

The MRT vs. Ratings of Kindergarten Teachers as Predictors of Success in First Grade¹

JOHN R. BOLIG*
GERALD O. FLETCHER

MANY educators have felt that teacher evaluations of pupils in kindergarten are as good as, if not better than, standardized test results in placement of children, or predicting their success in first grade. Research in this area for the most part has been centered on identifying particular developmental and/or academic tasks which may be effective as predictors of first grade achievement.

Studies by Ilg and Ames have indicated that school placement of children needs direct and continuing observation. Their Weston Study revealed that correspondence between teacher ratings and developmental examinations was high in the lower grades, but agreement decreased with older children and higher grade placement.

Dolores Durkin, Marion Monroe, and others concerned with child development have appealed to educators to use every available means to evaluate readiness for academic tasks such as reading and writing,

¹ The work presented or reported herein was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

including the use of standardized readiness tests.

Serious questions had been presented by parents, teachers, and administrators of the North Penn School District, Lansdale, Pennsylvania, as to the validity of the Metropolitan Readiness Test as an instrument for predicting first grade success. Advancement to first grade and placement into ability groups were frequently based upon the MRT² score, which was often in conflict with the opinion a parent or a teacher had about a child.

In an effort to present additional research relevant to predicting success in first grade, the North Penn School District initiated a study comparing kindergarten teacher ratings of pupils with the results obtained from the Metropolitan Readiness Test.

Further, it was considered that if teachers could predict achievement as well as, or better than, the MRT, then administration

² Metropolitan Readiness Test; hereafter referred to as MRT.

* John R. Bolig, Research Director, Del Mod System, Dover, Delaware; and Gerald O. Fletcher, Coordinator of Reading, North Penn School District, Lansdale, Pennsylvania

of this test (a time-consuming and costly project) would not be necessary.

Hypothesis

It was hypothesized that kindergarten teachers could predict first grade success as well as could the MRT. It was further hypothesized that they could do this for boys as well as they could for girls.

Procedure

A two-year study was initiated in April 1968. The first year was a pilot study, the second year became the basis of this report. It is noteworthy that the second year of the study confirmed the pilot study results.

Each kindergarten teacher rated every child in her class prior to administering the MRT (Form A). These ratings and the scores on the MRT were collected and stored by the experimenters.

One year later each first grade teacher rated every child in her class prior to administering the Stanford Achievement Test (Primary I Battery, Form W). These ratings and scores were collected and matched with the kindergarten ratings and scores of these children.

Statistical analysis is based upon those children with complete sets of ratings and scores for the two years of school. A total of 906 children were tested in kindergarten in April 1969. Of these, 625 were in first grade in April 1970. The loss of children can be attributed (largely) to their entrance into nonpublic schools for grade one.

Rating Procedure. Each kindergarten teacher rated her children on each of the six skills measured by subtests of the MRT. The subtests rated were: Verbal Concepts, Visual Perception, Listening, Alphabet, Numbers, and Copying.

A scale from 1 to 5 was used for each rating, and each rating was recorded on a 3 x 5 index card. The child's name was on one side of the card, his rating on the other. A description of each rating was provided for each teacher on separate sheets of paper.

For example: "Copying: the child has the ability to visualize sizes, shapes, and forms, and to reproduce them appropriately." Sheet one carried this description and the rating "1—Almost Never." Sheets two through five carried the description and the ratings: "2—Seldom," "3—Sometimes," "4—Frequently," and "5—Almost Always," respectively.

This was repeated for each subtest. The sheets were taped together for each multifold. These multifolds and the response cards were distributed to each teacher.

The following directions accompanied the materials:

Step 1. Place each pupil's name on a response card; identify school, teacher, and attendance session as indicated.

Step 2. Spread out the multifold paper for "Copying."

Step 3. Shuffle cards. Sort cards, name side up, placing each pupil's card on the description which *describes him best*.

Step 4. Turn cards over, pile by pile, and record the indicated number on the appropriate space on each card.

Proceed to the second multifold paper "Numbers" and repeat Steps 2, 3, and 4.

General Directions. Plan to complete any given round at one time. There is no need to do all six rounds at once.

Shuffling the response cards and rating subtests separately was intended to force each teacher to consider the skills of each child individually. Placing ratings on the reverse of the response card was intended to ensure independence of ratings.

The first grade teachers followed similar instructions one year later when they rated the children on the six skills measured by the SAT³: Word Reading, Paragraph Meaning, Vocabulary, Word Study Skills, Spelling, and Arithmetic.

There were 20 kindergarten teachers and 36 first grade teachers involved in rating children and administering tests. Several of the kindergarten teachers taught morning and afternoon sessions.

³ Stanford Achievement Test; hereafter referred to as the SAT.

Statistical Procedure. The ratings for each teacher were normalized. Converting the ratings into *t* scores served the purpose of equalizing the rating severity of the various teachers. This was done for both grades. The normalized ratings for the six subtests were summed, as were the subtests for the MRT and SAT. These gross measures became the basis for subsequent statistical analysis. Each child had four measures: an MRT score, an SAT score, a kindergarten rating, and a first grade rating.

Means, standard deviations, coefficients of correlation, and multiple coefficients of correlation were computed for the four variables. Appropriate statistical tests were applied to analyze differences between predictors and between sexes.

Results

The *z* test for independent sample means was used to compare boys and girls on the four variables. These comparisons are shown in Figure 1.

	MRT Score		Kindergarten Rating		SAT Score		First Grade Rating	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Girls	61.58	13.43	.6222	4.20	154.66	41.92	.2598	4.69
Boys	61.57	12.64	-.0147	3.97	151.56	41.56	-.2858	4.50
$M_G - M_B$.01		.6369		3.10		.5456	
$S_{M_G - M_B}$	1.07		.328		3.35		.368	
<i>z</i> ratio	0.01		1.94		.93		1.48	

Figure 1. Means and Standard Deviations for the MRT, the Kindergarten Ratings, the SAT, and the First Grade Ratings for 315 girls and 310 boys and Results of *z* Tests of Significance

Initially the boys and girls were quite similar as measured by the MRT, but the teacher ratings for kindergartners favored the girls. The difference in the kindergarten teacher ratings approached significance ($P = .052$).

The SAT scores for the first graders were not significantly different. The difference in the first grade ratings was not significant, but these also favored the girls.

The product moment coefficient of correlation for each variable with each other variable was computed to determine the efficiency of kindergarten teachers and the MRT as predictors of first grade success. First grade success is defined as either the SAT score or the first grade rating.

These coefficients of correlation are shown in Figure 2.

		Kindergarten Ratings	SAT Scores	First Grade Ratings
MRT Scores	Girls	.72	.75	.64
	Boys	.72	.68	.51
	Total	.76	.72	.58
Kindergarten Ratings	Girls		.66	.62
	Boys		.55	.47
	Total		.61	.55
SAT Scores	Girls			.76
	Boys			.72
	Total			.74

All coefficients reported are significantly different from zero beyond the 0.01 level.

Figure 2. Correlation Matrix: Coefficients of Correlation for the MRT, Kindergarten Ratings, the SAT, and First Grade Ratings by Sex and for the Total Sample

The *t* test for comparing two correlation coefficients involving a common variable was used to test differences between MRT scores and kindergarten ratings and the two first grade variables.

Predicting SAT Scores. The MRT is more highly related with first grade success on either criterion variable than are the ratings of kindergarten teachers. The coefficients of correlation between the MRT and the SAT were .09, .13, and .11 better than kindergarten teacher ratings for the girls, the boys, and the total sample, respectively. These differences are significant at or beyond the 0.01 level.

Predicting First Grade Ratings. The coefficients of correlation between the MRT and first grade teacher ratings were .04, .02, and .03 better than kindergarten teacher ratings for the boys, the girls, and the total sample, respectively. These values are not statistically significant.

Next, the MRT scores and kindergarten ratings were examined by sex. The *z* test for independent sample correlation coefficients was used to compare coefficients of correlation between the boys and the girls.

Sex Differences. The MRT was more highly related with the success of girls on both the SAT and first grade teacher ratings. The MRT coefficients of correlation were .07 better for girls on the SAT, and .13 better for girls on the first grade ratings. The differences were significant at the .05 level and the .01 level, respectively.

The kindergarten teacher ratings were also more highly related with the success of girls on the two first grade variables. The kindergarten rating coefficients of correlation were .11 better for girls on the SAT, and .15 better for girls on first grade ratings. Both differences were significant at or beyond the .01 level.

Multiple Coefficients of Correlation. While the MRT appears to be a better predictor of first grade success, whether defined as a score on a test or as an evaluation of a teacher, it is also true that kindergarten teacher ratings are significant predictors of first grade success. Combining these variables provides only a slightly more efficient predictor of first grade success. The multiple coefficients of correlation for predicting the SAT score and the first grade rating are shown in Figure 3.

	Boys	Girls	Total
SAT	.686	.771	.729
First grade rating	.531	.684	.612

All values are significantly different from zero beyond the .01 level.

Figure 3. Multiple Coefficients of Correlation of MRT Scores and Kindergarten Teacher Ratings with SAT Scores and First Grade Teacher Ratings

Combining the two kindergarten variables did not add much. The best predictor of first grade success in this study was the MRT score, and it is a better predictor for girls than it is for boys.

Conclusions

Initial differences between boys and girls in kindergarten are slight as measured by the MRT. The differences between ratings of boys and girls seem to indicate a teacher bias favoring girls. The use of the MRT score as a predictor of first grade success may

be preferable for either sex, but it is definitely preferable for boys.

The use of kindergarten ratings by teachers to predict first grade success for girls is reasonable, but kindergarten ratings by teachers for boys are not nearly as reliable an indicator of first grade success. They do not agree with either first grade variable as well as does the MRT score.

At the end of first grade there is not a significant difference between boys and girls as measured by the SAT. Since this difference is not significant, it may reflect the effects of teacher bias as evidenced by the lower ratings given boys by teachers in both grades.

Discussion

The hypotheses were not confirmed. The MRT as a predictor of first grade success for both girls and boys is as good as or better than the ratings of teachers on the six skills the MRT purports to measure. This is true whether the MRT is used as a predictor of scores on the SAT, or as a predictor of how a first grade teacher will rate children on the six skills purportedly measured by the SAT.

References

- Dolores Durkin. *Children Who Read Early*. New York: Teachers College Press, Columbia University, 1966.
- Frances L. Ilg and Louise Bates Ames. *School Readiness*. New York: Harper & Row, Publishers, 1965.
- Marion Monroe. *Children Who Cannot Read*. Chicago: University of Chicago Press, 1932.
- H. Alan Robinson. "Reliability of Measures Related to Reading Success of Average, Disadvantaged, and Advantaged Kindergarten Children." *The Reading Teacher* 20: 203-209; 1966.

Test References

- Gertrude H. Hildreth, Nellie L. Griffiths, and Mary E. McGauvran. *Metropolitan Readiness Test (Form A)*. New York: Harcourt Brace and World, Inc., 1969.
- Truman L. Kelley, Richard Madden, Eric F. Gardner, and Herbert C. Rudman. *Stanford Achievement Tests (Primary I Battery, Form W)*. New York: Harcourt Brace and World, Inc., 1966. □

Copyright © 1973 by the Association for Supervision and Curriculum Development. All rights reserved.