

**Keys to Reading Success™ Technical Report**

**August 2002**

**Steven G. Viger and Edward W. Wolfe**

**Michigan State University**

**Table of Contents**

<b><u>Topic</u></b>	<b><u>Page</u></b>
<b>Executive Summary.....</b>	<b>3</b>
<b>Reliability.....</b>	<b>6</b>
<b>Table 1: Test-Retest Stability.....</b>	<b>6</b>
<b>Concurrent Validity.....</b>	<b>7</b>
<b>Table 2: Concurrent Validity Estimates for Word Reading.....</b>	<b>8</b>
<b>Table 3: Concurrent Validity Estimates for Comprehension.....</b>	<b>8</b>
<b>Predictive Validity.....</b>	<b>8</b>
<b>Table 6: Predictive Validity Estimates for Keys Assessments.....</b>	<b>9</b>

## Executive Summary

Keys to Reading Success™ is a reading program that provides for screening, diagnostic and formative assessment as well as instruction and reporting in Phonics and Comprehension skills. Assessment data from 1,658 students, grades K-12, were collected on several measures. The census code designates the subject district is on the urban fringe of a large mid-western city. The data of primary interest were collected using the Keys to Reading Success™ comprehension and word reading tests. Word reading accuracy is assessed in both the phonics and passage reading portions of the tests. This technical report uses the data collected to evaluate the reliability of the Keys to Reading Success scoring system. Additionally, validity evidence was evaluated for the Keys to Reading Success instruments by way of relating the Keys to Reading Success scores to student level data obtained using other measures beyond the Keys to Reading Success assessments. These measures include the California Test of Basic Skills (CTBS) comprehension and total scores and the ACT.

Test-retest reliability estimates were computed using Cronbach's alpha coefficient, a standard in the field of instrument evaluation. Cronbach's alpha essentially evaluates the degree to which participants are rank ordered by the instrument across differing data collection periods. In order to accomplish that task, data must be available for a given instrument for at least two data collection instances. In the evaluated data set, such information was available for both Keys to Reading Success assessment types as well as the CTBS. In practice, it is desirable to have coefficients of at least .80. The evaluated data demonstrated coefficients above .90 for both Keys to Reading Success assessments. With the exception of the CTBS comprehension scores, coefficients for the other CTBS measures exceeded .90 with the CTBS comprehension

evaluations demonstrating reliabilities of .87 and .88 for grade equivalent and normal curve equivalent respectively.

Evidence for concurrent validity, evidence that the Keys to Reading Success assessments are correlated with tests measuring the same or similar constructs collected within the same time frame, was gathered for both the word reading and comprehension assessment of the Keys to Reading Success evaluations. Due to the high reliability of the aforementioned instruments, averages per student were used to compute the correlation coefficients. Correlation coefficients range from -1.00 to +1.00 and in this case suggest the amount of linear relatedness between the correlated measures. A negative coefficient in general suggests that as scores on one measure increase, scores on the measure it is being correlated with tend to decrease. The same type of logic suggests that a positive correlation is present in relationships where both measures tend to increase together. Since the point of establishing concurrent validity is to demonstrate that each instrument is measuring approximately the same construct, positive correlations are desirable. Additionally, coefficients greater than .40 suggest moderate to strong linear relatedness.

For the Keys to Reading Success word reading assessment, concurrent validity evidence was found to exist. Average word reading scores for the sampled students were moderately related to ACT scores ( $r = .47$ ), CTBS comprehension scores ( $r = .55$ ), and CTBS total scores ( $r = .58$ ). These relationships suggest that the Keys to Reading Success word reading test has a positive linear relationship with the aforementioned measures. Specifically, as word reading scores increased, the other measures tended to increase as well.

A similar, yet slightly stronger and more convincing pattern was observed with respect to the Keys to Reading Success comprehension scores. The CTBS total score was most strongly related to the Keys to Reading Success comprehension scores ( $r = .60$ ), followed by the CTBS

comprehension scores ( $r = .53$ ). Again, these observed coefficients suggest moderate to strong positive linear relationships.

Finally, predictive validity evidence was found by correlating the Keys to Reading Success comprehension and recognition data collected in the fall of 2000 with the ACT score obtained by those individuals in the spring of 2001. Coefficients were .43 and .42 for the comprehension and recognition scores respectively. This evidence suggests that information on Keys to Reading Success evaluations is useful in predicting performance on future assessments such as the ACT.

Overall, the evidence suggests that scores generated by the Keys to Reading Success assessments are reliable as evidenced by test-retest estimates. The scores from the assessments are also valid. Validity evidence was noted by concurrent and predictive types of evidence.

**Reliability**

Test-retest stability measures were computed for the following instruments: Keys to Reading Success (both word reading and reading comprehension subtests) and California Test of Basic Skills (comprehension and total scores in both grade equivalent and normal curve equivalent form). Internal consistency measures (coefficient alphas) are presented below in tabular form.

**Table 1: Test-Retest Stability**

Instrument	Test-retest Stability Estimate (Sample Size)
Keys to Reading Success Word Reading	0.95 (1392)
Keys to Reading Success Comprehension	0.96 (1393)
CTBS Comprehension Grade Equivalent	0.87 (109)
CTBS Comprehension Normal Curve Equivalent	0.88 (109)
CTBS Total Grade Equivalent	0.92 (109)
CTBS Total Normal Curve Equivalent	0.91 (109)

By convention, it is desirable to have reliability indices in the range of .80 to 1.00. As can be seen from Table 1, both the Keys to Reading Success scores as well as the California Test of Basic Skills (CTBS) scores were well within this range. These indices suggest that the students can be rank ordered in a consistent manner across multiple testing times on these instruments. Sample sizes for the Keys to Reading Success measures were more than adequate to get an accurate estimation. Although the CTBS samples were considerably smaller, they still are likely to be fairly precise estimates as to the reliability of the tests and are in accord with the published coefficients that are above .80 as well.

### **Concurrent Validity Evidence**

Evidence of concurrent validity (evidence that the Keys to Reading Success is correlated with measures of similar constructs measured at approximately the same time) was obtained by computing correlations between the average Keys to Reading Success grade level assignments for both the recognition and comprehension elements and scores on other measures of the verbal/reading construct (specifically, the average CTBS comprehension and total scores in grade equivalent form, ISAT total scores, and the ACT scores). Average scores were used to represent the various items due to the high reliability exhibited by the items in the aforementioned section. Correlation coefficients range from + 1.00 to – 1.00 with stronger relationships noted as the coefficient approaches each of the limits. For concurrent validity a moderate to strong positive correlation is what is desired, as this would suggest a positive linear relationship between the compared instruments. Those correlations are presented below (see tables 2 and 3) along with correlations corrected for attenuation when available. The attenuation correction takes into account the fact that perfect reliability is rarely, if ever, obtained and unless the reliability is perfect between two measures, the Pearson product moment correlation coefficient will be an underestimate of the observed relationship. The correction has the following formula:

$$r_{disattenuated} = r_{xy} \sqrt{\text{reliability}_x \times \text{reliability}_y}$$

As can be seen from the equation, unless the reliability of each instrument is 1.00, the denominator will always be a number less than one; therefore the corrected coefficient will always be greater than the uncorrected coefficient.

**Table 2: Concurrent Validity Estimates for Word Reading Subtest**

Concurrent Validity Estimates for Keys to Reading Success Word Reading		
Instrument	uncorrected coefficient	corrected coefficient
CTBS comprehension	0.50	.55
CTBS total score	0.55	.58
ACT score	0.47	NA

Table 2 reveals the observed relationships between the Keys to Reading Success Word Reading scores and the other tabled measurements. The coefficients reported suggest the potential for strong positive and substantively important relationships between the Keys to Reading Success Word Reading subtest and other measures of similar constructs.

**Table 3: Concurrent Validity Estimates for Comprehension Subtest**

Concurrent Validity Estimates for Keys to Reading Success Comprehension		
Instrument	uncorrected coefficient	corrected coefficient
CTBS comprehension	0.52	.57
CTBS total score	0.57	.60
ACT score	0.53	NA

Table 3 suggests a similar pattern of relationships as was noted in table 2. However, the correlation was stronger than reported in table 2 for the recognition scores suggesting that perhaps the comprehension portion is slightly more related to the same construct being measured by the other instruments.

### **Predictive Validity**

Finally, analyses were performed to determine if the Keys to Reading Success assessments were useful at predicting future performance on an achievement evaluation. For this sample, suitable assessment tools for such an analysis are the Keys to Reading Success word reading and comprehension scores recorded in the Fall of the school year as predictors with ACT scores (taken in the spring) as the criterion. Table 4 presents the correlations resulting from the analysis as well as an effect size indicator ( $r^2$ ). The effect size indicator describes the amount of variance accounted for by the linear relationship of the correlated variables.



**Table 4: Predictive Validity Estimates for Keys to Reading Success Assessments**

Instrument	Correlation with ACT	Effect Size
Keys Comprehension	0.43	.19
Keys Recognition	0.42	.18

Note: 98 participants had both ACT and Keys to Reading Success scores available

Table 4 reveals that performance on both the recognition and comprehension portions of the Keys to Reading Success assessment was positively related with ACT scores. Additionally the effect sizes were moderate for each. These findings suggest that higher scores on the Keys to Reading Success assessments predict higher scores on the ACT test given later in the same school year. The Keys to Reading Success scores account for 18-19% of the variability in their linear relationship with subsequent ACT scores.

-----

*Edward W. Wolfe is Assistant Professor, Measurement & Quantitative Methods Program, Michigan State University. Steven G. Viger is Graduate Instructor of Measurement & Quantitative Methods at Michigan State University.*