first full year of the program the major impact of the project was to be made upon kindergarten students. Their reading readiness at entry to first grade would be compared with that of their older siblings at their time of first grade entry, and their fourth grade reading achievement would be compared with that of the same older siblings when they were in fourth grade. This approach promised important control for family and other environmental variables. Advantages of matched longitudinal over cross-sectional data were also to be gained. An important disadvantage to this approach was that uncontrolled historical effects would be present because the siblings were not in the same grade at the same time. In total, however, fewer important sources of invalidity were found to be present in the sibling alternative.

The first phase of data collection led to the identification of 343 first graders of 1970 whose older siblings had also been tested with the Metropolitan Readiness Test at entry to first grade. Their older siblings included 157 who were firstborn and 275 who were second or later. A total of 268 target students were firstborn and thus had no older siblings. In addition, 246 were later born but did not have older siblings who had been tested in earlier years. It was possible to classify the 1970 examinee population by sex and by exposure to either program materials alone or to both materials and teachers instructed in their proper use.

Two analyses were performed on the first phase data. The first employed program exposure (P), sex (S) and birth order (B) as factors for grade cohorts tested in 1970. All main and first order interaction effects of this analysis were not significant. In the second analysis scores of the 343 target students were compared with those of their older siblings grouped by birth order; sex was also included as a factor. Results of this analysis showed a significant difference in favor of target pupils over their older siblings, but no difference between their first or later born family cohorts. Neither sex nor its interaction with program exposure was significant.

Historical effects associated with dates of test administration and age attainments required for entry to first grade partially explain the differences that were noted. However, application of the basic design for an additional three years can largely resolve uncertainties arising from these historical effects and can partially clarify the importance of other sources of invalidity. Formidable logistical obstacles to implementation of sibling methodology can only be overcome by effective long range planning within educational agencies.

Table 1

Analysis of Variance of Fall 1971 Metropolitan Readiness Scores of 857 Students Who Received Benefits from the Title I, ESEA, Pyramid Reading Program in Kindergarten During 1970-71

Source of Variation	Degrees of Freedom	Mean Square	F
Program Exposure (P)	1	1.276	.291
Birth Order (B)	1	3.775	.862
Six (S)	1	7.612	1.738
P x B	1	9.433	2.154
P x S	1	.571	.130
B x S	1	2.663	.608
P x B x S	1	37.050	8.458**
Error	849	4.380	

**F(1,849;.01)=6.67

Table 2

Analysis of Variance of Citywide Metropolitan Readiness Scores of 343 Title I, ESEA, Pyramid Reading Students and 157 Firstborn and 275 Other Older Siblings

Source of Variability	Degrees of Freedom	Sum of Squares	Mean Square	F
Program Exposure (P)	2	37.080	18.540	7.708**
Sex (S)	1	.040	.040	.017
P x S	2	.957	.479	.199
Error	769		2.405	

**F(2,769;.01=4.64)