

Technical Adequacy of Acadience Reading Pre-K: PELI[®]

This document is reprinted from the
Acadience Reading Pre-K PELI[®]
Assessment Manual Chapter 8.

Ruth A. Kaminski

Mary Abbott

Katherine Bravo Aguayo

Roland H. Good

acadience[®] reading pre-k: peli

Suggestion Citation:

Kaminski, R. A., Abbott, M., Bravo Aguayo, K., and Good, R. H. (2023). *Technical Adequacy of Acadience Reading Pre-K: PELI*. Acadience Learning Inc. www.acadiencelearning.org. (Reprinted from *Acadience Reading Pre-K: PELI assessment manual*, pp. 75–91 by R. A. Kaminski, M. Abbott, K. Bravo Aguayo, & R. H. Good, 2021, Voyager Sopris Learning.)

Table of Contents

Reliability of the PELI	1
Alternate-Form Reliability	1
Table 8.1. <i>Alternate-Form Reliability Coefficients by Book and Age Group for the PELI Composite Score and Language Index</i>	2
Table 8.2. <i>Median and Range of Alternate-Form Reliability Coefficients by Age Group for PELI Subtests and Composite Scores</i>	3
Inter-Scorer Reliability	3
Table 8.3. <i>Inter-Scorer Reliability for PELI Subtests and Composite Score</i>	3
Standard Error of Measurement	4
Table 8.4. <i>Median Standard Errors of Measurement Based on Alternate-Form Reliability Coefficients for Subtests and Composite Scores by Age</i>	4
Validity of the PELI	4
Evidence of Content Validity	5
Evidence of Criterion-Related Validity	8
Table 8.5. <i>Concurrent and Predictive Criterion-Related Validity Coefficients for PELI Alphabet Knowledge and Phonological Awareness with Acadience Reading Criterion Measures</i>	9
Table 8.6. <i>Concurrent and Predictive Criterion-Related Validity Coefficients of PELI Vocabulary-Oral Language, Comprehension, and Language Index with Language Criterion Measures</i>	10
Table 8.7. <i>Concurrent and Predictive Criterion-Related Validity Correlations of PELI Composite Scores with Early Literacy Composite Outcome Measures</i>	10
Evidence of Construct Validity	10
Table 8.8. <i>Rotated Factor Loadings of PELI Components for Children in Their 3/4 Preschool Year and Children in Their 4/5 Preschool Year</i>	11
Table 8.9. <i>Contingency Table for the End-of-Year PELI Composite Score (PCS) Benchmark Goal and Cut Point for Risk with Reading Composite Score (RCS) Outcome Goal for 4/5-Year-Old Children</i>	13
Table 8.10. <i>Contingency Table of the End-of-Year PELI Composite Score (PCS) Benchmark and Cut Point for Risk with PPVT-4 Outcome Goal for 4/5-Year-Old Children</i>	14
Table 8.11. <i>Contingency Table for End-of-Year PELI Language Index (PLI) Benchmark Goal and Cut Point for Risk with PPVT-4 Outcome</i>	14

Table of Contents (continued)

Figure 8.1. <i>Probability of Meeting or Exceeding the PPVT-4 40th Percentile Standard Given the End-of-Year PELI Language Index Value for 4/5-Year-Old Children</i>	15
Table 8.12. <i>Conditional Probabilities of Attaining the Next Benchmark or Outcome Goal Given Scores At or Above the Benchmark on the PELI</i>	16
Table 8.13. <i>Decision-Utility Metrics for the PCS End-of-Year Benchmark Goal with Respect the Benchmark Goal and Cut Point for Risk Levels for 4/5-Year-Old Children</i>	17
Table 8.14. <i>Receiver Characteristic Curve–Area Under the Curve (AUC) for the Outcome Goal and for the Outcome Cut Point for Risk</i>	18
Summary	19
References.....	20

Technical Adequacy of the PELI

It is important that assessment used for educational decision making be reliable and valid and adhere to accepted professional standards of measurement (American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME), 2014). This chapter explains the evidence gathered that supports the reliability, validity, and decision utility of the PELI in assessing a young child's early literacy and language skills.

Reliability of the PELI

The reliability of a test denotes the degree to which a test produces stable and consistent results over time and across different forms of the test and different testers. It is generally recommended that reliability coefficients be at least .60 if the scores are used for administrative purposes, .70 for progress monitoring decisions, .80 for screening decisions, and .90 for important individual decisions (Salvia, Ysseldyke, & Witmer, 2017). Two types of reliability were evaluated for the PELI: alternate form, and inter-scorer.

Alternate-Form Reliability

Alternate-form reliability is the degree to which two or more versions of the same test correlate with one another. For progress monitoring assessments, alternate-form reliability can provide the most valuable information. It represents the extent to which scores generalize across different samples of behavior, at different times, and potentially by different testers. To estimate alternate-form reliability, a sample of children are assessed with two different forms and the scores of the two forms are correlated.

Alternate-form reliability of the PELI was examined in a study in which a different PELI book (a reference form) was administered to a group of preschool-age children approximately two weeks after a designated benchmark book was administered at each benchmark time point (beginning-, middle-, and end-of-year).

The sample consisted of 3,229 preschool-age children from 106 schools and/or preschool programs from 22 states representing all geographic regions of the United States. Of the total sample, 2,416 children were 4–5 years old (i.e., one year from kindergarten entry at the beginning of the year) and 813 were 3–4 years old (i.e., two years from kindergarten entry at the beginning of the year).

Sites were matched based on geographic region, number of children, and program type (e.g., public school preschool program, Head Start, private preschool program) and assigned to one of three groups. Each group had designated books assigned for beginning-of-year, middle-of-year, and end-of-year benchmark time points. Within each group, one-third of the children received the benchmark assessment book followed by the PELI book that served as the alternate reference form (*On the Farm*) at the beginning of the year. One third of the children received the alternate book along with the benchmark book at the middle of the year and the final third at the end of the year. Sample sizes for subgroups of 3/4-year-old children ranged from 46 to 793 and for 4/5-year old children from 154 to 2,416.

Alternate-form reliability for each PELI book with the reference book (*On the Farm*) was computed using Pearson’s product-moment correlation. Reliability coefficients for the PELI Composite Score (PCS) and PELI Language Index (PLI) for each book are provided in *Table 8.1*.

For all books and for both age groups, the reliability coefficients for the PCS are greater than .80 (median = .88). Reliability coefficients for the PLI for individual books range from .65 to .86 for 3/4-year-old children (median = .78) and from .72 to .84 for 4/5-year-old children (median = .79).

Table 8.1

Alternate-Form Reliability Coefficients by Book and Age Group for the PELI Composite Score and Language Index

PELI Book	Alternate-Form Reliability Coefficients					
	3/4-Year-Old Children			4/5-Year-Old Children		
	<i>n</i>	PCS	PLI	<i>n</i>	PCS	PLI
Time for Bed	105	.89	.85	273	.91	.84
At the Playground	46	.88	.76	154	.86	.74
On the Farm	793	.89	.78	2,416	.88	.79
Getting a New Puppy	72	.87	.77	277	.88	.80
Grandma’s Birthday	93	.89	.78	345	.88	.79
Cooking With Mom	162	.85	.65	232	.86	.72
Show and Tell at School	133	.85	.81	282	.88	.80
A Day at the Beach	82	.90	.76	366	.91	.81
Trip to Outer Space	100	.92	.86	242	.89	.79
Off to the Grocery Store				245	.86	.77

Note. Alternate-form reliability of each book computed with respect to *On the Farm*. Alternate-form reliability of *On the Farm* is the median alternate-form reliability with all other PELI books. *Off to the Grocery Store* had fewer than 30 children who were 3–4 years old and so is not reported.

A summary of alternate-form reliability coefficients for the PCS, the PLI, and individual subtest scores is provided in *Table 8.2*. The median alternate-form reliability coefficients for individual subtests range from .66 to .95. The highest alternate-form reliability coefficients are for Alphabet Knowledge and Phonological Awareness (range = .80 – .95). The alternate-form reliability coefficients for Vocabulary-Oral Language and Comprehension are lower (range = .66 – .75).

Table 8.2

Median and Range of Alternate-Form Reliability Coefficients by Age Group for PELI Subtests and Composite Scores

Subtest/Composite Score	Median (Range) of Alternate-Form Reliability Coefficients	
	3/4-Year-Old Children (<i>n</i> = 46 – 162)	4/5-Year-Old Children (<i>n</i> = 154 – 366)
Alphabet Knowledge	.95 (.91–.97)	.94 (.90–.98)
Phonological Awareness	.81 (.64–.93)	.80 (.71–.84)
Vocabulary-Oral Language	.70 (.50–.83)	.75 (.62–.79)
Comprehension	.73 (.65–.79)	.66 (.62–.72)
PELI Language Index	.78 (.65–.86)	.79 (.72–.84)
PELI Composite Score	.88 (.85–.92)	.88 (.86–.91)

Inter-Scorer Reliability

Inter-scorer reliability indicates the extent of agreement among assessors who administer and score the test. Inter-scorer reliability of the PELI was examined in a study in which two independent observers simultaneously scored a PELI assessment administered to one child. The two scores were then correlated using Pearson’s product-moment correlation coefficient. The sample consisted of 74 preschool-age children across two sites in two different states in the Pacific Northwest and the Midwest. Inter-scorer reliability coefficients are presented in *Table 8.3*. For all subtests, the inter-score reliability coefficients are greater than .90.

Table 8.3

Inter-Scorer Reliability for PELI Subtests and Composite Score

PELI Book	First Scorer			Second Scorer		Inter-Scorer Reliability
	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	
Alphabet Knowledge	74	13.12	9.94	13.46	9.87	.96
Phonological Awareness	74	12.04	5.73	12.57	5.70	.96
Vocabulary-Oral Language	74	5.05	5.53	5.05	5.52	.98
Comprehension	74	18.59	7.78	19.66	8.07	.90
PELI Language Index	74					
PELI Composite Score	74	150.41	72.27	156.39	73.50	.98

Note. All correlations significant, $p < .001$. Mean differences were small for all subtests and composites, with Cohen’s *d* ranging from 0.03 to 0.13.

Standard Error of Measurement

The Standard Error of Measurement (SEM) is a statistic that estimates the amount of error in a child’s score. It is related to the test’s reliability and the variability of the test scores and may be used to estimate a confidence interval—a range of scores within which the child’s true score is likely to be.

The SEM for the PELI was computed using the alternate-form reliability coefficients provided in *Tables 8.1* and *8.2* and is based on the formula $SEM = SD (\sqrt{1-r})$ where SD is the standard deviation and r is the reliability. The median SEMs for all books for each PELI subtest and composite scores by age group are available in *Table 8.4*.

Table 8.4

Median Standard Errors of Measurement Based on Alternate-Form Reliability Coefficients for Subtests and Composite Scores by Age

Subtest/Composite	3/4-Year-Old Children	4/5-Year-Old Children
Alphabet Knowledge	1.93	2.20
Phonological Awareness	2.21	2.47
Vocabulary-Oral Language	3.58	3.56
Comprehension	2.88	2.69
PELI Language Index	18.38	16.66
PELI Composite Score	20.29	20.33

The SEM can be used to compute a confidence interval for each score on the PELI.

Computing a confidence interval is done in the following manner: a multiplier of 1 is used for a 68% confidence interval, a multiplier of 1.96 is used for a 95% confidence interval, and a multiplier of 2.58 is applied for a 99% confidence interval. For example, the standard error of measurement for a PELI Composite Score for a 4- to 5-year-old child is 20.33. This means that if a child has a composite score of 200, there is 68% confidence that the child’s true score is between 180 and 220, 95% confidence that the child’s true score is between 160 and 240, and 99% confidence that the child’s true score is between 148 and 252

Validity of the PELI

Validity of a test refers to “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests.” (AERA, APA, & NCME, 2014, p. 11). Thus, validation of a test should provide evidence of support for the interpretations of test scores that are related to the proposed use of a test (Salvia, Ysseldyke & Witmer, 2017). Types of evidence that can be considered in validating a test include evidence related to test content, internal structure, relationships between the test and other performances, convergent and discriminative power, and consequences of testing (AERA et al., 2014). While different experts use different terminology to describe these concepts, we chose to use Salvia, Ysseldyke, and Witmer’s (2017) terms: (a) *content validity*, which includes evidence related to test content; (b) *criterion-related validity*, which includes evidence of the relationships between the test and other performances; and (c) *construct validity*, which includes evidence related to internal structure, evidence of convergent and discriminant power, and evidence of the consequences of testing.

It should be noted that the process of test validation of the PELI is ongoing and will continue to occur throughout the life of the assessment.

Evidence of Content Validity

Content validity is the extent to which a test's items actually represent the domain or area of skills that are to be measured. Evidence of content validity for the PELI is provided by a detailed description of the underlying rationale and the research base for the selection of items and the format for each subtest. That research rationale is based on the premises that an assessment of preschool early literacy and language skills should: (a) measure skills that are predictive of future reading skills, (b) be relatively brief and efficient, (c) be formatted within a process that mirrors the literacy experiences of preschool children, and (d) support teachers in their efforts to make instructional decisions based on the assessment. Content validity of the PELI is strengthened by choosing indicators of skill areas that are predictive of future reading success and are developmentally appropriate for preschool children. Across more than three decades of research in early literacy and language, researchers have identified foundational preschool skills that are highly predictive of acquisition of reading skills and later reading success. These skills include alphabet knowledge, phonological awareness, vocabulary and oral language, and listening comprehension (Lonigan, 2006; National Early Literacy Panel, 2008; Neuman & Dickinson, 2001). Each of these skills is assessed by one of the subtests of the PELI.

The PELI subtests, by design, are *indicators* of each of the skill areas. For example, the phonological awareness subtest does not measure all possible phonological awareness skills such as rhyming, alliteration, and blending. Instead, the PELI phonological awareness subtest is designed to be an *indicator* of a child's progress toward the phonemic awareness outcome at the end of preschool of being able to identify the first sound in words. Similarly, Alphabet Knowledge does not measure all aspects of alphabet knowledge such as recognizing and matching letter shapes, letter names, and sounds. Rather, the child's ability to name the letters of the alphabet is an *indicator* of the child's broader alphabet knowledge. This notion of the skills assessed by the PELI being *indicators* is a critical one. Focusing on measuring indicators allows for a relatively efficient assessment that is reliable and valid for the purposes of identifying children who need instructional support and monitoring progress of all children in the acquisition of early literacy and language skills.

In addition to assessing critical skills, it is important that the assessment be conducted in a format that is "gathered from realistic settings and situations that reflect children's actual performance" (National Association for the Education of Young Children (NAEYC) and the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE), 2003, p. 3). After a review of current assessments for early literacy in preschool as well as research-based practices for facilitating early literacy and language skills in preschool (e.g., Justice & Ezell, 2002; Lonigan, Burgess, & Anthony, 2000; National Early Literacy Panel, 2000), we concluded that a shared book-reading experience would help to facilitate children's responses in an assessment of early literacy and language skills. The content of each subtest is set within the context of the book's theme to improve the flow and connection between tasks as well as to make the connection to reading.

Subtest items and formats for each of the PELI subtests were developed through an iterative research process in a series of pilot studies over a five-year period. Different items and versions of the subtests were field-tested with samples of preschool age children to determine feasibility of the items and formats for assessing young children. Descriptive statistics as well as concurrent and predictive correlations with criterion measures were used to determine which items and formats worked the best to identify children with low early literacy and language skills as well as which items and formats resulted in measures that were sensitive to growth and development over time (Bravo Aguayo & Kaminski, 2012; Kaminski, Abbott, Bravo Aguayo, & Latimer, 2013). In addition, a wide range of PELI consumers provided comments and feedback about test methodology and usability. These included comments from teachers, administrators, and university-level research scientists. This feedback informed decisions about how well the chosen skills and test format worked and also about the usefulness of the information provided for making educational decisions for individual children (Bravo Aguayo, Abbott, & Kaminski, 2015).

Alphabet Knowledge. During the Alphabet Knowledge subtest, a child is shown a page of letters and asked to name as many letters as possible. The letters are arrayed in a picture related to the theme of the book. For example, in *Cooking With Mom*, the letters are in a bowl of soup. Alphabet Knowledge includes a mix of upper- and lowercase letters rather than all uppercase or all lowercase letters. Our research on the PELI indicates that many children who attend preschool lack knowledge of letter names at the beginning of their preschool years, but have mastered all of their uppercase letters by the time of kindergarten entry. To provide a sensitive indicator of letter recognition that is neither too difficult for 3-year-olds at the beginning of preschool nor too easy for 5-year-olds at the end, approximately 75% of the letters are uppercase letters with the remaining 25% being lowercase letters. Among the lowercase letters, approximately half are identical to their uppercase counterpart, while the other half are unique characters. The inclusion of a small number of unique lowercase letters has been found to be an effective way to distinguish the extent of letter knowledge in preschoolers and monitor growth on the skill (Turnball, Bowles, Skibbe, Justice, & Wiggins, 2010).

Phonological Awareness. In the Phonological Awareness subtest, the child is shown a picture of a scene related to the theme of the book. For example, in *Cooking With Mom*, the scene is of a cupboard in a kitchen. The child is shown a series of 10 pictures of objects that go in the cupboard. After a model and practice item, the assessor names the picture and asks the child to say the first part of the word or the first sound in the word. After the child responds, he or she gets to put the picture in the cupboard. The format of the subtest is similar to the First Sound Fluency subtest of Acadience® Reading K–6 (Good et al., 2011, revised 2018), in which a word is presented to a kindergarten child who says the first sound in the word. The addition of the pictures and the game of putting the picture into the book pocket helps to keep preschool children engaged with the task.

Words for the Phonological Awareness subtest of the PELI were chosen that were relevant to the theme of the story and could be pictured. From the pool of words that met the criteria for each book, five words were selected for Part I, Word Parts. Three words are compound words (e.g., teapot) and two are two-syllable words (e.g., melon). Five one-syllable words were selected for Part II, First Sound. Of the five single-syllable words, two begin with continuous sounds (i.e., sounds that can be held out like /m/ and /s/), two begin with stop sounds (e.g., /p/, /k/), and one begins with a blend (e.g., /tr/, /bl/).

Vocabulary and Oral Language. To assess vocabulary and oral language skills, the child is shown a picture of a scene related to the theme of the book. Within each scene are 10 objects or items common to the theme. The child is asked to name all 10 items, and then is asked to tell everything he or she can about 5 of the items. Picture identification is a common task in preschool language assessments (e.g., Expressive Vocabulary Test, Williams, 2007). The ability to describe or tell about an object is similar to oral definition tasks which are commonly used in language assessments (e.g., Test of Language Development, Newcomer & Hammill, 2008).

Words for the Vocabulary-Oral Language subtest were initially selected from a word pool constructed by cross-referencing the following sources: (a) the Educator’s Word Frequency Guide (Zeno et al., 1995); (b) The Living Word Vocabulary (Dale & O’Rourke, 1981); (c) Words Worth Teaching (Biemiller, 2010); (d) Graves’ First 4,000 Words (Graves, Sales, & Ruda, 2009); (e) Hiebert’s Word Zones (Hiebert, 2005b); (f) MacArthur-Bates Communicative Development Inventories (Fenson et al., 2007); and (g) EDL Core Vocabulary (Taylor et al., 1989). Words highly ranked at early grade levels by multiple sources were considered more important and/or appropriate for preschool-age children and are included in our preschool word pool.

From the preschool word pool, words for the Vocabulary-Oral Language subtest of the PELI were chosen that: (a) were relevant to the theme of the story, and (b) could be pictured. For each PELI book, 10 words were selected based on the results of a study that examined the percentage of preschool children who could correctly identify potential PELI vocabulary words. Words were categorized as easy (correctly identified by over 80% of study participants), medium (50%–79% correct identification), and difficult (less than 50% correct identification). Three easy, three medium, and four difficult words were chosen for the picture-naming task for each PELI book, with consideration given to select words that represented a range of correct identification percentages within each difficulty category. From the pool of 10 words used for the picture-naming task, five were selected for the Tell About section of the Vocabulary-Oral Language subtest, including one easy, two medium, and two difficult words.

Comprehension. There are two parts to the Comprehension subtest of the PELI. Part I, Comprehension Questions, assesses a child’s skill in making predictions and inferences and correctly answering questions related to a simple story. The use of questions to assess comprehension has a long history in education and provides a way of directly assessing a child’s understanding of story elements that are explicitly stated as well as those that are not explicitly stated and require a child to make inferences based on the information provided. Comprehension questions have been found to provide a reliable and valid measure of preschool children’s story comprehension (Dempsey & Skarakis-Doyle, 2001; Paris & Paris, 2003). In addition, children’s answering of comprehension questions has been found to be sensitive to the effects of narrative interventions with preschool children (Spencer, Kajian, Petersen, & Bilyk, 2013).

Part II, the Shared Retell, assesses a child’s skill in providing the correct missing words in the story in an oral cloze task. Written cloze tasks, frequently used to assess reading comprehension, consist of text from which words have been deleted in a systematic manner. The child is asked to “fill in the blank” and provide the word or words that have been deleted from the text. Shared Retell consists of the assessor retelling the PELI story orally with missing words. During the retell, the assessor shows a series of pictures from the book and retells the story, pausing for the child to fill in the word or words that have been deleted. Oral cloze tasks like Shared Retell reduce the memory and expressive language demands of traditional comprehension question and/or story retell tasks (Skarakis-Doyle & Dempsey, 2008) and have demonstrated reliability and validity in assessing story comprehension of preschool-aged children (e.g., Joint Story Retell, Dempsey & Skarakis-Doyle, 2001).

Evidence of Criterion-Related Validity

Criterion-related validity is the extent to which a test relates to other tests that measure the same or similar constructs. Two types of criterion-related validity are commonly described. *Concurrent criterion-related validity* refers to how a child’s performance on the test relates to a criterion measure of the same construct administered at approximately the same time. *Predictive criterion-related validity* refers to how a child’s performance at one point in time predicts that child’s performance on the criterion measure at a later point in time.

Concurrent and Predictive Criterion-Related Validity of the PELI

The concurrent and predictive validity of the PELI were examined in a set of studies in which the PELI was administered at three benchmark time points (beginning-of-year, middle-of-year, and end-of-year) and correlated with a criterion measure administered at the end of the year. Subjects included children who were 4–5 years old. Criterion measures included the Core Language Index (CLI) of the Clinical Evaluation of Language Fundamentals–Preschool, Second Edition (CELF Preschool–2, Wiig, Secord, & Semel, 2004), the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4, Dunn & Dunn, 2007), and the beginning-of-year kindergarten measures of Acadience Reading (Good et al., 2011, revised 2018).

The PPVT-4 and CELF CLI were administered to different samples of children. The PELI Composite Score (PCS) represents a composite of language, alphabet knowledge, and phonological awareness skills. As such, the most appropriate outcome measure for the PCS is an early literacy composite outcome score composed of a language measure and Acadience Reading. Two early literacy composites were created as the most appropriate criterion measures for the PCS. One early literacy composite outcome was formed from the Reading Composite Score (RCS) for Acadience Reading and the PPVT-4 and a second early literacy composite outcome was formed from the RCS and the CELF CLI. Composites were formed by converting the language measure (e.g., PPVT-4 or CELF CLI) and the RCS to z-scores with a mean of 0 and standard deviation of 1 and then averaging the two z-scores.

Concurrent and predictive validity coefficients for the PELI Alphabet Knowledge and Phonological Awareness subtests with Acadience Reading Letter Naming Fluency and First Sound Fluency are presented in *Table 8.5*. The concurrent and predictive validity coefficients for the PELI Vocabulary-Oral Language and Comprehension subtests and PELI Language Index with CELF CLI and PPVT-4 are presented in *Table 8.6*. Validity coefficients for the PELI Composite Score with the early literacy composite outcome measures are reported in *Table 8.7*. The concurrent validity coefficients are at the end of the year; predictive validity coefficients are the beginning and middle of the year.

Table 8.5

Concurrent and Predictive Criterion-Related Validity Coefficients for PELI Alphabet Knowledge and Phonological Awareness with Acadience Reading Criterion Measures

Time of Year	Validity of PELI Alphabet Knowledge with Acadience Reading Letter Naming Fluency (N = 2,228)	Validity of PELI Phonological Awareness with Acadience Reading First Sound Fluency (N = 2,233)
Beginning of Year	.68	.56
Middle of Year	.76	.65
End of Year	.74	.66

Note. All criterion measures were administered at the end of the year. All validity coefficients significant at $p < .001$.

Table 8.6

Concurrent and Predictive Criterion-Related Validity Coefficients of PELI Vocabulary-Oral Language, Comprehension, and Language Index with Language Criterion Measures

PELI Subtest/Composite Score at Time of Year	CELF CLI (N = 174)	PPVT (N = 136)
Beginning of Year		
Vocabulary-Oral Language	.52	.72
Comprehension	.58	.70
PELI Language Index	.60	.77
Middle of Year		
Vocabulary-Oral Language	.56	.78
Comprehension	.46	.64
PELI Language Index	.57	.78
End of Year		
Vocabulary-Oral Language	.48	.67
Comprehension	.40	.62
PELI Language Index	.51	.72

Note. All criterion measures were administered at the end of the year. All validity coefficients significant at $p < .001$.

Table 8.7

Concurrent and Predictive Criterion-Related Validity Correlations of PELI Composite Scores with Early Literacy Composite Outcome Measures

PELI Composite Score at Time of Year	CELF–Acadience Reading Composite Outcome (N = 168)	PPVT-4–Acadience Reading Composite Outcome (N = 85)
Beginning of Year	.74	.83
Middle of Year	.72	.85
End of Year	.65	.80

Note. All criterion measures were administered at the end of the year. All validity coefficients significant at $p < .001$.

Evidence of Construct Validity

Construct validity of a test is the extent to which a test measures a theoretical trait or characteristic and includes evidence of convergent and discriminant power and evidence of the consequences of testing. The PELI is designed to be used to make educational decisions about children’s need for support in acquiring critical early literacy and language skills, and the consequences of the assessment should result in accurate identification of children who need instructional support. Evidence of construct validity of the PELI includes the PELI factor structure and decision-making utility of the PELI benchmark goals.

Factor Structure of the PELI

The factor structure of the PELI was investigated with 2,503 children in their 4/5 preschool year, and 813 children in their 3/4 preschool year. A principal component factor analysis was conducted using covariances with squared multiple correlation coefficients as initial estimates of commonalities. Principal components factoring with a quartimin oblique rotation were used to obtain rotated factor loadings. Rotated factor loadings of PELI components are shown in *Table 8.8*. The Comprehension and Vocabulary-Oral Language components loaded strongly on Factor 1, which corresponds to and correlates highly with the PELI Language Index ($r = .92$ to $.95$). Factor 2 represents alphabet knowledge, and Factor 3 corresponds to phonological awareness. This means that the PELI is composed of the three distinct skill areas of: (a) language skills, (b) alphabet knowledge, and (c) phonological awareness. The PELI data are interpreted and used for educational decision making using these three factors (see Chapter 7 of the PELI Assessment Manual).

Table 8.8

Rotated Factor Loadings of PELI Components for Children in Their 3/4 Preschool Year and Children in Their 4/5 Preschool Year

Book and Subtest	3/4 Preschool Year			3/4 Preschool Year		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Book 1						
Alphabet Knowledge	.00	.95	.03	-.01	.95	.02
Phonological Awareness	.00	.01	.88	.02	.03	.85
Comprehension	.77	-.07	.08	.70	-.05	.13
Vocabulary-Oral Language	.77	.00	.04	.85	.03	-.04
Book 2						
Alphabet Knowledge	.01	.96	.00	.01	.95	.01
Phonological Awareness	.02	.02	.85	.01	.01	.86
Comprehension	.88	-.02	-.05	.77	-.02	.04
Vocabulary-Oral Language	.79	.13	-.04	.86	.04	-.07

Note. Based on $N = 813$ 3/4-year-old children two years before kindergarten and $N = 2,503$ 4/5-year-old children in their year before kindergarten. Quartimin oblique rotation reported. Factor 1 correlated $.54$ with Factor 2 and $.72$ to $.83$ with Factor 3. Factors 2 and 3 correlated $.68$ to $.69$.

Decision Utility

A critical feature of any tool that is used for screening and progress monitoring is the establishment of benchmark goals, which can provide guidance about expected levels of performance and progress. Benchmark goals can be used both for identifying children who need additional instructional support and as a frame of reference for evaluating the adequacy of a child's progress. The development of PELI benchmark goals and studies evaluating their decision utility occurred in a series of studies over a two-year period. The first study included 274 3- to 4-year-old children and 2,472 4- to 5-year-old children from 217 preschool classrooms in 37 early childhood programs in nine states representing all four census regions of the United States. In the second study, 3,233 children participated from 106 schools located in 15 states representing all census regions of the United States.

PELI benchmark goals and cut points for risk were derived by examining the predictive utility of a score on the PELI at a particular point in time, compared with later PELI measures and to criterion measures that served as outcomes. Our fundamental logic for developing the benchmark goals was to begin with an external outcome goal and work backward following systematic step-by-step procedures. We started by determining a level of performance representing adequate early literacy and language skills on each outcome measure at the end of the year. We used the benchmark goal for the beginning-of-kindergarten RCS as our external outcome goal for early literacy skills and the 40th percentile on the PPVT-4 as our external outcome goal for language skills. Next, we examined the predictive utility of the end-of-year PCS for 4- to 5-year-olds with respect to the end-of-year external outcome goals (i.e., At or Above Benchmark on the RCS, 40th percentile on the PPVT-4) and used this data to specify a benchmark goal for the PCS. The primary specification for the PELI benchmark goals was to establish a level of skill where children scoring at or above the benchmark had a favorable probability (above 80%) of achieving subsequent literacy outcomes. Then, using the PCS end-of-year benchmark goal as an internal goal, we established the benchmark goals for the middle-of-year PCS. Finally, we established the benchmark goals on the beginning-of-year PCS using the middle-of-year PCS as an internal goal. Once the benchmark goals were established for the PCS, they were used to establish the benchmark goals for each individual PELI subtest using the same step-by-step procedures. The same step-by-step procedure was used for determining PELI goals for 3- to 4-year-olds using the 4- to 5-year-old beginning-of-year PCS as the starting point. Cut points for risk were derived using a similar step-by-step procedure. The primary specification for the PELI cut points for risk was to establish a level of skill where children scoring well below the benchmark had a low probability (less than 20%) of achieving subsequent literacy outcomes.

In addition to the primary specifications for the benchmark goals and cut points for risk, an important secondary consideration was based on the logistic regression predicting the probability of scoring at or above the benchmark on the outcome measure based on a child's score on the PELI. For all children with scores in the At or Above Benchmark range on the PELI, the overall probability of achieving subsequent early literacy goals may be greater than 80%, but for children with scores at the high end of the range, the probability is higher and for children at the low end of the range, the probability is lower. The logistic regression analysis was used to estimate the probability of achieving subsequent early literacy and language goals for children who obtained the exact benchmark score. We attempted to keep the predicted probabilities for children who obtained the exact benchmark score at 60% or higher.

Diagnostic efficiency of the PELI benchmark goals and cut points for risk also was evaluated using receiver-operator characteristic (ROC) curves for the PCS, the PLI, and each subtest of the PELI with subsequent PELI assessments and with each of the outcome measures. Additional indices were calculated to provide multiple perspectives on the effectiveness of the PELI scores, including sensitivity, specificity, positive and negative predictive power, and percent accurate classification.

In early childhood, we are operating from a prevention perspective. As such, our primary consideration in developing PELI benchmark goals was to identify a level of skill that is predictive of success so that we can set our goals and monitor progress toward those outcomes. Since the first step in our benchmark goal analyses was to predict from the end-of-year PELI goals to the outcome measure goals, we provide the detail from these specific analyses here.

Conditional Probabilities.Contingency tables for the PCS end-of-year benchmark goal and cut point for risk for 4/5-year-old children with the outcome measures are presented in *Tables 8.9* and *8.10*. The contingency table for the PLI end-of-year benchmark goal and cut point for risk for 4/5-year-old children with the PPVT-4 is presented in *Table 8.11*. (Contingency tables for each subtest and each age level at each time point (beginning of year, middle of year, and end of year) are available upon request from Acadience Learning.)

Table 8.9

Contingency Table for the End-of-Year PELI Composite Score (PCS) Benchmark Goal and Cut Point for Risk with Reading Composite Score (RCS) Outcome Goal for 4/5-Year-Old Children

RCS Outcome Status	PCS End-of-Year Screening Decision			Row Total (%)
	Well Below Benchmark	Below Benchmark	At or Above Benchmark	
At or Above Benchmark	130	300	1,218	1,648 (74%)
Below Benchmark	99	116	64	279 (13%)
Well Below Benchmark	231	50	19	300 (13%)
Column Total	460 (21%)	466 (21%)	1,301 (58%)	2,227 (100%)
Conditional probability of reaching benchmark	28%	64%	94%	

Note. The Reading Composite Score (RCS) outcome measure was administered at the end of the child’s pre-school year, using the beginning-of-kindergarten benchmark goals and cut points for risk.

As depicted in *Table 8.8*, there were 1,301 children who achieved a score at or above the benchmark on the PCS (58% of the total). This is a smaller proportion of children than those who scored at or above the benchmark on the RCS (1,648 or 74%). Of the children who scored at or above the benchmark on the PCS, 1,218 (94%) also scored above the benchmark on the RCS. Thus, the overall probability of achieving the RCS goal, given that a child’s PELI end-of-year composite score is above the benchmark, is 94%. This overall probability of achieving the RCS goal when a child’s score on the PELI is above the benchmark is known as the negative predictive power (i.e., the percentage of children identified as not at risk on the predictor who achieve the outcome goal).

Table 8.10

Contingency Table of the End-of-Year PELI Composite Score (PCS) Benchmark and Cut Point for Risk with PPVT-4 Outcome Goal for 4/5-Year-Old Children

PPVT Outcome Status	PCS End-of-Year Screening Decision			Row Total (%)
	Well Below Benchmark	Below Benchmark	At or Above Benchmark	
At or above 40th percentile	4	13	73	90 (66%)
20th–39th percentile	3	6	14	23 (17%)
Below 20th percentile	16	3	4	23 (17%)
Column Total	23 (17%)	22 (16%)	91 (67%)	136 (100%)
Conditional probability of reaching benchmark	17%	59%	80%	

With respect to the PPVT-4 (*Table 8.10*), 91 of 136 children scored above the benchmark on the PCS (67%). Of the children who scored at or above the benchmark on the PCS, 73 (80%) also scored above the benchmark on the PPVT-4. Thus, the overall probability of achieving the PPVT-4 goal if a child’s PELI end-of-year composite score is at or above benchmark (negative predictive power) is 80%. In addition to the PCS, the decision utility of the PELI Language Index with respect to the PPVT-4 also was examined (*Table 8.11*) with a similar pattern of results. For the PLI, 89 of 136 children scored above the benchmark (65%). Of those who scored at or above the benchmark on the PLI, 72 (81%) also scored above the benchmark on the PPVT-4. Thus, the overall probability of achieving the PPVT-4 goal if a child’s end-of-year PELI Language Index is at or above benchmark (negative predictive power) is 81%.

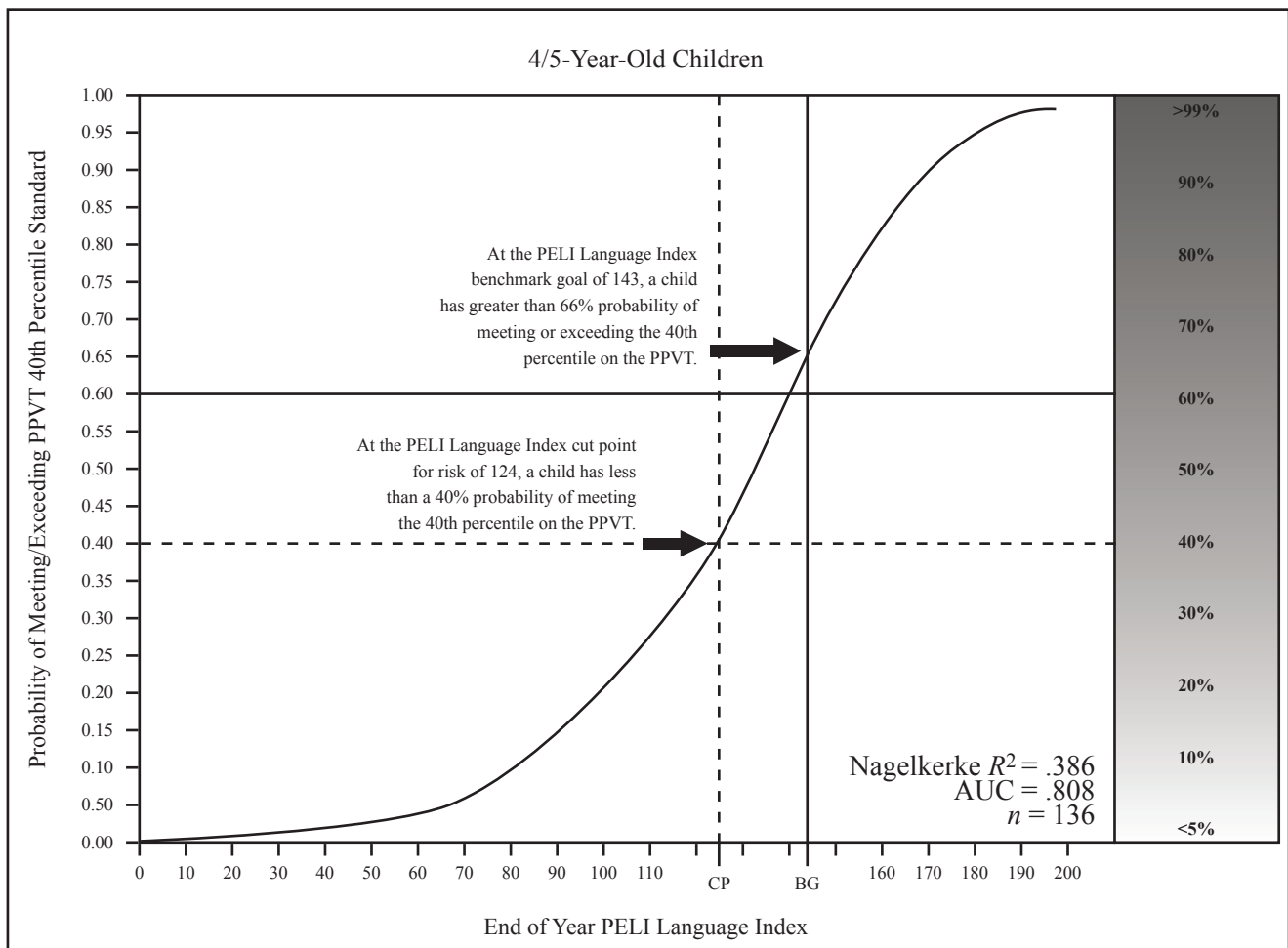
Table 8.11

Contingency Table for End-of-Year PELI Language Index (PLI) Benchmark Goal and Cut Point for Risk with PPVT-4 Outcome

PPVT-4 Outcome Status	PLI End-of-Year Screening Decision			Row Total (%)
	Well Below Benchmark	Below Benchmark	At or Above Benchmark	
At or above 40th percentile	3	15	72	90 (66%)
20th–39th percentile	2	8	13	23 (17%)
Below 20th percentile	17	2	4	23 (17%)
Column Total	22 (16%)	25 (18%)	89 (65%)	136 (100%)
Conditional probability of reaching benchmark	14%	60%	81%	

While the overall probability of achieving the PPVT-4 goal if a child’s end-of-year PELI Language Index is at or above benchmark is 81%, the probability is lower if a child is just barely at the benchmark goal (about 66%) and increases as the child achieves scores that are further above the benchmark goal (approaching 100%). The probability of meeting or exceeding the 40th percentile standard on the PPVT-4 for each PELI Language Index value is illustrated in *Figure 8.1*. Within each PELI benchmark status, higher PLI values yield higher probability of meeting or exceeding the 40th percentile standard on the PPVT-4. Children who score just above the benchmark goal, for example, do not have substantially different probability of meeting the PPVT-4 standard than children who score just below the benchmark goal.

Figure 8.1
Probability of Meeting or Exceeding the PPVT-4 40th Percentile Standard Given the End-of-Year PELI Language Index Value for 4/5-Year-Old Children



The PELI was developed to inform decisions within an Outcomes-Driven Model. For this purpose, the most important decision utility metric is the probability of attaining the next goal or benchmark as illustrated by the conditional probabilities in *Tables 8.9, 8.10, and 8.11* as well as in *Figure 8.1*. Those values for each age and time of year for the PELI Composite Score and PELI Language Index are summarized in *Table 8.12*. (The probabilities for each subtest, the Language Index, and the Composite Score are available upon request from Acadience Learning.)

Table 8.12

Conditional Probabilities of Attaining the Next Benchmark or Outcome Goal Given Scores At or Above the Benchmark on the PELI

Assessment and Time of Year	Outcome and Time of Year	Probability if Well Below Benchmark	Probability if Below Benchmark	Probability if At or Above Benchmark
PCS BOY 3/4-year-old children	PCS MOY (<i>n</i> = 1,439) 3/4-year-old children	13%	40%	89%
PCS MOY 3/4-year-old children	PCS EOY (<i>n</i> = 1,451) 3/4-year-old children	14%	34%	90%
PCS EOY 3/4-year-old children	PCS BOY (<i>n</i> = 315) 4/5-year-old children	15%	41%	84%
PCS BOY 4/5-year-old children	PCS MOY (<i>n</i> = 4,995) 4/5-year-old children	9%	36%	86%
PCS MOY 4/5-year-old children	PCS EOY (<i>n</i> = 4,995) 4/5-year-old children	7%	34%	87%
PCS EOY 4/5-year-old children	PPVT-4 EOY (<i>n</i> = 136) 4/5-year-old children	17%	59%	80%
PLI EOY 4/5-year-old children	PPVT-4 EOY (<i>n</i> = 136) 4/5-year-old children	14%	60%	81%
PCS EOY 4/5-year-old children	Acadience Reading Benchmark Goal for BOY Kindergarten (<i>n</i> = 2,227)	28%	64%	94%

Note. BOY = Beginning of year. MOY = Middle of year. EOY = End of year. Sample sizes are based on subsamples of children who had PELI scores at both time points.

Sensitivity and Specificity. Sensitivity and specificity metrics, long used as screening and diagnostic criteria within the field of medicine, have recently been applied to educational decisions. When applied to educational assessments, sensitivity and specificity are used as indices of the accuracy of a screening measure in identifying children who are at risk for learning difficulties from children who are not. The sensitivity and specificity of a measure are typically evaluated by using children's performance on one measure to predict performance on an outcome measure. A level of performance on the outcome measure is selected as an outcomes goal. Children who perform below the outcome goal are considered to be at risk for learning difficulties in a particular area, for example, reading. Children who meet the outcome goal are considered to be not at risk.

The sensitivity of a measure is the proportion of children who do not meet the outcomes goal who were identified as at risk on the predictor measure. Sensitivity is generally considered to be important because it is the proportion of children who were correctly identified by the screener as at risk. These are the children who are likely to need additional instructional support in order to achieve subsequent early literacy and language goals. Specificity is the proportion of children who achieve the outcomes goal who were not at risk on the predictor measure. Classification accuracy is a summative metric that describes the overall proportion of children who are correctly identified by the predictor measure as at risk or not at risk.

Applying sensitivity and specificity decision metrics to PELI is complicated because rather than a single decision level, there are two decision levels: (a) the benchmark goal level where children are considered to have adequate early literacy skills and are likely to continue to attain future early literacy goals (80% or better probability), and (b) the cut point for risk where the level of children’s early literacy skills place them at risk and where they are unlikely to attain future early literacy goals (20% or less probability) without additional instructional support. In between the benchmark goal and cut point for risk is a range of scores where future performance is harder to predict (about 50% +/- 10% probability of attaining future early literacy goals).

Because risk status can be changed for preschool children by providing effective support and learning opportunities, we believe that sensitivity and specificity are interpretable only when the measures are administered at the same time. Between the administration of a screening measure at the beginning of the year and an outcome measure at the end of the year, there is a gap in time in which intervention may occur and the prediction may be ruined. In fact, it is our role and function as preschool educators to use our screening data to reduce the risk status of children through learning opportunities and support. That is not a failure of the screening, but rather an intervention success.

In *Table 8.13* we provide the sensitivity, specificity, and classification accuracy for decisions using the benchmark goal as the predictor and the outcome and for decisions using the cut point for risk as the predictor and the outcome.

Table 8.13

Decision-Utility Metrics for the PCS End-of-Year Benchmark Goal with Respect to the Benchmark Goal and Cut Point for Risk Levels for 4/5-Year-Old Children

Criterion Measure	At or Above Benchmark on PELI Composite Score and on the Criterion			Below Cut Point for Risk on PELI Composite Score and on the Criterion		
	Sensitivity	Specificity	Classification Accuracy	Sensitivity	Specificity	Classification Accuracy
RCS	.86	.74	.77	.77	.88	.87
PPVT-4	.61	.81	.74	.70	.94	.90

Note. $N = 2,227$ for Reading Composite Score (RCS) for Acadience Reading. $n = 136$ for the PPVT-4.

Receiver Operator Characteristic (ROC) Curve–Area Under the Curve (AUC). Sensitivity and specificity generally involve a tradeoff such that choosing a higher decision level results in greater sensitivity but lower specificity. Receiver operator characteristic (ROC) curves provide a way to evaluate the decision utility of a measure across all possible decision levels. The area under the curve (AUC) ranges from .50 (decisions no better than chance) to 1.00 (very high sensitivity and specificity regardless of decision level). In general, an AUC greater than .90 is considered excellent; .80–.90, good; .70–.80 fair; and below .70 poor (Compton, Fuchs, Fuchs & Bryant, 2006). ROC curves for the outcome goal and outcome cut point for risk are provided in *Table 8.14*. The AUC for the PELI are all in the good to excellent range.

Table 8.14

Receiver Characteristic Curve–Area Under the Curve (AUC) for the Outcome Goal and for the Outcome Cut Point for Risk

Predictor, Time of Year, and Age Group	Outcome, Time of Year and Age Group	AUC for Outcome Benchmark Goal	AUC for Outcome Cut Point for Risk
PCS BOY 3/4-year-old children	PCS MOY (<i>n</i> = 1,439) 3/4-year-old children	.90	.90
PCS MOY 3/4-year-old children	PCS EOY (<i>n</i> = 1,451) 3/4-year-old children	.92	.93
PCS EOY 3/4-year-old children	PCS BOY (<i>n</i> = 315) 4/5-year-old children	.88	.86
PCS BOY 4/5-year-old children	PCS MOY (<i>n</i> = 4,995) 4/5-year-old children	.91	.93
PCS MOY 4/5-year-old children	PCS EOY (<i>n</i> = 4,995) 4/5-year-old children	.92	.94
PCS EOY 4/5-year-old children	PPVT-4 EOY (<i>n</i> = 136) 4/5-year-old children	.81	.94
PLI EOY 4/5-year-old children	PPVT-4 EOY (<i>n</i> = 136) 4/5-year-old children	.81	.93
PCS EOY 4/5-year-old children	Acadience Reading Benchmark Goal for BOY Kindergarten (<i>n</i> = 2,227)	.87	.90

Note. BOY = Beginning of year. MOY = Middle of year. EOY = End of year. Sample sizes are based on subsamples of children who had PELI scores at both time points.

Summary

This chapter presented evidence of the reliability, validity, and decision-making utility of the PELI. The information in this chapter supports the conclusion that the PELI is a reliable and valid measure of early literacy skills for preschool-aged children. Research on the technical adequacy and decision-making utility of the PELI is ongoing with different samples of children and statistical procedures, and using different external outcome measures. Ongoing research studies will provide additional evidence regarding the technical adequacy of the PELI as well as provide guidance for future revisions of the PELI. Current technical reports on the PELI are available from Acadience Learning.

References

- Bravo Aguayo, K., Abbott, M., & Kaminski, R. A. (2015). *Analysis of PELI user satisfaction questionnaire* (Technical Report No. 19). Dynamic Measurement Group. https://acadiencelearning.org/wp-content/uploads/2020/07/Analysis_of_PELI_User_Satisfaction_QNR_2015.pdf
- Bravo Aguayo, K., & Kaminski, R. A. (2012, April 11–14). *Reliability and validity of the preschool early literacy indicators (PELI)* [Conference session]. Council for Exceptional Children Convention, Denver, CO, United States.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. <https://www.apa.org/science/programs/testing/standards>
- Biemiller, A. (2010). *Words worth teaching: Closing the vocabulary gap*. SRA/McGraw Hill.
- Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98(2), 394–409.
- Dale, E., & O'Rourke, J. (1981). *The living word vocabulary*. World Book/Childcraft International.
- Dempsey, L. & Skarakis-Doyle, E. (2001). The validity of the Joint Story Retell as a measure of young children's comprehension of familiar stories. *Journal of Speech-Language Pathology and Audiology*, 25(4), 201–211.
- Dunn, L.M., & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test*, (4th ed.). Pearson.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates CDO words and sentences*. Brookes.
- Good, R. H., III, Kaminski, R. A., Cummings, K., Dufour-Martel, C., Petersen, K., Powell-Smith, K., Stollar, S., & Wallin, J. (2011, revised 2018). *Acadience Reading K–6*. Dynamic Measurement Group. <https://acadiencelearning.org/>
- Graves, M., Sales, G., & Ruda, M. (2009). *Source list for the first 4,000 words project*. Seward, Inc.
- Hiebert, E. H. (2005b). *Word Zones for 4,000 simple word families*. TextProject. https://textproject.org/wp-content/uploads/resources/WordZones_4000-simple-word-families.pdf
- Justice, L. M., & Ezell, H. K. (2002). Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology*, 11, 17–29.
- Kaminski, R., Abbott, M., Bravo Aguayo, K., & Latimer, R. (2013, February 12–15). *Examining the validity and utility of an assessment of preschool early literacy skills* [Paper presentation]. National Association of School Psychologists Conference, Seattle, WA, United States.
- Lonigan, C. J. (2006). Development, assessment, and promotion of preliteracy skills. *Early Education and Development*, 17(1), 91–114.
- Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36(5), 596–613.

- National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003, November). *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8*. <https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/pscape.pdf>
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. <https://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Neuman, S. B., & Dickinson, D. K. (2001). *Handbook of early literacy research*. Guilford Press.
- Newcomer, P. L., & Hammill, D. D. (2008). *Test of language development-primary* (4th ed.). Pro-Ed.
- Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, 38(1), 36–76.
- Salvia, J., Ysseldyke, J. E., & Witmer, S. (2017). *Assessment in special and inclusive education* (13th ed.). Cengage.
- Skarakis-Doyle, E., & Dempsey, L. (2008). Assessing story comprehension in preschool children. *Topics in Language Disorders*, 28, 127–144.
- Spencer, T. D., Kajian, M., Petersen, D. B., & Bilyk, N. (2013). Effects of an individualized narrative intervention on children’s storytelling and comprehension skills. *Journal of Early Intervention*, 35(3), 243–269.
- Taylor, S. E., Frackenpohl, H., White, C. E., Nieroroda, B. W., Browning, C. L., & Birsner, E. P. (1989). *EDL core vocabularies in reading, mathematics, science, and social studies*. Steck-Vaughn Company.
- Turnball, K. L., Bowles, R. P., Skibbe, L. E., Justice, L. M., & Wiggins, A. K. (2010). Theoretical explanations for preschoolers’ lowercase alphabet knowledge. *Journal of Speech, Language, and Hearing Research*, 53, 1757–1768.
- Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical evaluation of language fundamentals preschool* (2nd ed.). Harcourt Assessment.
- Williams, K. T. (2007). *Expressive vocabulary test (EVT-2)* (2nd ed.). Pearson.
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, T. (1995). *The educator’s word frequency guide*. Touchstone Applied Science Associates, Inc.