Summary of Information on Reliability and Validity

Pre-Reading Screening Procedures and Slingerland Screening Tests for Identifying Children with Specific Language Disability

by Susanne P. Fulmer

Educators Publishing Service
A division of Delta Education, LLC
31 Smith Place
Cambridge, MA 02138
1-800-225-5750

Summary of Information on Reliability and Validity

In the fall of 1978 a study was undertaken in order to gather information on the reliabilities and validities of the *Pre-Reading Screening Procedures* and *Slingerland Screening Tests for Identifying Children with Specific Language Disability.* Test scores on these tests were obtained for a total of 1021 subjects in six school districts in Alaska, California, Ohio, and Washington. A detailed psychometric analysis of the tests as well as a description of the 1978 study and its results are contained in the *Pre-Reading Screening Procedures and Slingerland Screening Tests for Identifying Children with Specific Language Disability: Technical Manual*.* The following is a brief summary of the information available in the *Technical Manual*.

Reliability

Table 1 gives three estimates of the reliability, or consistency, of test scores on the *Pre-Reading Screening Procedures* and the *Slingerland Screening Tests*. The alpha reliability coefficients are estimates of homogeneity, or consistency of performance over all the items on the test. The test-retest reliability coefficients are estimates of the stability of test scores over time. The inter-rater reliability coefficients are estimates of the consistency of test scores when a student's test and corresponding retest are scored by different raters. The standard errors of measurement give an indication of the amount of variability (or error) one could theoretically expect in scores if a student were retested many times (assuming sources of error, such as practice effect, were eliminated).

*A copy of the *Pre-Reading Screening Procedures and Slingerland Screening Tests for Identifying Children with Specific Language Disability: Technical Manual* can be obtained from Educators Publishing Service, 31 Smith Place, Cambridge MA 02138 -1089.

TABLE 1
RELIABILITIES AND STANDARD ERRORS OF MEASUREMENT (SE_M)

TEST	NUMBER of Items	SE* M	Reliability		
			alpha (N)	test-retest (N)	inter-rater (N)
Pre-Reading Screening Procedures	106	3.9	.94 (209)	.78 (183)	.78 (137)
Slingerland Screening Tests					
Form A	114	3.4	.94 (227)	.71 (180)	.69 (141)
Form B	148	3.4	.94 (239)	.78 (221)	.78 (183)
Form C	154	4.3	.96 (137)	.85 (185)	.91 (89)
Form D	186	3.9	.93 (126)	.80 (120)	.83 (64)

*These calculations were done using alpha reliability estimates. Standard errors of measurement using test-retest or inter-rater reliability estimates would appear larger.

For the *Pre-Reading Screening Procedures* and the *Slingerland Screening Tests* coefficient alpha ranged from .93 to .96. These figures indicate that for each test the items are strongly related to each other: they tend to measure the same trait or combination of traits. Test-retest reliability coefficients ranged from .71 to .85, indicating a good degree of stability in test scores over time. Inter-rater reliability coefficients ranged from .69 to .91, indicating a good degree of consistency even when tests were scored by different raters. Since the inter-rater reliabilities were comparable to test-retest reliabilities, it is evident that the scoring of test papers by different raters was not an important source of errors in scores. However, it should be noted that these reliability coefficients are most applicable when tests are scored and evaluated by trained raters, as done in this study. In general, these three types of estimates of reliability for each test indicate that students' scores are not greatly affected by possible sources of distortion or error, such as the circumstances of test administration, the characteristics of the students, the scoring of items, and item characteristics.

The standard errors of measurement provide easily interpretable measures of the amount of error one can expect in a given student's test score.

The actual score a student obtains on a test is called "the obtained score." The student's true score can be thought of as "the hypothetical score" that would be obtained if the test were perfect and involved no error in measuring the characteristics of interest. The standard error of measurement is the standard deviation of the errors of measurement for the test. It is assumed that these errors of measurement have a normal distribution; thus, characteristics of the Normal Curve are used in determining the amount of error one can expect in estimating true scores. The standard error of measurement can be interpreted as follows:

> Suppose a student obtains a score of 83 items correct on the *Pre-Reading Screening Procedures.* The standard error of measurement for the *Pre-Reading Screening Procedures* (as indicated in Table 1) is 3.9. Reference to a table of areas of the Normal Curve indicates that probability is .68 that this student's true score lies within one standard error of measurement of the obtained score of 83 (or between 83-3.9 and 83 + 3.9). That is, there is a 68% probability that the student's true score lies between 79.1 and 86.9. And there is a 95.4% probability that the true score lies within two standard errors of measurement of 83 (or between 75.2 and 90.8).

Thus, the standard error of measurement can be used to obtain a range within which one has a certain degree of confidence that the student's true score would lie. The smaller the standard error of measurement, the smaller the range becomes. When this range is small, the measurement of the characteristic or trait in question is more precise. The standard errors of measurement which appear in Table 1 are relatively small when compared to the total number of items for each test, indicating that the tests exhibit a good degree of precision in measurement.

Additional specific information on the reliabilities of the tests, as well as standard errors of measurement, can be found in the *Technical Manual.*

Validity

For the *Pre-Reading Screening Procedures* and the *Slingerland Screening Tests* we are concerned with providing evidence of criterion-related or predictive validity. As stated in the *Teacher's Manual to Accompany Slingerland Screening Tests*, the "purpose of the Screening Tests is to screen from among a group of children those with potential language difficulties and those with

already present specific language disabilities who are in need of special attention *at the moment*." (Slingerland, 1970, p.xx). A similar statement is made concerning the *Pre-Reading Screening Procedures* (see Slingerland, 1977, p.1). Criterion-related validity is important when measuring or screening for "readiness," or when selecting students for special programs of study (or "special attention").

According to Slingerland, Specific Language Disability "refers to children of average to high intelligence whose difficulties in reading, spelling, handwriting, written and sometimes oral expression interfere with academic achievement" (1970, p. xx). The correlation of scores on the PSP and the Slingerland Screening Tests with the CTBS scores and other measures suport the ability of these tests to detect difficulties in the above academic areas. These correlations are presented in Table 2.

TABLE 2

CORRELATIONS OF TOTAL SCORES ON THE *PRE-READING SCREENING PROCEDURES* (PSP) AND THE *SLINGERLAND SCREENING TEST S* WITH THE *COMPREHENSIVE TESTS OF BASIC SKILLS* (CTBS) AND RATINGS OF ABILITY IN PENMANSHIP †

		Slingerland Screening Tests Form			
	PSP	*A*	*B*	*C*	*D*
CTBS(raw scores)					
Alphabet Skills	.71(28)***				
Vis. & aud. discr.	.42(28)**				
Pre-reading	.69(28***				
Spelling		-.55(71)***	-.74(48)***	-.61(38)***	-.53(11)*
Total reading	.70(54)***	-.63(71)***	-.63(43)***	-.72(38)***	-.71(12)**
Total language	.28(54)*	-.58(46)***	-.75(15)***	-.68(9)*	-.86(11)***
Ability in penmanship					
Teacher ratings	-.52(57)***	.48(58)***	.41(95)***	.29 (57)**	.53(66)***
Independent ratings ††	-.44(214)***	.24(228)***	.29(238)***	.36(139)***	.31(74)**

*p<.05
**p<.01
***p<.001

†The numbers given in parentheses indicate the number of subjects for which the given correlation was calculated.

††These ratings of handwriting samples were done by trained raters, using Zaner-Bloser evaluation scales and other criteria. For more detailed information, see the *Technical Manual.*

Note that many correlations in Table 2 are negative. These negative correlations reflect differences in methods of scoring (i.e., *Comprehensive Tests of Basic Skills* scores are expressed as the number correct, while *Slingerland Screening Test* scores reflect the number incorrect). However, it is the magnitude of the correlation which is important in judging the strength of the relationship between the two variables (i.e., a correlation of -.71 between two variables would indicate a stronger relationship than a correlation of .56). Also, the probability that an obtained correlation is due to chance rather than to actual relationships between variables is indicated by the value of p. For example, a p value of .01 indicates that the odds are only one in one hundred that the indicated correlation coefficient would occur merely by chance. The smaller the p value, the more confidence we can place in our results.

Table 2 shows high positive correlations of the *Pre-Reading Screening Procedures* with *Comprehensive Tests of Basic Skills* (CTBS) scores, especially in alphabet skills, pre-reading, and total reading. Thus, children who obtain low CTBS scores in these areas tend to obtain low scores on the Pre-Reading Screening Procedures, and those who obtain high CTBS scores tend to obtain high *Pre-Reading Screening Procedures* scores. Therefore, it appears that the *Pre-Reading Screening Procedures* total scores would tend to indicate which children might have difficulty in reading, language, and readiness skills. The *Slingerland Screening Tests (Forms A, B, C,* and *D*) exhibit strong correlations with CTBS scores in spelling, total reading, and total language. Thus, children who obtain low scores on the CTBS in spelling, total reading, and total language tend to make a relatively large number of errors on the *Slingerland Screening Tests*. Also, for both the *Pre-Reading Screening Procedures* and the *Slingerland Screening Tests* there is a relatively strong relationship between scores on these tests and teacher ratings or independent ratings of ability in penmanship.

In general, the correlations presented in Table 2 support the ability of the *Pre-Reading Screening Procedures* and the *Slingerland Screening Tests* to screen out children who may have difficulties in reading, spelling, handwriting, language, and readiness skills. Please consult the *Technical Manual* for additional correlations of these tests with other variables, as well as for a discussion of homogeneity, correlations among subtests, and factorial composition as they relate to validity. The findings of this study are also related to those of previous studies as listed in the *Technical Manual.*

References

Burns, W.J. and Burns, K.A. "The Slingerland Screening Tests: Local Norms," *Journal of Learning Disabilities* 10 (1977): 7.

Comprehensive Tests of Basic Skills. Monterey, CA: CTB/ McGraw-Hill, 1973.

Evaluation Scales for Cursive and Manuscript Writing, Grades One through Six. Columbus, OH: Zaner-Bloser, Inc., 1974.

Oliphant, G. *A Study of Factors Involved in Early Identification of Specific Language Disability.* The Orton Society Reprint Series (Reprint No. 34). Pomfret, CT: The Orton Society, Inc., 1969.

Slingerland, B.H. *Teacher's Manual* to accompany *Slingerland Screening Tests for Identifying Children with Specific Language Disability, Forms A, B, and C.* Cambridge, Mass: Educators Publishing Service, 1970.

Slingerland, B.H. *Teacher's Manual* to accompany *Slingerland Screening Tests for Identifying Children with Specific Language Disability, Form D.* Cambridge, Mass.: Educators Publishing Service, 1974.

Slingerland, B.H. Revised *Pre-Reading Screening Procedures*: *Teacher's Manual*. Cambridge, Mass: Educators Publishing Service, 1977.

Snyder, L.; Brown, B.; and Gates, S. "Pilot Study to Determine the Utility of the Screening Tests for Identifying Children with Specific Language Disability in the Diagnosis of Specific Learning Disability." Unpublished manuscript. Board of Education of the Borough of North York, Ontario, Canada, 1972.

ISBN 0-8388-61686

Copies of this summary are available without
charge from Educators Publishing Service,
31 Smith Place, Cambridge, MA 02138-1089