

October 1967

Evaluation of Title I Projects - Some Constructive Criticism

Richard Elder

Follow this and additional works at: <https://scholarworks.gvsu.edu/mrj>

Recommended Citation

Elder, Richard (1967) "Evaluation of Title I Projects - Some Constructive Criticism," *Michigan Reading Journal*: Vol. 1: Iss. 2, Article 7.

Available at: <https://scholarworks.gvsu.edu/mrj/vol1/iss2/7>

This work is brought to you for free and open access by ScholarWorks@GVSU. It has been accepted for inclusion in Michigan Reading Journal by an authorized editor of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

Evaluation of Title I Projects — Some Constructive Criticism

by Richard Elder

In June, 1967, the Department of Education of the State of Michigan distributed a draft for evaluation of ESEA Title I projects at a meeting in Lansing. In this document there are specifications for the reporting of test data. The specifications take the form of sample Tables. Two of the Tables, 4A and 4B on pages 31 and 32 respectively, have caused much real concern among school administrators, reading specialists, and other professionals who are involved in evaluating ESEA Title I projects. This presentation is an analysis of the specifications implicit in Tables 4A and 4B, including the directions given orally at the above cited meeting. The writer offers his analysis in order to assist all concerned professionals to direct needed constructive criticism to Michigan's State Department of Education in this matter.

Tables 4A and 4B are identical in all respects except one -- the matter of which students' scores are to be included. Table 4A is used "...for reporting Achievement Batteries or Reading Tests **only for students who took both pre- and post-tests.**" Table 4B is used "...for reporting Achievement Batteries or Reading Tests for all students who took either/or both pre- and post-tests." Thus, schools must submit a set of data on the

children for whom there are both pre-test and post-test scores. Then, the schools must retabulate all data to include the few children who took a pre-test but no post-test, and the very few children who took no pre-test but did take the post-test.

The purpose of a pre-test/post-test design is to provide evidence of change in behavior over a given period of time. Table 4A requires the reporting of data from both pre-testing and post-testing for every child included. This procedure is consistent with pre-test/post-test design, and Table 4A is thus defensible. Table 4B require the inclusion of data for those children who were available for only one of the testing sessions. Data from only a pre-test or a post-test does not fit into the pre-test/post-test design. Including such data from a single testing violates the design pattern, and therefore contaminates the total tabulation of data. Hence, Table 4B, as presently defined, is not defensible. It is reasonable to ask for the number of children who did not take both the pre-test and the post-test, but that information can be supplied better and more easily elsewhere in the school's total report.

Now, let us turn to the categories of information required in Table 4A. The category head-

ings do not make use of co-identifying numerals or letters. Since Arabic numerals and the capital letter "A" already appear elsewhere on the page, this writer has chosen, for convenience, to assign Roman numerals to the seven categories. In addition, since Category VII has four sub-sections, the writer will at the appropriate point co-identify each sub-section with a lower-case letter.

I. Name & Form No. of Test & Each Subsection

This information is essential. However, having a column devoted to this information encourages a variety of listing procedures. Thus, some schools may organize by type of data (as, reporting all Reading data, then all Math data), whereas other schools may focus upon one grade level at a time). To avoid confusion, it would be better to ask schools to report each different type of data on separate pages (as, Reading testing on one page, Math testing on the next page, etc.). If this suggestion is implemented, then using Column I for Category I could be eliminated, for the information on type of testing could be placed on a line after the name of the Table (as, "Table 4A, Standardized Test Results, Reading").

II. Month & Year Administered (Pre-Test, Post-Test)

This information is also essential. If the precise dates of the testing sessions are deemed to be important, then they can

be mentioned elsewhere in the total report. As was the case for Category I, the presence of a column for this Category does invite a mixture of listings. For example, a school might inter-mix two or three time periods, such as a regular school year, a part of a regular school year, and a special summer session. While such inter-mixing might occur infrequently, it would seem to be desirable organizational procedure to ask schools to report only equivalent time periods on each page.

III. Grade

Since the norms for tests are typically developed and reported by grade level rather than by age, it is realistic to use grade level here.

IV. Number of Students

Since ESEA Title I projects are intended to be for the special benefit of certain specified numbers of disadvantaged children within any given school district, it is consistent with public policy to ask schools to report the number of students serviced.

V. Raw Score Mean (Pre-Test, Post-Test)

It is useful to know the central tendency of growth or power. However, the mean raw score is not useful information when requested alone, as raw scores have meaning only when related to some normative scaling of growth or power. Unfortunately, there is no provision, now or later, for the mean raw score to be converted into some

meaningful scale. (A number of school personnel who attended the meeting in June, 1967, have personally told the writer that they were given explicit oral directions to report only the mean raw score, and especially not to report a mean grade score. This is very puzzling, particularly in view of notation No. 2 at the bottom of this form: "If not raw score, indicate type score reported for each test.")

In addition, the reporting of mean raw scores invites misinterpretations. For example, the mean raw score of a vocabulary sub-test may be smaller than the mean raw score of a Comprehension sub-test, and yet the smaller mean Vocabulary raw score could easily reflect a higher level of growth or power than does the somewhat larger mean Comprehension raw score.

Category V could be converted into a useful category by relabeling it "Mean Grade Score" or "Mean Grade-Equivalent Score."

VI. Raw Score Standard Deviation (Pre-Test, Post-Test)

The standard deviation of a set of scores is the positive square root of the variance of that set of scores. The standard deviation can be used as an index of the dispersion (spread, or variability) of a distribution of data. If the purpose for requiring the raw score standard deviation to be reported is to provide an index of dispersion, then the purpose is noble, but

not fulfilled. Again, we should remember that raw scores have meaning only when converted into some normative scaling of growth or power. Hence, computing the raw score standard deviation in order to provide an index of dispersion is meaningless, and therefore unnecessary computation. It should be further noted that with small-size samples (which we could reasonably expect from most school systems) the range of grade scores could be used as an index of dispersion. The range of grade scores is much easier for school personnel to identify and report than is the standard deviation.

If the purpose for requiring the raw score standard deviation to be reported is to provide the information needed to test for a significant difference between the pre-test and the post-test mean raw scores, then the wrong information has been required. Instead, the variance is needed.

In short, there does not appear to be valid justification for requiring schools to report raw score standard deviations.

VII. Number of Students Scoring

- a. 25th Percentile & Below
- b. 26th to 50th Percentile
- c. 51st to 75th Percentile
- d. 76th Percentile & Above

A percentile specifies the point below which a given per-

centage of the scores fall. One may suppose, then, that the purpose of this tabulation is to show evidence regarding individual change (hopefully improvement) in growth or power. Percentiles are one kind of scaling frequently used to indicate rank position with respect to other children of the same grade level. Percentiles are useful for that purpose, and for putting scores from different tests on a comparable basis. It should be noted, however, that percentiles do not necessarily indicate a particular level of growth or power -- merely a comparative level of growth or power. The same is true of standard scores and stanines. Only a grade score indicates a particular level of power, or probable learning level which can be translated into instructional materials' graded levels. In addition, percentiles over-emphasize the relative position of children who are within, or close to, grade level in achievement. At the same time, percentiles underestimate the differences between individual children who are much above, or much below, grade level in achievement. Thus, a child who pre-tests 07th percentile and post-tests 17th percentile has actually made more improvement than has a child who pre-tests 17th percentile and post-tests 27th percentile, even though each child advanced the same number of percentile points. Both of the cases show a positive change in achievement, but the design of the sub-

sections of Category VII would reveal only one case of improvement. The case of the lesser improvement would be revealed, and the case of the greater improvement would be concealed.

If the kind of tabulation involved in Category VII is to be meaningful, then more sub-sections should be used, and the designation of percentile ranges within sub-sections should take into account the facts pointed out in the preceding paragraph. In this regard, it should be asked that the existing sub-sections contain equal percentages of children -- 26%, 25%, 25%, and 24% respectively. Such an arrangement is not typical in a four-cell division. It would seem that in preparing the existing four-cell arrangement someone has confused percentiles with percentages. If such be the case, then the confusion should be corrected before any multi-cell arrangement is substituted for the existing, and highly limited, four-cell one.

The cumbersome, time-consuming procedure of tallying percentile ranges can be avoided by reporting two facts: (1) the average gain in achievement, and (2) the percentage of pupils whose scores reflect a gain. If mean grade scores were reported under Category V, then one has only to subtract the pre-test mean grade score from the post-test mean grade score to obtain the mean gain for the group. To determine the percentage of pupils whose scores

reflect a gain, one has only to review the individual pre-test grade scores and to tally the instances in which a gain is indicated on the face of things. These procedures are far simpler for school personnel to carry out than would be the elaborate tallying required even under the existing arrangement of Category VII, not to mention the more detailed tallying required under some realistic modification of Category VII.

In short, the only valid justification for retaining Category VII seems to be an intent to put scores from different tests on a comparable basis.

In summary, the writer has offered the following constructive criticisms of Tables 4A and 4B:

1. Table 4A requires the reporting of data which are consistent with the rationale for pre-test/post-test design. Hence, Table 4A, as a Table, should be retained.

2. Table 4B requires the inclusion of data which do not fit a pre-test/post-test design. Hence, Table 4B should be discarded.

3. The Categories of information specified in Table 4A were evaluated as follows:

- a. Categories I, II, III, and IV are defensible, and should be retained. At the same time, several suggestions for the improvement of organizational procedures were offered.

- b. Categories V and VI are seen to be indefensible, and should therefore be discarded. Substitutions were recommended.

- c. Category VII is seen to be of very restricted value as presently arranged. Suggestions for improvement were made. In addition, it was pointed out that this Category is not needed unless there is a clear intent to put scores from different tests on a comparable basis.

To the preceding analysis, the writer now adds some recommendations for the improvement of Table 4A. There are several Categories which could, and probably should, be added to the Table 4A, as follows:

1. Show the percentage of pupils whose individual pre-test and post-test scores reflect some gain in achievement.

2. Show the average current gain (assuming positive change) in achievement for the duration of the project.

3. Show the average previous rate of gain in achievement for comparable time periods. This is easily computed from the pre-test mean grade score, and the computation permits the use of data on years of retention -- information neglected in the existing Table 4A. The average previous rate of gain in achievement may then be used as an estimate of expected current progress.

4. Show the average change in rate in achievement. This is the difference between average current gain and average previous rate of gain in achievement for a comparable time period. Each positive difference may be interpreted to indicate project effectiveness in the particular area. Each negative difference may be interpreted to indicate a lack of project effectiveness in the particular area. The computation of average change in rate of gain in achievement tends to control the factor of the influ-

ence of ongoing classroom instruction.

This presentation has offered a number of recommendations for strengthening and extending the existing guidelines for the evaluation of ESEA Title I projects. The writer respectfully submits his analysis to his colleagues in the Michigan Reading Association for their consideration and possible action.

(Dr. Elder is a member of the staff of The Child Study Center at Kent State University in Kent, Ohio.)

1967-68 Memberships Are Due Now!

MEMBERSHIP APPLICATION MICHIGAN READING ASSOCIATION

Name _____
 Address _____
 City _____ State _____ Zip _____
 Position _____ Phone _____
 School _____ School Address _____
 City _____ State _____ Zip _____ *

MEMBERSHIP YEAR
1967-68

CHECK ONE

MRA Dues: \$2.00

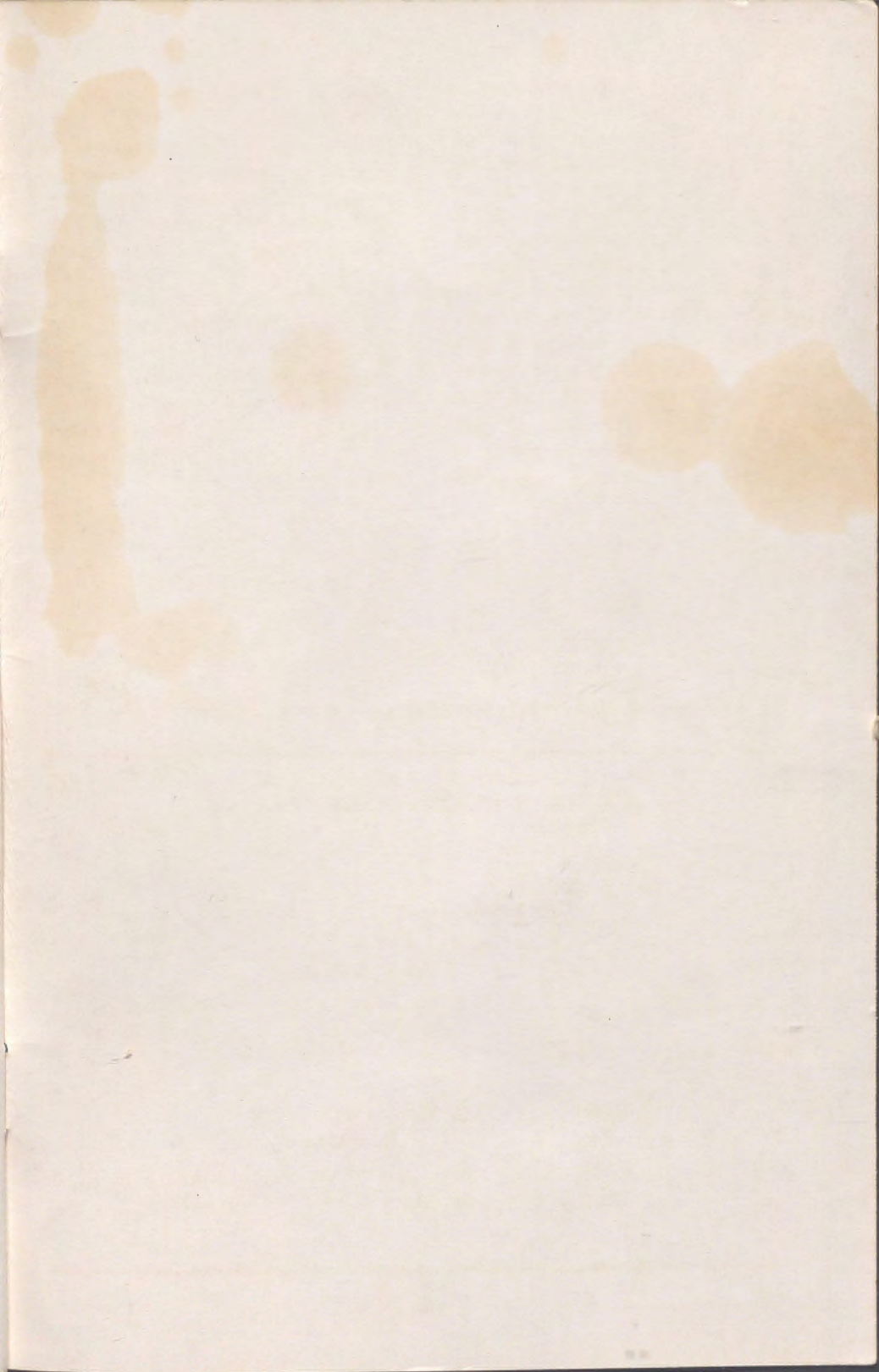
New member:

Check no. _____

Renewal:

MAIL TO: Sybil Michener, Treasurer
 Michigan Reading Association
 P.O. Box 665
 Flint, Michigan

*Required for mailing
 MRA Journal



The
Michigan Reading Journal

MARYGROVE COLLEGE

DETROIT, MICHIGAN 48221



Mrs. M. Dygert
1930 Sylvan Ave.
Grand Rapids, Mich. 48503

THIRD CLASS
MAIL