READING IN KINDERGARTEN?

A Critique of THE DENVER STUDY

DARLENE WEISBLATT MOOD
Assistant Professor, Department of Education
Eastern Michigan University, Ypsilanti

McKee and Brzeinski (1966) have recently reported the final results of a six year study. They indicate a high degree of success with an experimental program that included a formal program in "beginning reading" introduced in kindergarten. This was followed by an "adjusted program" in Grades One through Five designed to maintain and enhance the gains made as a result of the kindergarten experience.

With the current concern on the part of classroom teachers, especially in the lower elementary grades, for improving their techniques of teaching reading, this report is certain to stir a great deal of excitement. It is unfortunate that much of that excitement will have been generated by findings based on a weak research foundation.

Let us review briefly the findings of McKee and Brzeinski in their study conducted in Denver (Colo.) from 1960-66. They reported that beginning reading could be taught in the kindergarten with success; that gains made in the kindergarten could be maintained through Grade Five with a program using appropriate materials and an accelerated pace compared to the regular program; that gains made in reading tended to be accompanied by gains made in some other academic areas; that introducing the experimental program in the kindergarten did not increase the incidence of certain physical or emotional disabilities; that gains made in kindergarten could not be maintained if not followed by an "adjusted," accelerated program; and that greater achievement was made by students who started the experimental program in the first grade than those who were in the regular program in Grades One through Five, regardless of the nature of their kindergarten experience.

The results, as impressive as they seem, however, must be considered with extreme caution.

Evaluation of the Research Design: The Kindergarten Program

In any experimental research, the results are only as meaningful as the design of that research allows them to be. This is an area of evaluation which few teachers or school administrators are able to deal with effectively. Typically the educator who reads research must assume that the research was carefully conducted and results accurately interpreted. The educator would be apt to miss, therefore, the fact that this study by McKee and Brzeinski violates some of the most basic rules of good research design and good research reporting.

Fundamental to a research design which calls for the comparison of an experimental and control group is the necessity that all conditions under which the two groups are studied be held constant, except for the variable being examined which is manipulated in the experimental group and not in the control group. When these conditions are met, an observed difference between the two groups can be presumed to be the effect of the manipulated variable.

In the kindergarten program of the Denver study under consideration, at least two variables were manipulated without any means of separating their effects. The two variables were: (a) the method, and (b) the materials.

The experimental group was given prescribed training in basic reading skills for 20 minutes per day in kindergarten, followed by an accelerated reading program in Grades One through Five where they moved through the series of basal readers at their own pace. The kindergarten materials they used were developed by the senior author of the Denver study and an associate (McKee and Harrison, 1960a).

The control group followed the regular programs in kindergarten through Grade Five. While the report is somewhat unclear as to the nature of the "regular program" (the reader is constantly referred to the Denver Public School curriculum guides), it is specifically described as involving a kindergarten program in which the learning of basic reading skills is incidental, and a reading program in Grades One through Five which limited the children to the use of basal readers intended for their grade level. The materials were the basal series (unidentified) regularly used by the school system.[1]

Considering the two groups described, and the resulting superior achievement of the experimental group, to what can the favorable results be attributed? McKee and Brzeinski refer to the accelerated program—the method—especially when begun in kindergarten, as the primary determinant of the greater academic achievement in the experimental group.

Are the results, on the other hand, to be attributed to the experimental materials? There is an implication that these materials are especially well planned, and that the same results

[1] There were actually four groups in the study. Only the true experimental group (experimental program throughout study) and the true control group (regular program throughout study) are described here, but criticisms of design can be extended to comparisons with the other two groups as such.

might not be achieved with other available materials. If this were not so, why would the authors confound and confuse the results of their research by using materials other than those already used by the Denver Public Schools?

Still a third interpretation could be made; namely, that the results are due to the *combination* of the method and the materials. Again, this is impossible to determine, because the research design used does not allow the experimenters to separate the effects of the method from the effects of the materials.

Finally, it should be noted that the test used to measure kindergarten achievement was developed by McKee and Harrison (1960b) and corresponds to their program of learning to which the experimental group was exposed. The superior performance of the children in the experimental group may have resulted from their having an appropriate "response set"; that is, they were familiar with the type of questions being asked and the type of answers which were expected.

The Later Grades

Another fundamental aspect of research which must be seriously questioned is that of random sampling. From the descriptions available, the original selection of the groups seems to have met the criteria of randomness. A loss of subjects in any longitudinal study would be expected and presumably distributed among the various groups, maintaining randomness.

In the Denver study, a higher attrition rate was anticipated in the experimental groups which therefore began the study with 1250 subjects, compared with 750 subjects in the control group. At the conclusion of the study, these N's had dropped to 759 and 225 respectively, representing a 39 percent loss in the experimental group and a 70 percent loss in the control group. The difference in attrition rates is highly significant statistically.[2]

The questions that arise are these:

1. Why was a higher attrition rate expected in the experimental groups?

2. Why was the observed attrition rate so much higher in the control group?

3. What were the relative rates of loss of subjects from year to year?

4. Was any attempt made to examine the data on subjects who dropped out to determine the possible existence of some common factors which might bias the results of the study?

The investigators fail to deal with these questions, leaving the research-wise reader to wonder whether the groups of subjects who completed the study could actually be considered random samples—a requirement basic to any statistical analysis.

There also exists the possibility of a "Hawthorne" effect accounting for the findings. The authors suggested that steps were taken to minimize these effects. It is highly questionable whether in-service meetings with teachers would actually compensate for the greater enthusiasm of teachers in the experimental conditions, if such enthusiasm existed. Certainly the duration of the study is not an adequate control, as the investigators suggest, if the subjects had new teachers for each grade.

[2] Chi square test of significance was equal to 176.99, which is significant beyond the .001 level.

Even when the groups entered Grade Five, beginning the sixth year of the study, the fifth grade teacher was being directly involved in the study for the first time.

The Research Report

The most frustrating problem to anyone attempting to evaluate carefully the complex details of the report of the Denver study is the incomplete and misleading presentation of the results. Any study of this magnitude must of necessity be selective in the material that actually appears in the report. The selection in the McKee and Brzeinski report, however, is careless and extremely difficult to interpret.

Analysis of variance and covariance was indicated to be the most appropriate statistics, but few of these analyses are discussed. Data presented in tables do not always correspond to the analyses reported in succeeding paragraphs purporting to interpret those tables. The values of p which they report indicate a vacillation between one- and two-tailed tests. In reporting achievement test results, it is impossible to determine the actual performance levels of the groups being compared.

Scores of separate tests are combined into a single score for which no normative data are available. The size of the samples for the various comparisons is never given. In many instances, no averages are reported, rendering impossible an evaluation of the groups by the concerned reader. One must either accept the misleading inference that statistical significance is the same as educational significance—or reject the data entirely.

Because of the careless reporting of this study, the possibility of replicating the experiment is also seriously hampered. An independent researcher would be unable, on the basis of the report, to establish a comparable experimental program. The description of that program is vague. Children are to be allowed to progress at their own pace in the experimental program—is this individualized reading? Or did the children move through the series in groups? Or total classes?

The list of important questions this study leaves unanswered is alarming.

Theoretical Implications

The pedagogical aspects of the Denver study have intentionally been deemphasized. There can be no argument with the premise that innovative teaching methods need to be sought. One could question, however, whether the Denver study is innovative.

It is for the specialists in reading to decide whether the experimental kindergarten program used in this study is appropriately called "beginning reading" or merely constitutes another set of "reading readiness" materials. The authors, however, indicate that the ability to teach beginning reading in the kindergarten is not a new concept, but has already been established, and that they are primarily concerned with the long-range effects.

Proponents of individualization in reading may be surprised to see a program in which the child is allowed to proceed at his own pace called innovative. In theory, if not in practice, this educational concept has been recommended by some educators for well over

50 years. Under any circumstance the theoretical hypotheses have not, because of research weaknesses described above, been adequately evaluated.

In summary, the poor research design and reporting of the McKee and Brzeinski study exemplify a situation which occurs too frequently in the field of educational research. This is unfortunate, indeed, because conscientious professional educators are groping for methods to meet the challenges of modern education. They look to research such as the Denver study for answers, but are generally not prepared to recognize the weaknesses which may reside in such a research report.

The researchers, who should recognize the limitations of their work, seem to indicate no obligation to point these out —to be somewhat self-critical. Perhaps an investigator not so directly involved with the experimental materials, which represent a vital factor in the study, might have been more objective in evaluating his results.

Caveat emptor—let the buyer beware —is the word of caution that the consumer of educational research might do well to remember.

References

Paul McKee and Joseph E. Brzeinski. "The Effectiveness of Teaching Reading in Kindergarten." Cooperative Research Project No. 5-0371. Denver: Denver Public Schools and Colorado State Department of Education, 1966.

Paul McKee and M. Lucile Harrison. *Program in Skills Basic to Beginning Reading.* Boston: Houghton Mifflin Company, 1960a.

Paul McKee and M. Lucile Harrison. *Test of Skills Basic to Beginning Reading.* Boston: Houghton Mifflin Company, 1960b.

The Supervisor: Agent for Change in Teaching

Edited by

James Raths and Robert R. Leeper

This new ASCD booklet presents eight papers in which eminent scholars share their research and their views concerning supervision and the work of the supervisor.

Contents

Helping Teachers Change
James B. Macdonald

An Interaction Model Applied to Supervision
Robert D. Boyd

Effects of Supervisor's Knowledge of Student Teacher Personality Evaluations
Thomas A Ringness

Interaction Analysis as a Feedback System in Teacher Preparation
Edmund J. Amidon and Evan Powell

Professional Integration and Clinical Research
Ted W. Ward

Strategies for Instructional Change: Promising Ideas and Perplexing Problems
Ben M. Harris

Challenges for Supervisors
Leslee J. Bishop

Organizational Factors in Supervision
James G. March

138 pages; $3.25

Order from:
ASCD
1201 Sixteenth St., N.W.
Washington, D.C. 20036

Copyright © 1967 by the Association for Supervision and Curriculum Development. All rights reserved.