# Development of a Vocabulary General Outcome Measure

Susan M. Rattan, Ph.D. | Ruth A. Kaminski, Ph.D. | Jacob D. Gray, Ph.D.

## Introduction

The simple view of reading (Gough & Tunmer, 1986, Hoover & Gough, 1990) provides a framework for reading instruction that includes skills for accurate word recognition (i.e., phonemic awareness, alphabet knowledge including letter names and letter sounds, basic and advanced phonics) as well as skills involved in the comprehension of language. While there currently exists an array of general outcome measures with technical adequacy to measure decoding skills and the precursors to them, there are no current measures that have technical adequacy for screening and progress monitoring of language comprehension, including vocabulary, within a general outcomes approach to assessment.

The assessment of vocabulary and oral language skills is critical to improving reading outcomes for students as there is strong evidence that children with a history of speech/language delay in preschool are at risk for future difficulties in reading, especially when those speech/language delays persist into kindergarten (Hulme et al., 2015; Pennington & Bishop, 2009; Snowling et al., 2016).

### Purpose and Research Questions

The purpose of this study was to examine initial evidence of reliability and validity of an experimental general outcome measure of vocabulary and oral language (VOL) skills. The research questions were:

1. What are the descriptive statistics and distributions of scores on the VOL measure administered at the middle and end of grades K–3?
2. What is the relationship between performance on the VOL measure and performance on Acadience Reading K–6 measures administered concurrently?

## Method

### Participants

Participants in this pilot study included students in kindergarten through third grade from one diverse elementary school in the upper Midwest (50% Hispanic, 46% White, 1% Black, 3% multiracial; 79% eligible for free or reduced lunch). The sample included 73 kindergarten students, 74 first-grade students, 91 second-grade students, and 75 third-grade students (52 students at the end of the year).

### Procedures

*Measure Development*

All words appearing on the VOL measure were selected from a carefully constructed word pool. This pool was developed with consideration of three characteristics of words: frequency of words in text, difficulty of words, and utility of words.

Frequency of words in text was determined by examining *Word Zones for 4,000 Simple Word Families* (Hiebert, 2012). This database classifies words into different zones based on frequency in text and ranges from Zone 1 (most frequent) to Zone 4 (least frequent).

We used three sources of information to determine word difficulty. First, we examined *Words Worth Teaching* (Biemiller, 2010). Words in this database are assigned to one of six levels based on the percentage of students who knew the meaning of each word at various grade levels. Words that we considered to be easy were known by a high percentage of students in the early grades, whereas words that we considered to be difficult were known by a low percentage of students in the upper grades. Second, we examined *Age of Acquisition Norms for 30,000 English words* (Kuperman et al., 2012). Age of acquisition refers to the age at which a word is typically learned. We included a wide range of words that are typically learned from the preschool years up through middle school. Third, we considered the concreteness of the words as another way to judge difficulty. We used concreteness ratings from a study conducted by Brysbaert and colleagues (2014) that resulted in ratings for 40,000 English words. Words were rated on a 5-point scale with ratings of 1 indicating that a word is very abstract to ratings of 5 indicating that a word is very concrete.

We were interested in selecting words that would be useful for students to know. We used a model developed by Beck and colleagues (2002) to judge the utility of the words in our word pool. This model includes 3 tiers of words. Tier 1 includes basic words that are known by most students when they enter kindergarten (e.g., walk, wet). Tier 2 includes words that are conceptually easy to understand but that use a label that is at a higher level than the basic labels found in Tier 1 (e.g., saunter, drenched). Finally, Tier 3 includes words with a low frequency of use that are usually specific to a particular domain (e.g., photosynthesis, hypotenuse).

For each time point at each grade level, we selected 15 words including five nouns, five verbs, and five adjectives. Words within each grade level were categorized as easy, medium, or difficult based on age-of-acquisition ratings. All grade level lists began with two easy nouns. After that, words were assigned to triads. Each triad included a noun, verb, and adjective and an easy, medium, and difficult word (e.g., a hard noun, easy verb, medium adjective).

*Assessor Training*

Assessors were selected by the reading specialist at our partnering school and consisted of paraprofessionals and reading specialists. Ten assessors completed an online, asynchronous training in January that included an hour-long video, three practice activities, and a reliability check to ensure accurate scoring. The same assessors participated in a refresher training in May that consisted of a review of administration and scoring rules, a practice activity, and another reliability check. All assessors demonstrated a level of reliability > .8 and were cleared to administer the measure at the middle and end of the school year.

*Data Collection*

Assessors administered the VOL measure individually to students during the third week of January (middle of year, or MOY) and the third week of May (end of year, or EOY). All score forms were collected and sent to the Acadience Learning offices for data entry.

### Measures

*Acadience Vocabulary Oral Language (VOL)*

Acadience VOL is designed based on the principles of General Outcome Measurement (Fuchs & Deno, 1991) and is administered three times per year in grades K–3. Assessors administer the measure one-on-one to students by asking students to provide a definition of a word that is read out loud, (e.g., "What does dull mean?"). Responses are scored as correct (2 points), partially correct (1 point), or incorrect (0 points). The assessor presents the 15 words on the list one at a time until they get to the end or the student receives five consecutive scores of zero.

### Acadience Reading K–6

The Acadience Reading K–6 measures used in this study included all of the measures that comprise the MOY and EOY benchmark assessments for each grade level. Benchmark assessments contained various combinations of the following measures: First Sound Fluency (FSF), Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Oral Reading Fluency (ORF), and Maze. These measures are combined to create the Reading Composite Score (RCS) that we used to examine the relationship between VOL and Acadience Reading K–6 measures. Additional information on the design specifications of Acadience Reading K–6 measures and the formulas for calculating the RCS are available in the *Acadience Reading K–6 Technical Manual* (Good et al., 2019), available at www.acadiencelearning.org.

### Analyses

Data analyses included descriptive statistics, including distributions of VOL scores, and concurrent correlations between VOL and Acadience Reading K–6 RCS scores.

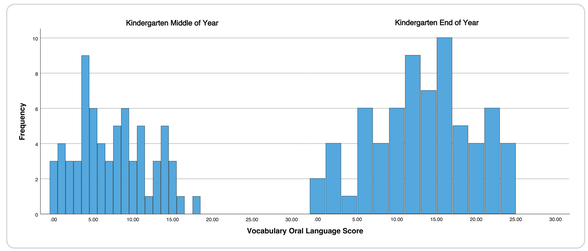
## Results

### Descriptive Statistics

Means, standard deviations, and distributions of scores on the VOL measure are presented in Figures 1–4.

**FIGURE 1**

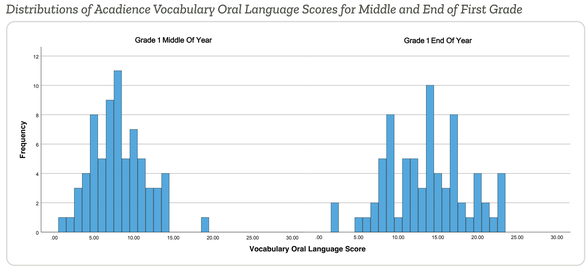
*Distributions of Acadience Vocabulary Oral Language Scores for Middle and End of Kindergarten*




*Note:* Summary statistics for the middle of kindergarten are as follows. Mean = 7.49. Standard deviation = 4.596. Sample size = 68. Summary statistics for the end of kindergarten are as follows. Mean = 12.93. Standard deviation = 6.313. Sample size = 68.

**FIGURE 2**

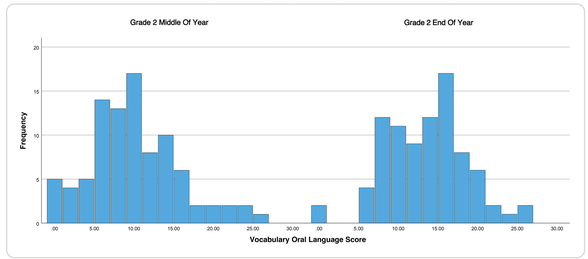
*Distributions of Acadience Vocabulary Oral Language Scores for Middle and End of First Grade*




*Note:* Summary statistics for the middle of first grade are as follows. Mean = 8.13. Standard deviation = 3.383. Sample size = 70. Summary statistics for the end of first grade are as follows. Mean = 13.61. Standard deviation = 4.978. Sample size = 72.

**FIGURE 3**

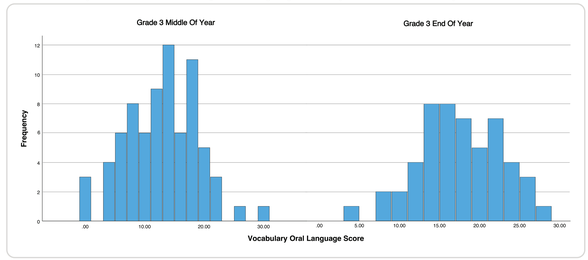
*Distributions of Acadience Vocabulary Oral Language Scores for Middle and End of Second Grade*




*Note:* Summary statistics for the middle of second grade are as follows. Mean = 9.48. Standard deviation = 5.654. Sample size = 91. Summary statistics for the end of second grade are as follows. Mean = 12.95. Standard deviation = 5.09. Sample size = 86.

**FIGURE 4**

*Distributions of Acadience Vocabulary Oral Language Scores for Middle and End of Third Grade*




*Note:* Summary statistics for the middle of third grade are as follows. Mean = 12.43. Standard deviation = 5.953. Sample size = 75. Summary statistics for the end of third grade are as follows. Mean = 16.85. Standard deviation = 5.127. Sample size = 52.

### Reliability

Reliability data are presented in Table 1. Cronbach's alphas range from .609–.858.

**TABLE 1**

*Reliability of Acadience Vocabulary Oral Language*

| Grade | Cronbach's α |
|---|---|
| Kindergarten | |
| Middle of Year | .803 |
| End of Year | .860 |
| Grade 1 | |
| Middle of Year | .609 |
| End of Year | .754 |
| Grade 2 | |
| Middle of Year | .823 |
| End of Year | .771 |
| Grade 3 | |
| Middle of Year | .858 |
| End of Year | .809 |

*Note:* $N = 15$.

### Concurrent Validity of Vocabulary Oral Language

Results of the correlational analysis are reported in Table 2. The concurrent correlations between VOL and the Acadience Reading K–6 RCS scores are in the moderate range (.30-.54).

**TABLE 2**

*Correlations between Vocabulary Oral Language Total Score and Acadience Reading Composite Score*

| Time of Year | Grade | | | |
|---|---|---|---|---|
| | K | 1 | 2 | 3 |
| Middle of Year | .47** | .34** | .49** | .47** |
| End of Year | .30* | .54** | .44** | .35* |

*Note:* *$p < .05$. **$p < .01$.

## Discussion

This study provides initial evidence to support the use of Acadience VOL.

- Distributions show a range of scores at each grade level.
- Reliability coefficients fall in an acceptable range for a screening measure (Salvia & Ysseldyke, 2017).
- VOL was moderately correlated with the RCS, suggesting a valid assessment of early literacy. Acadience Reading K–6 measures assess the skills that are highly related to and necessary for decoding. It is likely that these correlations would be higher with a measure of literacy that includes an assessment of language skills.

Future research should examine the following:

- The degree to which VOL can predict scores on the RCS and additional language measures;
- The relationship between VOL scores in the early grades and scores on measures of reading comprehension in later grades; and
- Item-level analyses including the contribution of each word on the word lists to determine if substitutions should be made.

References available by contacting the authors.