R E P O R T R E S U M E S

ED 019 205 RE 001 227

A COMPARATIVE STUDY OF THE VALIDITY OF THE BOTEL READING INVENTORY AND SELECTED STANDARDIZED TESTS.

BY- BOTEL, MORTON

PUB DATE 26 APR 68

EDRS PRICE MF-$0.25 HC-$0.84 19P.

DESCRIPTORS- *TEST VALIDITY, *AGE GRADE PLACEMENT, *INDIVIDUAL TESTS, STANDARDIZED TESTS, SILENT READING, ORAL READING, *READING LEVEL, READING DIAGNOSIS, PENN VALLEY SCHOOL, BOTEL READING INVENTORY,

THE EXTENT OF THE RELATIONSHIP OF THE BOTEL READING INVENTORY, FORMS A AND B, AND SELECTED STANDARDIZED TESTS TO THE INSTRUCTIONAL READING LEVEL OF PUPILS IN GRADES 1 TO 6 WERE INVESTIGATED TO PROVIDE EVIDENCE CONCERNING THE VALIDITY OF THE TESTS. THE COOPERATIVE CHECKOUT WAS EMPLOYED AS A CRITERION. TEACHER OBSERVATION AND PUPIL ORAL READING FLUENCY WERE ALSO USED AS VALIDITY CHECKS. CORRECT GRADE PLACEMENT WAS BASED UPON EASE OF READING AT, BELOW, OR ABOVE GRADE LEVEL. ALL READING TEST SCORES WERE CORRELATED WITH THE CRITERION, THE INSTRUCTIONAL LEVEL OF THE PUPILS, AND THE BOTEL INVENTORY WAS CORRELATED WITH THE STANDARDIZED TESTS TO PROVIDE INFORMATION ON THE CONGRUENT VALIDITY OF THE INVENTORY. RESULTS INDICATED THAT, EXCEPT FOR THIRD GRADE, THE BOTEL READING INVENTORY MIGHT BE SUPERIOR TO THE STANDARDIZED SILENT READING TESTS USED IN THIS STUDY IN PLACING PUPILS AT THEIR CORRECT INSTRUCTIONAL LEVELS. IT WAS FOUND THAT WHILE THE STANDARDIZED TESTS OVERPLACED MORE PUPILS THAN DID THE BOTEL READING INVENTORY, MANY PUPILS WERE EITHER CORRECTLY PLACED OR UNDERPLACED BY THE STANDARDIZED SILENT READING TESTS RATHER THAN OVERPLACED. IT WAS NOTED THAT SOME ABLE PUPILS WERE LIMITED TO GRADE LEVEL ON BASAL READING MATERIAL OR, AT MOST, TO ONE GRADE LEVEL BEYOND GRADE PLACEMENT, DESPITE THE FACT THAT THEIR PERFORMANCE IN ORAL READING AND COMPREHENSION WAS ALMOST PERFECT. REFERENCES AND TABLES ARE PROVIDED. THIS PAPER WAS PRESENTED AT THE INTERNATIONAL READING ASSOCATION CONFERENCE (BOSTON, APRIL 24-27, 1968). (MC)

A COMPARATIVE STUDY OF THE VALIDITY OF THE BOTEL READING INVENTORY AND SELECTED STANDARDIZED TESTS*

by

Morton Botel

PROBLEM

Kerlinger has observed that often "Tests are...used as substitutes for perhaps less convenient, more difficult, or otherwise cumbersome modes of measurement and evaluation." (5)

The problem of testing pupils to find their instructional levels in reading is a case in point. Ideally such a determination should be made directly in the materials themselves. Given the fact that basal reading materials are scaled into carefully controlled sequences and given the generally accepted notion that pupils ought to be reading with a certain minimum degree of comprehension and oral reading fluency, such an evaluation can be readily made if there is sufficient time. Each pupil would read, orally and silently, random passages at a number of levels of difficulty until he performed with the required degree of success. But the question of convenience and difficulty for the classroom teacher is very much a problem. With 25 to 35 pupils in a class, many teachers believe that there is insufficient time to prepare such

*Research Report delivered at the Thirteenth Annual Convention, International Reading Association, Boston, Massachusetts, April 26, 1966.

a test and to evaluate pupils in this ideal way.

Where the performance of pupils cannot be evaluated directly in their materials to find their level, what test should be substituted? For some, this question is answered simply: Test them on an informal reading inventory. An informal reading inventory is, after all, a series of samples from a particular series of graded basal readers. By having pupils read silently and aloud and by having them respond to questions based on the reading, it is possible to determine the levels at which the desired fluency is obtained by the pupils. They are then given materials to read at this level.

There are at least four problems, however, that need to be considered concerning the validity of an informal reading inventory. The informal reading inventory may not be comprised of a good set of samples. The samples may be taken from a series not actually to be used for instruction. The questions used to measure comprehension may vary in difficulty depending on the formulator. The notions of the examiners may vary with respect to the degree of fluency and comprehension needed to establish the pupil's instructional level.

The four problems posed above are validity questions which have usually been ignored in developing and using informal reading inventories. Yet, such queries suggest the complexity of the problem of placement of pupils if anything other than the instructional materials themselves are used to place pupils at the right level. Furthermore, they indicate the need to establish the extent of validity for instruments that purport to determine the instructional level of pupils. The informal reading inventory may not be exempted from the requirement of other tests that its

validity (and reliability) must be determined.

PURPOSE OF THE STUDY

It was the main purpose of this study to determine the extent of relationship of the Botel Reading Inventory (Forms A and B) (2) and several standardized tests (the California Reading Test in Grade 2 (3), the average of the reading sub-tests of the Iowa Tests of Basic Skills in Grade 3 (4), and the STEP Tests in Grades 4-6 (6)) to the instructional reading level of pupils in grades 1 to 6, thus providing some evidence with respect to the concurrent validity of these tests. The instructional reading level or criterion in this study is the placement of pupils in their basal readers at the time the several tests were administered. The validity of the criterion is thoroughly discussed under "Procedures."

A second purpose of this study was to determine the relationship of the Botel Reading Inventory to the standardized silent reading tests used in this study, thus providing some evidence with respect to the congruent validity of these tests.

HYPOTHESES

1. There is no significant relationship between the Botel Reading Inventory raw scores and the criterion.
2. There is no significant relationship between the several standardized tests grade equivalent scores and the criterion.
3. There is no significant difference between the relationship of the Botel Reading Inventory with the criterion and the standardized tests with the criterion.

4. There is no significant relationship between the Botel Reading Inventory raw scores and the standardized reading tests grade equivalent scores.

5. There is no significant difference between the Botel Reading Inventory instructional level scores and the standardized silent reading test grade equivalent scores in their relative ability to place a pupil at his instructional reading level.

PROCEDURES

Establishing a Valid Criterion

According to Kerlinger "The single greatest difficulty of prediction validation is the criterion. Often criteria don't exist or validity is doubtful." (5) Certainly the same thing can be said of the criterion in concurrent validity studies. In connection with this study of the validity of the Botel Reading Inventory (Forms A and B) and various standardized reading tests, an attempt was made to carefully establish the most valid criterion possible. This was done by selecting a school for the validation study which carefully places its pupils in graded readers at their instructional levels. From the pre-primer level on, pupils are advanced from one reading level to another based upon their mastery of earlier levels. Mastery is established when the pupil can read aloud with 95 per cent or better fluency on the average and with 75 per cent or better comprehension on the average in the books they have completed. A description of this method will be found in How To Teach Reading (1). The procedure is called the Cooperative Checkout.

In the case of the Penn Valley School, used in this study, the principal, Dr. Dorothy Ingalls, had checked and approved the advancement

of each pupil from instructional level to instructional level. Each pupil's advancement was based on his observed fluency and comprehension in his basal readers and workbooks following the standards described above. In other words, pupils in the Penn Valley School may be said to be reading at their proper instructional levels.

To further establish the validity of this criterion, two further checks were made by the investigator. First, at his request, a record was kept by each teacher for one semester of the percentage of accuracy of pupils on all workbook exercises in their reading program. In every case the average pupil-performance in these workbooks pages (which were completed independently by the pupil) was 75 per cent or higher. Second, the investigator checked each pupil's oral reading fluency in stories they had not been exposed to immediately following those they were working on currently. In each case the pupil was able to read orally at sight with 95 per cent fluency or better.

While correct placement was confirmed for those pupils reading below grade level, those reading at grade level and above presented a problem. For, in some classrooms, pupils who could read above grade level were not advanced to these higher levels. And in no case were pupils allowed to read more than one grade level beyond grade placement in the basal reader program. (Able readers, however, were unrestricted in their choice of library books.) The effect of this restriction of placement in basal readers from the point of view of this study is to limit the effectiveness of the criterion for those reading at grade level and beyond.

Thus, for pupils who are reading below grade level and to a lesser extent for those reading at and above grade level we have established the

correctness of their placement in basal readers. This serves as a valid criterion for determining the relative power of several tests to predict the correct placement of pupils.

Two basic types of comparisons were made among the variables: (1) correlations and (2) matching. (All tests were administered by the classroom teachers in early February 1967.)

Correlations. All reading tests were correlated (using the Pearsonian, product-moment correlation) with the criterion described earlier -- the instructional reading level of the pupils. Further, the Botel Reading Inventory (Forms A and B) were correlated with the standardized tests to provide information on the congruent validity of the Botel Reading Inventory.

In the case of the criterion the score used was grade level in tenths of a year depending upon how far pupils had advanced in the Ginn readers which were used as basal in the Penn Valley School.

Matching. The instructional levels given by the Botel Reading Inventory (Forms A and B) and the grade equivalent scores of the standardized tests were compared with the criterion on the basis of the extent of match. That is, the extent to which each test placed pupils at, above, and below their instructional levels, was mapped out. The Botel Reading Inventory provides a table in which the instructional levels of the pupils are given. In the case of the standardized tests the following table was used to convert the grade equivalent score to an instructional reading level score:

Grade Equivalent Score	Instructional Reading Level
1.0-1.2	Pre-Primer
1.3-1.5	Primer
1.6-1.9	1^2
2.0-2.4	2^1
2.5-2.9	2^2
3.0-3.4	3^1
3.5-3.9	3^2
4.0-4.9	4
5.0-5.9	5
6.0-6.9	6
7.0 *	7

(Since the STEP tests do not provide a grade equivalent score, these had to be determined. The method used was to regard the mean score at each grade level given in the table of norms as having the grade equivalent designation for the time of the year the test was administered to the norming group. Intermediate points were then determined assuming average growth of pupils to be regular throughout the ten months in which school is in session.)

CORRELATIONAL FINDINGS

Correlations found between the Botel Reading Inventory (Forms A and B) and the standardized tests with the criterion are shown in Tables 1 and 2. It will be seen the tests were highly correlated with the criterion.

Table 1

Correlation[1,2] Of The Performance Of Pupils Using Their Raw Score On The Botel Reading Inventory (Forms A and B) With Their Reading Instructional Levels

Grade	N	$\bar{M}$ A	$\bar{M}$ B	SD A	SD B	r A	r B
2	103	136.1	140.4	58.1	60.0	.95	.95
3	127	184.2	187.6	36.9	36.2	.82	.86
4	97	217.2	222.1	21.8	20.2	.86	.84
5	111	229.1	230.1	19.8	19.7	.84	.86
6	103	239.5	240.0	14.5	14.6	.74	.73

Table 2

Correlation[1,2] Of The Performance Of Pupils Using Their Grade Equivalent Scores[3] On The Indicated Standardized Tests With Their Reading Instructional Levels

Grade	Test	N	$\bar{M}$	SD	r
2	California	103	2.9	.94	.92
3	Iowa	127	3.9	.92	.81
4	STEP	97	5.4	1.9	.66
5	STEP	111	6.1	2.0	.63
6	STEP	103	7.4	2.1	.51

[1]BMD03D - Correlation with item deletion - version of March 1, 1966. Health Sciences Computing Facility, UCLA.
[2]All r's are significant at the .001 level.
[3]The lack of raw scores for the standardized tests may have lowered the magnitude of the r's.

Inspection of the r's in Tables 1 and 2 suggest that the r's between the Botel Reading Inventory (Forms A and B) and the criterion might be significantly greater than the r's between the standardized tests and the criterion. These differences were compared by the "t"[1] test and the results recorded in Tables 3 and 4. (Since grade equivalent scores were used for the standardized tests rather than raw scores as in the case of the Botel Reading Inventory, this may have resulted in a lower magnitude of r's for the standardized tests than was actually found.)

[1]Hotelling formula -- cited in J. P. Guilford, Fundamental Statistics in Psychology and Education, New York: McGraw Hill, 1965, p. 190.

Table 3

Comparison Of r's Between (1) Botel Reading Inventory Form A And The Criterion And (2) The Standardized Reading Tests And The Criterion

Grade	N	Botel Form A	Standardized Silent Reading Test	t
2	101	.946	.922	2.126[1]
3	127	.820	.807	.479[2]
4	97	.863	.657	4.886[3]
5	111	.837	.633	4.315[3]
6	103	.737	.510	3.600[3]

Table 4

Comparison Of r's Between (1) Botel Reading Inventory Form B And The Criterion And (2) The Standardized Reading Tests And The Criterion

Grade	N	Botel Form B	Standardized Silent Reading Test	t
2	101	.951	.922	2.669[3]
3	127	.858	.807	1.924[2]
4	97	.840	.657	4.088[3]
5	111	.859	.633	5.216[3]
6	103	.728	.510	3.709[3]

[1]Significant at the .05 level.
[2]Not significant.
[3]Significant at the .01 level.

The results of this comparison indicate that except for grade 3, in which no significant difference was found, all comparisons show that the Botel scores relate more closely to the criterion than the standardized silent reading test scores. Seven of these differences are at the .01 level of significance, one is at the .05 level of significance, two are not significant at the .05 level.

Correlations found between the Botel Reading Inventory and the standardized tests are shown in Table 5. It will be seen that the correlations are highest in the lower grade levels ranging from .93 in second grade to .55 in grade 6. That is to say the lower the grade level the higher the concurrent validity.

Table 5

Correlations Of Botel Reading Inventory (Forms A And B)
With Selected Standardized Reading Tests In Grades 2-6*

Grade	N	Standardized Silent Reading Tests	Botel Form A	Botel Form B
2	103	California Reading Test	.93	.93
3	127	Iowa Reading Test	.80	.81
4	97	STEP-Reading	.65	.65
5	111	STEP-Reading	.60	.59
6	103	STEP-Reading	.55	.55

*All r's are significant at the .01 level of confidence.

Table 6 shows how the scores of pupils on the Botel Reading Inventory (Forms A and B) and the standardized reading tests match the actual placement of those pupils in grades 2-6 who are reading below grade level. This group was selected for study for two reasons. First, as we have noted, the group reading in grade level texts or above frequently were able to read at higher levels. Second, the slower readers have been believed in the past to be overplaced in reading by standardized checks. Verification of this notion could be checked by this comparison.

Table 6

Percentage Of Pupils In Readers Below Grade Level Placed Correctly, Underplaced, And Overplaced By The Botel Reading Inventory (Forms A And B) And Standardized Silent Reading Tests Using Placement By Teacher As Criterion (February 1967)

Grade	N	Test	Underplacement -2	-1	On Level	+1	Overplacement +2	+3	+4
2	22	Botel Form A			82	18			
2	22	Botel Form B			68	32			
2	22	Calif.		4	18	64	14		
3	25	Botel Form A		12	80	8			
3	25	Botel Form B	4	12	80	4			
3	25	Iowa		20	44	24	12		
4	24	Botel Form A		13	75	12			
4	24	Botel Form B			67	25	8		
4	24	STEP	13	17	17	25	25		3
5	26	Botel Form A		8	88		4		
5	26	Botel Form B		4	88	8			
5	26	STEP	4	8	37	23	27	4	
6	24	Botel Form A	13	21	62	4			
6	24	Botel Form B	8	29	58		5		
6	24	STEP	4	33	25	17	13	4	4

Since inspection of Table 6 suggests that the Botel Reading Inventory (Forms A and B) more frequently place pupils at their instructional reading level than do the standardized tests, the significance of each difference was tested using Chi Square (including Yates Correction).[1] This data was grouped in two categories: on level and off level in making these comparisons. Tables 7 and 8 summarize the significance of these differences for pupils reading below grade level.

[1]Formula (11.10) in J. P. Guilford, Fundamental Statistics in Psychology and Education, New York: McGraw Hill, 1965, p. 240.

Table 7

Relative Ability Of Botel Reading Inventory (Form A)
And Standardized Silent Reading Tests To Place
Pupils On Their Instructional Reading Levels

Grade	N	Botel A		Stand. Test		Chi Square
		on	off	on	off	
2	20	18	4	4	18	15.364[1]
3	35	20	5	14	11	5.435[2]
4	24	18	6	4	20	14.250[1]
5	26	23	3	9	17	13.731[1]
6	24	15	9	6	18	5.410[2]

Table 8

Relative Ability Of Botel Reading Inventory (Form B)
And Standardized Silent Reading Tests To Place
Pupils On Their Instructional Reading Levels

Grade	N	Botel A		Stand. Test		Chi Square
		on	off	on	off	
2	20	15	7	4	18	9.263[3]
3	35	20	5	11	14	5.432[2]
4	24	16	8	4	20	10.371[3]
5	26	23	3	9	17	13.731[1]
6	24	14	10	6	18	4.200[2]

[1]Significant at the .001 level.
[2]Significant at the .05 level.
[3]Significant at the .01 level.

CONCLUSIONS

All five hypotheses stated in null form were rejected, leading to the following conclusions for the population studied:

1. The Botel Reading Inventory (Forms A and B) are more closely related to the criterion than the standardized silent reading tests used in this study.

2. The Botel Reading Inventory (Forms A and B) are more highly related to the standardized silent reading tests used in this study at the lower grade levels.

3. The Botel Reading Inventory (Forms A and B) place pupils reading below grade level at their instructional level more effectively than do the standardized silent tests used in this study.

Analysis of the data suggests that for the purpose of placing a student at his instructional level, the Botel Reading Inventory might be superior to the standardized silent reading tests used in this study. Of course, the findings need to be verified with other populations before any more general statement of the relative efficiency of these tests for such purposes could be made.

The findings of this study shed some light on the extent to which standardized tests place pupils at their instructional levels. The generalization is made by some that standardized silent reading tests place pupils at their frustration level. In this study it was found that while the standardized tests overplaced more pupils than did the

Botel Reading Inventory, many pupils were either correctly placed or underplaced by the standardized silent reading tests rather than overplaced.

Each school or school system ought to determine for itself the validity or relative ability of informal reading inventories and standardized tests to place its pupils at their instructional levels. Tabular procedures like those used in this study for determining extent of match should prove useful for such a purpose.

A more ideal study of the extent of match between various informal and standardized tests with the pupils' instructional levels can be accomplished by further improving the criterion measures. It seems that this would best be done by determining for each student how well he reads silently and orally at various levels of one or more scaled or graded reading programs. It was noted in this study that some able pupils were limited to grade level on basal reading material or at most to one grade level beyond grade placement, despite the fact that their performance in oral reading and comprehension was almost perfect. In an ideal classroom each student must be allowed to read as far beyond grade level as his ability and maturity justifies.

REFERENCES

1. Botel, Morton. How To Teach Reading. Chicago: Follett Educational Corporation, 1968.

2. Botel, Morton. Botel Reading Inventory, Forms A and B. Chicago: Follett Educational Corporation, 1966.

3. California Reading Test, Subtest of the California Achievement Tests, 1957 Edition with 1963 Norms, Form W. Los Angeles: California Test Bureau, 1963.

4. Iowa Reading Tests - Vocabulary and Comprehension, Subtests of the Iowa Tests of Basic Skills. Boston: Houghton Mifflin Co., 1956.

5. Kerlinger, Fred N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1966.

6. School and College Ability Tests. Princeton: Educational Testing Service, 1957.