EVALUATING TITLE I EARLY CHILDHOOD PROGRAMS: PROBLEMS,THE APPLICABILITY OF MODEL C, AND SEVERAL EVALUATION PLANS

Keith McNeil and Emily Findlay

NTS Corporation

The enactment of the 1974 amendments to Title I expresses the growing concern for developing adequate evaluation strategies for assessing the effectiveness of Title I programs. The current Title I Evaluation and Reporting System (TIERS) was designed for use in grades 2-12. However, Title I programs are also present in the early childhood grades (prekindergarten, kindergarten, and first grade). Huron Institute is currently investigating alternative evaluation strategies to assess Title I programs in these grade levels (Bryk, Strenie, and Weisberg, 1979).

Although many evaluation models exist for evaluating early childhood programs, this paper will focus on the "Special Regression Model," Model C in the TIERS. We first present factors which make evaluating early childhood programs particularly difficult. The paper concludes with several feasible evaluation plans utilizing Model C.

Factors Which Make Evaluating Early Childhood Programs Difficult

Several factors make assessing the effectiveness of Title I early childhood

programs more difficult than in the upper grades. Factors that contribute to this difficulty include: 1) Scope of Programs, 2) Instrumentation for Early Childhood, 3) Developmental Characteristics of Early Childhood, and 4) Student Selection.

Scope of programs. The scope and content of early childhood programs vary widely across programs. Some of the variation is a result of grade levels. Most prekindergarten programs are global in nature, placing the greatest emphasis on early intervention and exposure to the school environment. Parental involvement in the students' education is also stressed. In contrast, kindergarten programs emphasize readiness skills: prereading, mathematics, language, and socio-emotional development. First grade programs begin to approximate programs found in upper grade Title I programs, by concentrating on supplementing regular school programs and bolstering the readiness skills taught in kindergarten.

Instrumentation. Serious concern has been voiced regarding the technical excellence of early childhood measures. The majority of measures do not meet minimum standards for validity, reliability and appropriateness of norms (Hoephner, Stern, and Nummedal, 1973). Measures which meet minimum criteria are of a cognitive nature, the majority being "IQ" type instruments. In many instances, these instruments are inappropriate measures of the goals and objectives of Title I programs. Furthermore, cognitive growth is so rapid at this age level that it cannot be measured, pre and post, by most instruments.

Developmental Characteristics. Young children have minimal experience in formal assessment situations. As a result, many of the prerequisite skills for test taking are lacking. In addition, egocentricity, emotional reactivity, and fluctuations in attention span may influence performance on assessment instru-

ments. Therefore, scores obtained from these instruments may not accurately measure the student's achievement level in the content area. Other assessment techniques may be needed to measure the student's achievement.

Student Selection. The last factor which contributes to the difficulty in assessing the effectiveness of Title I programs is student selection. Because instrumentation is problematic, test scores used as the sole criteria for selection, may identify many students as needing the program, when in fact some students do not. Likewise, many students are identified as not needing the program, when in fact some of those students do. Thus, the selection of the appropriate students is problematic. The design one chooses to evaluate the early childhood program must avoid this hazard, as well as deal with the previously mentioned factors.

The Special Regression Model -- Model C

The Special Regression Model (Model C) is a form of the regression projection model proposed by Campbell and Stanley (1963). The remarks in this paper actually pertain to Model C1, the "norm-referenced version of Model C" (referred to as Model C in this paper-see Tallmadge and Wood for complete discussion of the models). Many of the remarks do not apply to Model C2, the Model which uses a non-normed test for posttest. Expected posttest performance of the Title I group is based on the projection of the regression line from the comparison group. (See Figure 1.) If the Title I program is not effective, over and above the regular program, then the Title I effect will be zero. If the Title I program is effective, then the performance of the Title I students will be higher than predicted from the comparison group. Model C requires that all students in grades served by Title I in a target school be tested, and that the students who

score lowest on the pretest be placed in the Title I program with the remaining students serving as comparison students. The cutoff would be based on the number of students that can be served, and ideally would be the same in all buildings in a school district. Actually, the pretest/selection measure may be achievement test scores, independently made teacher ratings, classroom grades, or some combination of these or other types of measures.

Reasons for Using Model C

The first major reason for using Model C is the fact that students learn quite rapidly in the early years of schooling. This very often necessitates the use of a different test at posttest time than that used at pretest time. Model C allows for the use of a different test at the two sessions.

Secondly, selecting students for the Title I program in these early years can be difficult. Since most testing must be accomplished individually, testing takes a lot of time, it is usually very costly, and there are very few good screening devices available. In addition, the longer the screening takes, the less time there is for instruction. Because Model C uses the pretest as the selection device, Model C is preferable over other Title I evaluation Models which require that pretest and selection be separate. Because it is difficult to get one good screening device, it would be an additional burden to get another device which could be used as the pretest. Model C does not require two procedures for student selection and pretest, but requires that student selection be based on the pretest.

A third advantage of Model C is that it does not make any assumptions about the effectiveness of the regular program. Indeed many schools are not as effective with their regular curriculum as the average curriculum in the country. Because Model C uses the actual results from students in the regular

program, it yields an accurate indication of the Title I effect, over and above the regular program, in a particular school.

Fourthly, the procedure used for pretest/selection in Model C does not have to be a norm-referenced test. Therefore, teacher judgements, norm-referenced tests, or non-norm referenced tests could be used individually, or as part of a composite score. A composite score would most likely identify more accurately the students who are most in need of additional educational services. One way in which some Title I programs are implementing Model C is to use a non-norm referenced test at pretest, and a norm-referenced test at posttest. The pretest ranks students for student-selection and provides diagnostic/prescriptive information in terms of the kinds of skills the students lack. Thus posttest norm-referenced cutoff situation have pretest scores below some Title I students. Thus, there is a wider range in the comparison group over which the regression slope is determined.

Model C Evaluation Plans

Figure 2 contains three possible evaluation plans using Model C. For each plan, the testing time, test, and purpose of test are identified. For instance, those interested in testing only every Spring would find evaluation plan 2 appropriate. Since it would be unlikely that a pre-kindergarten spring score would be available, the kindergarten evaluation would probably have to be a Fall-Spring evaluation. But the kindergarten Spring score could also be used for the pretest for the grade 1 evaluation.

Figure 2 indicates that evaluation plans 1 and 2 both result in once a year testing. One should also note that the test used at any testing time may be the same as previously used, or a different test.

Summary

Because the purpose of Title I evaluation is to determine the effect of Title I over and above the regular curriculum, Model C seems most appropriate. The advantages of the Special Regression Model (Model C) for early childhood evaluation seem to outweigh the disadvantages. Those advantages are:

- Scope of Program - Model C allows for the use of a different test at posttest than used at pretest

- Instrumentation - non-norm referenced test can be administered at pre-test, providing diagnostic and prescriptive information as well as baseline and student selection information

- Developmental Characteristics - Different test levels can be used pre and post, even if they are not linked by a common scale

- Student selection - a separate selection/pretest is not needed and the selection/pretest can be a composite score
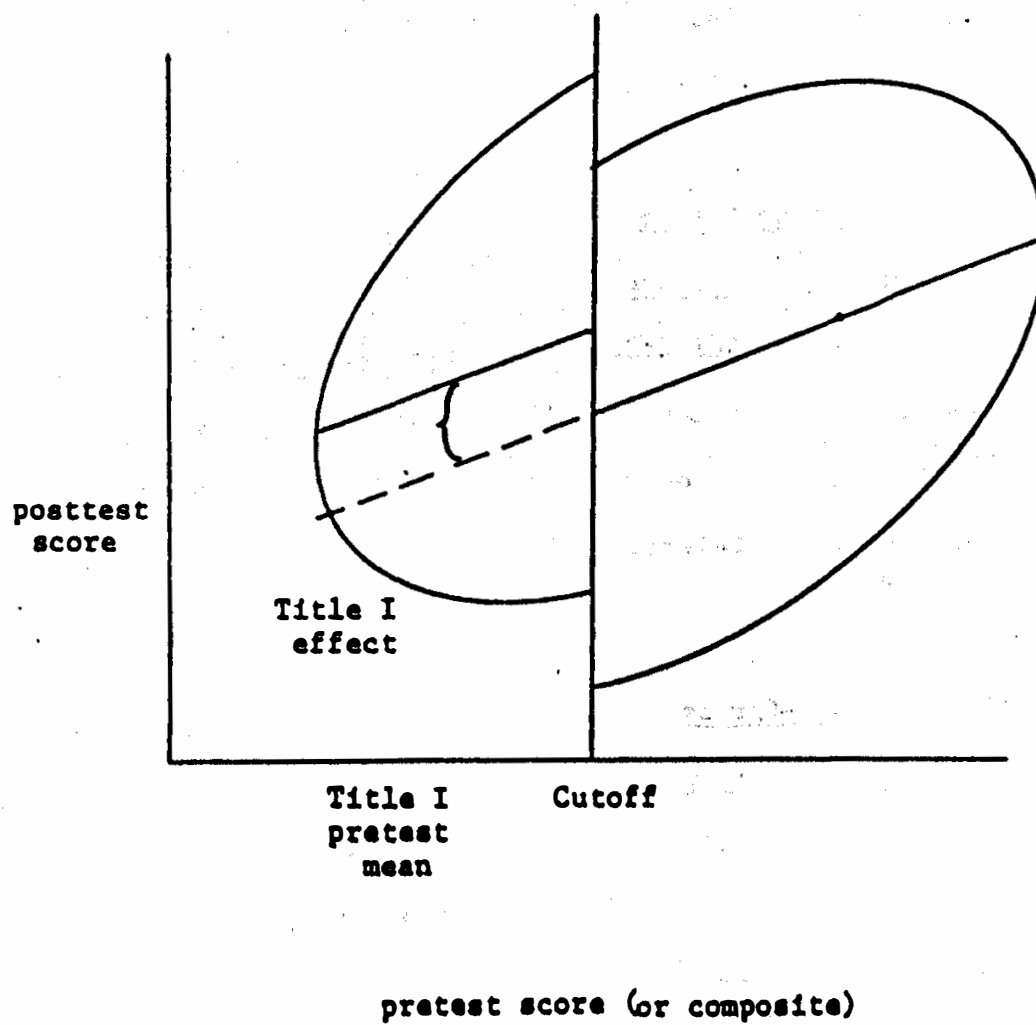
Figure 1. Model C approach to determining Title I effect

48

Evaluation Plan # 1 - - - Fall-Fall

Testing Time	Test	Purpose	K	Grade 1	Grade 2
Fall of K	A	Selection (K)	pre (K)		
Fall of 1	A or B	Selection (1)	post (K)	pre (1)	
Fall of 2	A or B or C	Selection (2)		post (1)	pre (2)

Evaluation Plan #2 - - - Spring-Spring

Testing Time	Test	Purpose	K	Grade 1	Grade 2
Fall of K	A	Selection (K)	pre (K)		
Spring of K	A or B	Selection (1)	post (K)	pre (1)	
Spring of 1	A or B or C	Selection (2)		post (1)	pre (2)
Spring of 2	A or B or C or D	Selection (3)			post (2)

Evaluation Plan #3 - - - Fall-Spring

Testing Time	Test	Purpose	K	Grade 1	Grade 2
Fall of K	A	Selection (K)	pre (K)		
Spring of K	A or B		post (K)		
Fall of 1	A or B or C	Selection (1)		pre (1)	
Spring of 1	A or B or C or D			post (1)	
Fall of 2	A or B or C or D or E	Selection (2)			pre (2)
Spring of 2	A or B or C or D or E or F				post (2)

Figure 2. Model C evaluation plans for early
childhood programs

References

Bryk, A.S., Strenio, J.F. and Weisberg, H.I. A method for estimating treatment effects when individuals are growing. Cambridge, MA: Huron Institute, 1979.

Campbell, D.T. and Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Hoephner, R., Stern, C., and Nummedal, S.G. CSE-ECRC preschool/kindergarten test evaluations. Los Angeles: UCLA Graduate School of Education, 1971.

McNeil, K.A. Problems with Model C, solutions, and an argument for the fuzzy cutoff. Unpublished manuscript, 1977.

Tallmadge, G.K. and Wood, C.T. User's Guide: ESEA Title I evaluation and reporting system (revised). Washington, D.C.: U.S. Office of Education, 1978.