

TACL-4/TEXL Comprehensive Scoring Supplement

Elizabeth Carrow-Woolfolk and Elizabeth A. Allen



pro·ed
An International Publisher

8700 Shoal Creek Boulevard Austin, Texas 78757-6897
800/897-3202 Fax 800/397-7633 www.proedinc.com

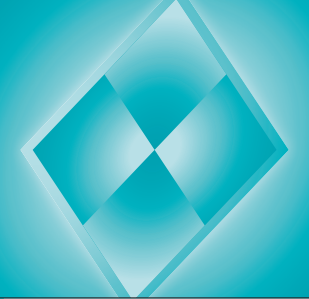


© 2014 by PRO-ED, Inc.
8700 Shoal Creek Boulevard
Austin, Texas 78757-6897
800/897-3202 Fax 800/397-7633
www.proedinc.com

All rights reserved. No part of the material protected by this copyright notice may be reproduced or used in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without prior written permission of the copyright owner.

Printed in the United States of America

1 2 3 4 5 6 7 8 9 10 23 22 21 20 19 18 17 16 15 14



Combining the TACL-4 and TEXTL Results to Produce a Comprehensive Measure of Oral Language

Clinicians who have given both the TACL-4 and TEXTL to the same child may want to combine these scores to obtain a comprehensive measure of oral language abilities. The procedures described in this document will produce measures of language ability across linguistic features (semantics and grammar) and linguistic systems (receptive and expressive) as well as overall oral language ability. The reproducible TACL-4/TEXTL Summary Form (found in Appendix C of this document and provided as a free download on the tests' product pages at www.proedinc.com) provides space to record the TACL-4 and TEXTL scores, calculate additional composite scores, conduct receptive–expressive discrepancy analyses, and profile these scores. These scores provide comprehensive information on oral language skills comparable to scores derived from the *Test of Language Development–Primary, Fourth Edition* (TOLD-P:4; Newcomer & Hammill, 2008), the *Test of Language Development–Intermediate, Fourth Edition* (TOLD-I:4; Hammill & Newcomer, 2008), the *Oral and Written Language Scales, Second Edition* (OWLS-II; Elizabeth Carrow-Woolfolk, 2011), the *Clinical Evaluation of Language Fundamentals, Fifth Edition* (CELF-5; Semel, Wiig, & Secord, 2013), and the *Comprehensive Assessment of Spoken Language* (CASL; Carrow-Woolfolk, 2008).

The following sections will discuss (a) how to complete the TACL-4/TEXTL Summary Form; (b) how to interpret the results; and (c) the psychometric properties of the features, systems, and overall composite scores.

Completing the TACL-4/TEXTL Summary Form

This form has six sections: Identifying Information, Summary of Subtest Scores, Summary of Composite Scores, Language Index Comparisons, Profile of Standard Scores, and Descriptive Terms. A completed example is provided in Figure 1. Each of these sections are described next.

Section 1. Identifying Information

In Section 1, record the examinee's name, age at the time of the TACL-4 and TEXTL testing, gender, grade, and school (if appropriate). Record the examiner's name and title, and reason for testing.

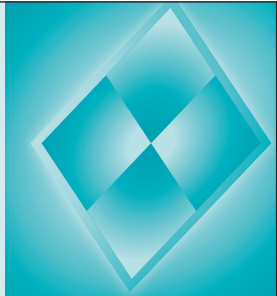
Section 2. Summary of Subtest Scores

In Section 2, record each of the TACL-4 and TEXTL subtest scaled scores in all the blank spaces provided for the subtests, then sum these scaled scores across each row to arrive at the sums of scaled scores for the composite scores.

TACL-4/TEXL

Summary Form

Elizabeth Carrow-Woolfolk and Elizabeth A. Allen



Section 1. Identifying Information

Name Alex Female Male Grade 4

Year 10 Month 5 School Pillow School

Age at TACL-4 Testing 10 5 Examiner's Name T. Cox

Age at TEXL Testing 10 5 Examiner's Title CCC-SLP

Reason for Testing Alex has been identified as dyslexic; his parents are concerned about language skills, as well.

Section 2. Summary of Subtest Scores

TACL-4/TEXL Composites	Subtest Scaled Scores						Sum of Scaled Scores
	TACL-4			TEXL			
	V	GM	EPS	V	GM	EPS	
Vocabulary	11			5			16
Morphology		10			6		16
Syntax			9			6	15
Receptive Language	11	10	9				30
Expressive Language				5	6	6	17
Oral Language	11	10	9	5	6	6	47

Section 3. Summary of Composite Scores

	Composite Scores	Sum of Scaled Scores	%ile Rank	Index Score	SEM	Confidence Interval	Descriptive Term
Linguistic Features	Vocabulary Index (VI)	<u>16</u>	<u>27</u>	<u>91</u>	3	<u>88 - 94</u>	<u>Average</u>
	Morphology Index (MI)	<u>16</u>	<u>30</u>	<u>92</u>	3	<u>89 - 95</u>	<u>Average</u>
	Syntax Index (SI)	<u>15</u>	<u>23</u>	<u>89</u>	3	<u>86 - 92</u>	<u>Below Average</u>
Linguistic Systems	Receptive Language Index (RLI)	<u>30</u>	<u>50</u>	<u>100</u>	3	<u>97 - 103</u>	<u>Average</u>
	Expressive Language Index (ELI)	<u>17</u>	<u>5</u>	<u>75</u>	3	<u>72 - 78</u>	<u>Borderline Impaired</u>
	Oral Language Index (OLI)	<u>47</u>	<u>25</u>	<u>90</u>	3	<u>87 - 93</u>	<u>Average</u>

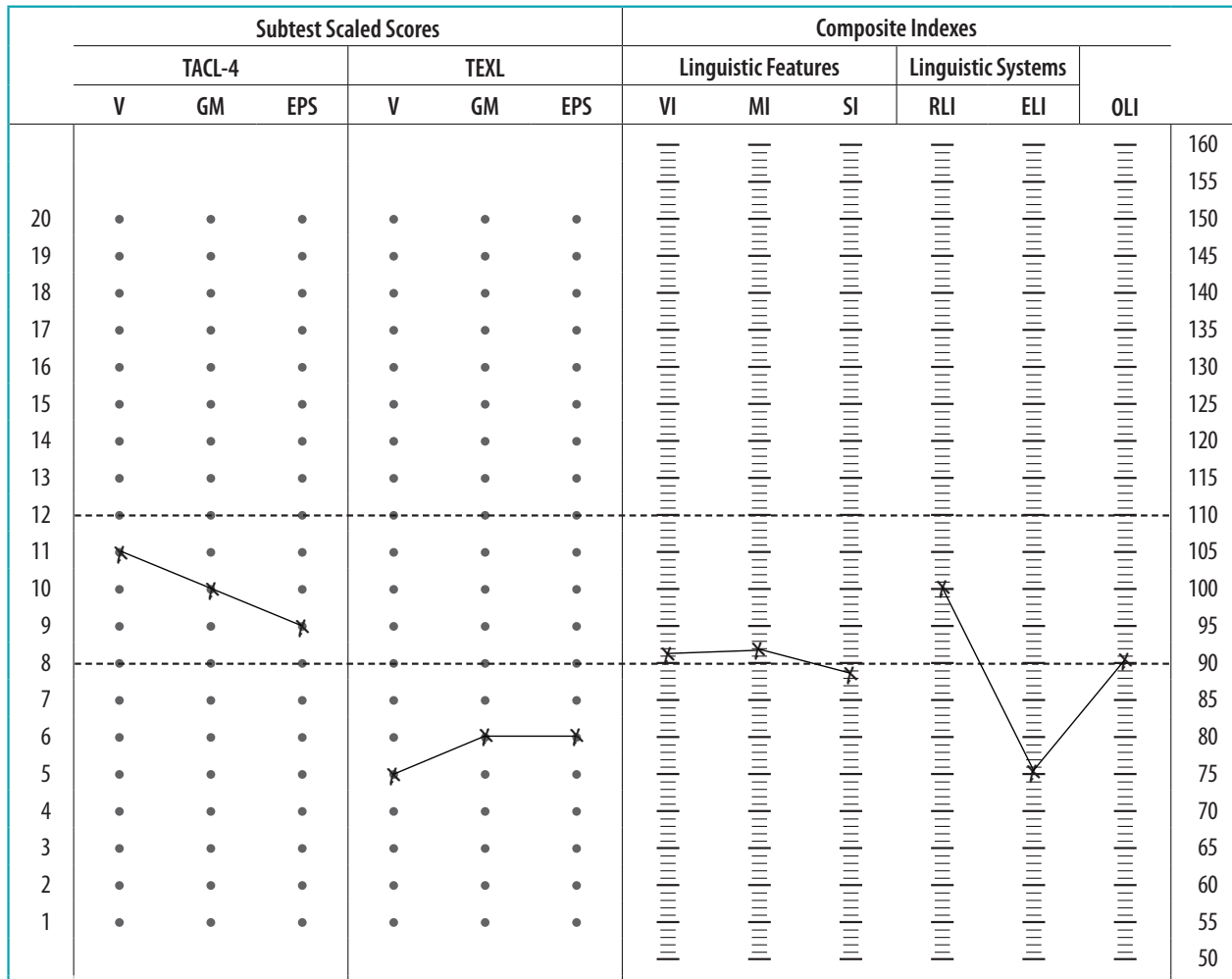
© 2014 by PRO-ED, Inc. PRO-ED, Inc. grants permission to reproduce this form for use with your clients. This license does not grant the right to reproduce these materials for resale, redistribution, or any other purposes. This Summary Form is available as a free PDF download on the TACL-4 and TEXL product pages at www.proedinc.com.

Figure 1. Example Summary Form completed for Alex.

Section 4. Language Index Comparisons

RLI vs. ELI Comparison	Index Score		Difference	Statistically Significant	Clinically Meaningful	% of Sample With This Difference
	RLI	ELI				
	100	75	= 25	>6	>20	3.6

Section 5. Profile of Standard Scores



Section 6. Descriptive Terms

Scaled Score	1–3	4–5	6–7	8–12	13–14	15–16	17–20
Descriptive Term	Impaired or Delayed	Borderline Impaired or Delayed	Below Average	Average	Above Average	Superior	Gifted or Very Advanced
Index Score	<70	70–79	80–89	90–109	110–119	120–129	>129

© 2014 by PRO-ED, Inc. PRO-ED, Inc. grants permission to reproduce this form for use with your clients. This license does not grant the right to reproduce these materials for resale, redistribution, or any other purposes. This Summary Form is available as a free PDF download on the TACL-4 and TEXL product pages at www.proedinc.com.

Figure 1. (continued)

Section 3. Summary of Composite Scores

Record the appropriate sum of scaled scores from Section 2 in the corresponding space provided in Section 3. These summed values are converted into percentile ranks and index scores using Table C.1 in Appendix C of the Examiner’s Manual (of either test) and Tables A.1 and A.2 in this document. For example, the sum of Alex’s Vocabulary subtest scaled scores was 16. By consulting Table A.1 in Appendix A, the examiner locates the “Sum of 2 Vocabulary subtests” column and reads down until 16 is found. The corresponding percentile rank is 27, and the corresponding Vocabulary Index is 91. Note: The Receptive Language and Expressive Language Indexes are converted using the table in Appendix C of each respective Examiner’s Manual.

Next, calculate the confidence interval by subtracting and adding the *SEM* (i.e., 3) to each index score and record these numbers in the space provided for confidence intervals. Finally, record the descriptive term corresponding to each index score. Section 6 provides the descriptive terms that correspond to each scaled and index score.

Section 4. Language Index Comparisons

The Receptive Language Index (RLI) and the Expressive Language Index (ELI) are recorded in Section 4 for comparison purposes. Subtract the ELI from the RLI and record the difference in the appropriate space. If the difference is greater than 6, the difference is statistically significant and the examiner should circle the “>6” under the “Statistically Significant” column. If the difference is greater than 20, the difference is clinically meaningful and the examiner should circle the “>20” under the “Clinically Meaningful” column. By consulting Table B.1, the examiner locates the difference score (i.e., 25) in the “Amount of difference” column and determines that the corresponding cumulative percent of the sample with a difference this large is 3.6 percent. Record this number in the appropriate space in this section.

Section 5. Profile of Standard Scores

The child’s performance on the TACL-4 and TEXTL and related supplemental composites is profiled in this section.

Section 6. Descriptive Terms

Descriptive terms that correspond to the scaled and index scores are provided in Section 6. These terms range from *impaired or delayed* to *gifted or very advanced*.

Interpreting the Results of the TACL-4/TEXTL Combined Testing

The following steps should be used to interpret the results of the TACL-4/TEXTL combined testing. In this section, we review three basic steps, shown below.

- Step 1—Describe the Oral Language composite score.
- Step 2—Conduct discrepancy analysis on the Receptive Language Index (RLI) and Expressive Language Index (ELI) composite scores.
- Step 3—Profile the TACL-4/TEXTL results.

The procedures to complete each of these steps are described in the following sections.

Step 1—Describe the Oral Language Composite

The first step to interpreting the combined TACL-4/TEXL scores requires describing the examinee's quantitative performance on the Oral Language composite (i.e., index score, confidence interval, percentile rank) and describing the performance qualitatively (i.e., descriptive term). This composite score is the most reliable data source in the TACL-4/TEXL combination because it includes sources of variation from all the scores that contribute to the RLI and the ELI. As such, this molar data source is more reliable than the more molecular scores that compose it. It may be used to refer to a child's general language ability and is the best estimate of an individual's current oral language ability. Evaluation of receptive and expressive discrepancies and supplemental composite performance, however, should be used to generate hypotheses or speculations about a child's receptive and expressive abilities. This information should be a basis for further in-depth study of the area that appears impaired.

Step 2—Conduct Discrepancy Analysis on the Receptive and Expressive Composites

The second step of the combined TACL-4/TEXL interpretation focuses on the variability between the RLI and the ELI to determine the representativeness of the Oral Language Index (OLI) as an estimate of overall oral language ability. When an examinee's scores on these two tests exhibit significant variability, the oral language composite may not serve as a good estimate of general oral language ability, and performance on the RLI, ELI, and supplemental composites should be interpreted.

Once the examinee's performance on the receptive and expressive composites has been described both quantitatively (i.e., index score, confidence interval, percentile rank) and qualitatively (i.e., descriptive term), these scores should then be examined for statistically significant and meaningful differences. Anastasi and Urbina (1997) provided a formula (p. 111) for determining how large a difference score must be to be statistically significant. *Statistical significance* is the probability that an observed score difference will occur in the population by chance alone. We used Anastasi and Urbina's formula to compute the minimal difference score (i.e., 7) required for a significant (at the $p < .05$ level) difference between the RLI and the ELI. However, a statistically significant difference score does not necessarily mean that the difference is large enough to be clinically useful. Reynolds (2003) provided a formula (#7, p. 483) to determine how large the difference score must be to be considered clinically useful. This formula was used to compute the difference required for clinical usefulness (i.e., 21). Minimal difference scores between the RLI and ELI are reported in Section 4.

Because the use of statistical significance alone identifies too many false-positive cases, it is important to note the relative frequency—the commonness or rarity—of a statistically significant score difference within a given population to interpret that discrepancy's meaningfulness. The relative frequency of differences between scores addresses an issue separate from statistical significance. As mentioned previously, statistical significance is the probability that an observed score difference will occur in the population by chance alone. However, differences of that magnitude may occur in the normative population with considerable frequency because of real differences that exist in the abilities of individuals within the population rather than because of chance factors. A statistically significant discrepancy that occurs very frequently is generally of less concern than a statistically significant discrepancy that is also rare.

Alex's test results (see Figure 1) provide us with an example of how to compare the difference scores. Alex's performance on the RLI was 100 and on the ELI was 75. His

difference score is 25 points. This 25-point difference is statistically significant at the $p < .05$ level and is clinically meaningful. One would then consult Table B.1 in Appendix B. The contents of Table B.1 also indicate that this large of a difference occurred in 3.6 percent of the TACL-4/TEXL normative sample.

Step 3—Profile the TACL-4 and TEXL Results

We recommend clinicians interpret combined TACL-4 and TEXL results at the most global level possible. However, if statistically significant or clinically meaningful variability exists between the TACL-4 and TEXL, the examinee's performance across the linguistic features assessed by these tests (i.e., vocabulary, morphology, syntax) should be examined for areas of strength and weakness. Section 5 provides space to profile subtest and composite performance on these tests. Scores falling below the shaded area that represents average performance (8–12 for subtests and 90–109 for composites) of the normative sample should be considered as areas of concern.

Psychometric Properties of the TACL-4/TEXL Composite Scores

Because the psychometric properties of the TACL-4 and TEXL were reviewed extensively in each respective manual, this section briefly reviews the psychometric properties of the linguistic features, and overall oral language composite scores that can be created using the TACL-4/TEXL Summary Form.

Reliability

Content sampling error (i.e., internal consistency reliability) for the TACL-4/TEXL composite scores was investigated by applying Cronbach's (1951) coefficient alpha method. Coefficient alphas for the composite were derived using Guilford's (1954, p. 393) formula. These coefficients are reported in Table 1. The coefficients were averaged using the Fisher z -transformation technique. The averaged coefficients are listed at the bottom of the table. As can be seen, these averaged coefficients exceed .90, a most desirable level of reliability.

The standard error of measurement (SEM) is used to estimate the confidence interval that surrounds a particular test score. The SEM estimates the amount of error in an individual's test score due to less-than-perfect reliability of a test. The SEM is based on the formula $SEM = SD \sqrt{1 - r}$ (SD = standard deviation; r = reliability), and establishes a zone within which an individual's true score probably lies. Due to the exceedingly high reliability coefficients, the $SEMs$ associated with these composite scores were all 3.

One cannot always assume that because a test is reliable for a general population it will be equally reliable for every subgroup within that population. Therefore, the alphas for selected subgroups within the normative sample were calculated and are reported in Table 2. The subgroups represent a broad spectrum of populations, embracing gender, ethnic, and exceptionality categories. These consistently large alphas demonstrate that the TACL-4/TEXL composites are equally reliable for all the subgroups investigated and support the idea that the test contains little or no bias relative to these groups.

Table 1
Coefficient Alphas at 10 Age Intervals (Decimals Omitted)

Age (in years)	TACL-4/TEXL composite values			
	Vocabulary	Morphology	Syntax	Oral Language
3	97	98	95	99
4	96	97	96	99
5	96	97	96	99
6	96	97	97	99
7	96	97	97	99
8	96	97	97	99
9	95	97	97	99
10	95	96	97	98
11	95	96	97	98
12	95	95	96	98
Average ^a	96	97	97	99

^aFisher's average of alpha coefficient across all ages.

Criterion-Prediction Validity

Anastasi and Urbina (1997) described criterion-prediction validity as “the effectiveness of a test in predicting an individual’s performance in specific activities” (p. 118). They stated that performance on a test is checked against a criterion that can be either a direct or an indirect measure of what the test is designed to predict. Thus, to be valid, a test like the TACL-4, for example, which is presumed to measure receptive language, and oral language in general, should correlate highly with other tests that measure the same ability.

To establish the criterion-prediction validity of the TACL-4/TEXL composite scores, we conducted two studies of its combined relationship to criterion measures of oral language as follows:

- *Clinical Evaluation of Language Fundamentals, Fourth Edition* (CELF-4; Semel, Wiig, & Secord, 2003)/*Clinical Evaluation of Language Fundamentals Preschool, Second Edition* (CELF Preschool-2; Wiig, Secord, & Semel, 2004) Receptive Language, Expressive Language, Language Content, Language Structure, and Core Language indexes
- *Oral and Written Language Scales, Second Edition* (OWLS-II; Carrow-Woolfolk, 2011) Listening Comprehension scale, Oral Expression scale, and Oral Language composite
- *Diagnostic Achievement Battery, Fourth Edition* (DAB-4; Newcomer, 2014) Spoken Language composite

Table 2
Coefficient Alphas for the TACL-4/TEXL Composite Scores for Selected Subgroups
(Decimals Omitted)

Subgroup	Vocabulary	Morphology	Syntax	Oral Language
Gender				
Male	98	99	99	99
Female	98	99	99	99
Race/ethnicity				
White	98	99	99	99
Black/African American	98	99	98	99
Hispanic	98	99	99	99
Asian/Pacific Islander	98	98	99	99
Exceptionality				
Gifted and talented	95	96	97	98
Learning disability	97	98	98	99
Attention-deficit/hyperactivity disorder	96	98	98	99
Deaf or hard of hearing	98	99	98	99
Articulation disorder	98	99	98	99
Language impairment	98	99	98	99
Autism spectrum disorder	98	99	98	99
Intellectual disability	97	98	97	99

Children in Study 1 were administered the TACL-4, the TEXL, the OWLS-II and either the CELF-4 or the CELF Preschool-2, depending on the age of the examinee. Note: For the purposes of this study, the CELF-4 and the CELF Preschool-2 have been treated as one measure and will be referred to as the CELF. Children in Study 2 were administered the TACL-4, the TEXL, and the DAB-4 Spoken Language Index, a measure of expressive language. The children in Study 1, the OWLS II and CELF study, were 59% male, 41% female, 72% White, 18% Black/African American, and 10% two or more races. Fifteen percent of the sample was of Hispanic origin and 71% of the sample had some exceptionality, primarily articulation disorder (40%) and/or language impairment (57%). The children in Study 2 were 53% male, 47% female, 69% White, 28% Black/African American, and <1% Asian/Pacific Islander. Twenty-eight percent of the sample was of Hispanic origin and 53% of the sample had some exceptionality.

The evidence relating to criterion-prediction validity has been organized into three sections: (a) a review of the correlations between the TACL-4/TEXL composites and the criterion language measures, (b) a comparison of the mean scores and standard deviations of the Oral Language Index with the scores of the criterion measures, and (c) a series of diagnostic accuracy analyses (including sensitivity, specificity, and receiver operating characteristic/area under curve [ROC/AUC] analyses).

Correlations With Criterion Measures

In this investigation of criterion-prediction validity, we report correlation coefficients showing the relationship of the TACL-4/TEXL composites to nine measures of oral language. The correlation coefficients between the TACL-4/TEXL composites and the criterion measures are reported in Table 3 and are organized into three categories of oral language (receptive, expressive, and general). In this analysis, we are asking a theoretical question: Do the TACL-4/TEXL composite scores measure receptive, expressive, and oral language, in general? Because the question is theoretical, one should correct coefficients to account for any range effects that might artificially repress or inflate the size of the coefficients. Both corrected and uncorrected coefficients are reported in the table.

Table 3
Corrected (and Uncorrected) Correlation Coefficients Between the TACL-4/TEXL Composites and Criterion Test Scores (Decimals Omitted)

Content/ criterion test	Score	Study	N	Vocabulary	Morphology	Syntax	Oral Language	Magnitude ^a
Receptive language								
OWLS-II	Listening Comprehension	1	68	79 (64)	74 (65)	72 (63)	80 (71)	Very large
CELF	Receptive Language	1	68	71 (63)	71 (69)	76 (75)	79 (77)	Very large
Average				75 (63)	73 (67)	74 (69)	79 (74)	Very large
Magnitude ^b				Very large	Very large	Very large	Very large	
Expressive language								
OWLS-II	Oral Expression	1	68	79 (68)	79 (75)	82 (79)	87 (83)	Very large
CELF	Expressive Language	1	68	61 (59)	68 (72)	72 (76)	75 (77)	Very large
Average				71 (64)	74 (74)	77 (78)	82 (80)	Very large
Magnitude ^b				Very large	Very large	Very large	Very large	
General language								
OWLS-II	Oral Language	1	68	84 (71)	83 (76)	84 (77)	90 (84)	Nearly perfect
CELF	Core Language	1	68	65 (64)	68 (72)	75 (79)	79 (81)	Very large
CELF	Language Content	1	68	67 (59)	75 (74)	73 (72)	79 (77)	Very large
CELF	Language Structure	1	68	65 (66)	66 (71)	75 (80)	77 (80)	Very large
DAB-4	Spoken Language Index	2	32	85 (85)	81 (78)	74 (69)	87 (85)	Very large
Average				75 (70)	75 (74)	77 (76)	83 (82)	Very large
Magnitude ^b				Very large	Very large	Very large	Very large	

Note. Children in Study 1 were administered either the CELF-4 or the CELF Preschool-2, depending on the age of the examinee. For the purposes of this study, the CELF-4 and the CELF Preschool-2 have been treated as one measure and will be referred to as the CELF. Coefficients inside the parentheses are uncorrected; coefficients outside the parentheses are corrected for range effects. OWLS-II = *Oral and Written Language Scales, Second Edition* (Carrow-Woolfolk, 2011); CELF = *Clinical Evaluation of Language Fundamentals, Fourth Edition* (Semel, Wiig, & Secord, 2003) and *Clinical Evaluation of Language Fundamentals Preschool, Second Edition* (Wiig, Secord, & Semel 2006); DAB-4 = *Diagnostic Achievement Battery, Fourth Edition* (Newcomer, 2014).

^aMagnitude of the corrected coefficient for the Oral Language Index based on Hopkins's (2002) criteria. ^bMagnitude of the average corrected coefficients based on Hopkins's (2002) criteria.

In interpreting the magnitude of these coefficients, we are guided by Hopkins (2002). He suggested that coefficients between .00 and .09 are very small or trivial, coefficients between .10 and .29 are small, coefficients between .30 and .49 are moderate, coefficients between .50 and .69 are large, coefficients between .70 and .89 are very large, and coefficients between .90 and 1.00 are nearly perfect. Because all these criterion tests measure important aspects of oral language, one would expect that their relationship to the TACL-4/TEXL composites would be large or very large, which proved to be the case.

As one might expect, the relationship between the TACL-4/TEXL composites and the criterion measures range in magnitude from very large to nearly perfect, providing ample evidence for the criterion-prediction validity of these composite scores. These analyses involved different criterion tests and a diverse sample of participants. Regardless of the criterion test or category of spoken language, the correlations with the TACL-4/TEXL composites were uniformly large.

Comparison of the Oral Language Index and Criterion Test Means and Standard Deviations

When two tests are highly correlated, it usually means that they are likely to be measuring the same or a similar ability. It does not necessarily mean, however, that the tests yield the same results. For example, one test may score consistently higher than another test even though they correlate highly with each other. The validity of both tests is supported when the two tests produce similar means as well as correlate highly with each other.

The standard score means, standard deviations, and comparative information for the Oral Language Index and the criterion language tests are presented in Table 4. The descriptive terms used to describe the means are listed in Section 6 of the TACL-4/TEXL Summary Form. The differences between the Oral Language Index means and the corresponding criterion test score means were analyzed using the dependent samples *t* test (Guilford & Fruchter, 1978) and effect size *r* from *d* (Borenstein, 2009) estimates and effect size *d* for correlated designs (Formula #3; Dunlap, Cortina, Vaslow, & Burke, 1996). The average values depict the weighted means and weighted standard deviations.

The findings reported in this table support the idea that, for all practical purposes, regardless of the sample's characteristics or criterion test administered, the Oral Language Index score will likely be similar to those obtained from giving other tests of oral language.

Diagnostic Accuracy Analyses

Diagnostic accuracy refers to the precision with which a test differentiates individuals with a disorder from those without a disorder. Researchers such as Betz, Eickhoff, and Sullivan (2013) and Dollaghan (2004) have suggested that this is “the most important criterion for evaluating a diagnostic measure” (Dollaghan, 2004, p. 395). Methods for establishing diagnostic accuracy involve the computation of a test's sensitivity and specificity indexes. In the current context, the *sensitivity index* reflects the ability of a test to correctly identify students who have a language impairment. The *specificity index* refers to the ability of a test to correctly identify examinees who do not have a language impairment.

The results for sensitivity and specificity are reported as proportions (i.e., percentages). The size of the proportions necessary to be considered acceptable varies depending on the purpose of the analysis (e.g., when screening for cancer, a relatively high number of false positives is tolerable in order to ensure that the number of true positives identified is high). Educational researchers vary in their opinions about how large a test's sensitivity and specificity indexes should be. Wood, Flowers, Meyer, and

Table 4
Standard Score Means (and Standard Deviations), Related Statistics, and *t* Values
for the TACL-4/TEXL Oral Language Index and Criterion Tests

Oral Language Index/ Criterion test	<i>N</i>	Study	<i>M</i> (<i>SD</i>)	Descriptive term	<i>t</i> ^a	Effect size <i>r</i> ^b	Effect size <i>d</i> ^c	Magnitude ^d
TACL-4/TEXL Oral Language	68	1	89 (12)	Below average	6.72***	.19	.38	Small
OWLS-II Oral Language			84 (14)	Below average				
TACL-4/TEXL Oral Language	68	1	89 (12)	Below average	7.07***	.32	.67	Moderate
CELF-4 Core Language			78 (20)	Below average				
TACL-4/TEXL Oral Language	68	1	89 (12)	Below average	4.70***	.20	.41	Small
CELF-4 Language Content			83 (17)	Below average				
TACL-4/TEXL Oral Language	68	1	89 (12)	Below average	6.46***	.31	.64	Moderate
CELF-4 Language Structure			78 (21)	Below average				
TACL-4/TEXL Oral Language	34	2	103 (14)	Average	-2.35*	-.10	-.21	Trivial
DAB-4 Spoken Language			106 (15)	Average				
TACL-4/TEXL Oral Language Average	306	1–2	91 (12)	Average	12.35***	.23	.46	Small
General Language Criterion Average			84 (17)	Below average				

Note. Children in Study 1 were administered either the CELF-4 or the CELF Preschool-2, depending on the age of the examinee. For the purposes of this study, the CELF-4 and the CELF Preschool-2 have been treated as one measure and will be referred to as the CELF. OWLS-II = *Oral and Written Language Scales, Second Edition* (Carrow-Woolfolk, 2011); CELF = *Clinical Evaluation of Language Fundamentals, Fourth Edition* (Semel, Wiig, & Secord, 2003) and *Clinical Evaluation of Language Fundamentals Preschool, Second Edition* (Wiig, Secord, & Semel 2006); DAB-4 = *Diagnostic Achievement Battery, Fourth Edition* (Newcomer, 2014).

^aValues of *t* were computed by the dependent samples method (Guilford & Fruchter, 1978). ^bEffect size *r* was calculated using Borenstein's (2009) formula for *r* from *d*. ^cEffect size was calculated using Dunlap et al.'s (1996) formula #3 that corrects for inflated effect size due to correlated design *t* tests. ^dValues of magnitude of the effect size correlation between the Oral Language Index and the criterion tests according to Hopkins's (2002) criteria.

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

Hill (2002) recommended that the sensitivity and specificity indexes should be at least .70. Jansky (1978), Gredler (2000), and Kingslake (1983) preferred .75 for both indexes. Carran and Scott (1992) and Plante and Vance (1994) recommended a more rigorous standard of .80 or higher. Jenkins and others (Jenkins, 2003; Jenkins, Hudson, & Johnson, 2007; Johnson, Jenkins, Petscher, & Catts, 2009) recommended that sensitivities be high—perhaps as high as .90—and that specificity levels be relatively high as well.

The receiver operating characteristic/area under the curve (ROC/AUC) “is a measure of the overall performance of a diagnostic test and is interpreted as the average value of sensitivity for all possible values of specificity” (Park, Goo, & Jo, 2004, p. 13). ROC/AUC values range from 0 (representing no predictive ability) to 1 (representing perfect predictive ability). Zhou, Obuchowski, and Obuchowski (2002) recommended that the screening measures designed to distinguish between students with good and poor receptive language ability should have ROC/AUC values that are close to 1. Compton, Fuchs, Fuchs, and Bryant (2006) suggested that ROC/AUCs of .90 and above are excellent, .80 to .89 are good, .70 to .79 are fair, and .69 or below are poor.

Because the Oral Language Index is a measure of general oral language, a series of analyses was conducted to examine its ability to predict other measures of general oral language. The data that were used for these analyses are the same as those used in Tables 3 and 4. In addition, analyses were performed to examine the Oral Language Index's ability to differentiate students in the normative sample who were diagnosed by local school personnel as having specific language impairment ($n = 127$) from a matched sample of students with no language impairment ($n = 127$).

The diagnostic accuracy of the Oral Language Index was investigated at a variety of cutoff points selected to correspond to cutoff scores required by school systems and to maximize sensitivity and specificity. School system policies often require that students score at the extreme low end of the normal distribution (i.e., 1, 1.5, or 2 *SDs* or more below the mean) while researchers (e.g., Dolloghan, 2004; Gray et al., 1999; Merell & Plante, 1997; Plante & Vance, 1994, 1995; Rescorla, 1989; Rescorla & Alley, 2001; Rice & Wexler, 2001; Spaulding et al., 2006) advocate for empirically based cutoff scores that maximize sensitivity and specificity. Betz, Eickoff, and Sullivan (2013), however, argued that, because SLPs are often bound to arbitrary criteria for diagnosing SLI (due to strict guidelines for qualifying children for services), publishers should include the diagnostic accuracy corresponding to a range of cutoff scores. For these reasons, the diagnostic accuracy of the Oral Language Index was examined at six different cutoff scores—index scores of 70 (-2 *SD*), 74 (-1.75 *SD*), 78 (-1.5 *SD*), 85 (-1 *SD*), 90 (-0.7 *SD*), and 92 (-0.5 *SD*). Each of these cutoff scores was used to predict a criterion that was dichotomized into either “at risk” (i.e., standard score below 90) or “not at risk” (standard score 90 or above) based on the student's scores on the criterion measures. In an additional study, each of these cutoff scores was used to analyze the Oral Language Index's diagnostic accuracy in identifying students who received a clinical diagnosis of language impairment from those who had no diagnosis of language impairment.

Using the two dichotomous groups that were created based on the selected cutoff scores, thirty-six 2×2 frequency matrices were created—six for each criterion. Table 5 reports the results of the diagnostic accuracy analyses. The Oral Language Index met minimal criteria for diagnostic accuracy when compared to other measures of general oral language and when used to differentiate students who had been diagnosed in a school program as having a specific language impairment diagnosis from those who did not.

In all comparisons, the cutoff score that best maximized sensitivity and specificity was a cutoff score of 90 or 92 (i.e., -0.7 and -0.5 *SD*, respectively). When using cutoff scores of 90 to 92, the Oral Language Index's diagnostic accuracy meets or exceeds the minimum standards recommended by the authorities mentioned earlier in this section. The medians reported at the bottom of the table are particularly encouraging. These findings suggest that, across a variety of measures and diagnostic categories, the Oral Language Index is a valid and reliable predictor of students exhibiting symptoms of language impairment.

Construct-Identification Validity

Another way of establishing a test's validity is to study the performance of different groups of people on the test. For example, in the case of the TACL-4/TEXL composites, one would expect that the mean standard scores of different gender subgroups (males, females) and different racial subgroups (e.g., Whites and Black/African Americans) would be in the average range (90–109 for index score). Moreover, one would expect that means for groups of people who we know or suspect of having a language disability (i.e., learning

Table 5
Diagnostic Accuracy Analyses for the TACL-4/TEXL Oral Language Index

Criterion	Cutoff index score	SD	Percentile rank	Sensitivity index	Specificity index	ROC/AUC	Classification accuracy	True positives	False positives	True negatives	False negatives
OWLS-II Oral Language (n = 68)	92	−0.5	30	.90	.96	.95	.93	38	1	25	4
	90	−0.7	25	.74	.96		.82	31	1	25	11
	85	−1	16	.57	.96		.72	24	1	25	18
	78	−1.5	7	.21	1.00		.51	9	0	26	33
	74	−1.75	4	.12	1.00		.46	5	0	26	37
	70	−2	2	.07	1.00		.43	3	0	26	39
CELF Core Language (n = 68)	92	−0.5	30	.76	.82	.92	.78	35	4	18	11
	90	−0.7	25	.67	.95		.76	31	1	21	15
	85	−1	16	.54	1.00		.69	25	0	22	21
	78	−1.5	7	.20	1.00		.46	9	0	22	37
	74	−1.75	4	.11	1.00		.40	5	0	22	41
	70	−2	2	.07	1.00		.37	3	0	22	43
CELF Language Content (n = 68)	92	−0.5	30	.79	.77	.92	.78	33	6	20	9
	90	−0.7	25	.74	.96		.82	31	1	25	11
	85	−1	16	.60	1.00		.75	25	0	26	17
	78	−1.5	7	.21	1.00		.51	9	0	26	33
	74	−1.75	4	.12	1.00		.46	5	0	26	37
	70	−2	2	.07	1.00		.43	3	0	26	39
CELF Language Structure (n = 63)	92	−0.5	30	.77	.80	.91	.78	33	4	16	10
	90	−0.7	25	.67	.95		.76	29	1	19	14
	85	−1	16	.53	1.00		.68	23	0	20	20
	78	−1.5	7	.21	1.00		.46	9	0	20	34
	74	−1.75	4	.12	1.00		.40	5	0	20	38
	70	−2	2	.07	1.00		.37	3	0	20	40
DAB-4 Spoken Language (n = 34)	92	−0.5	30	.67	.89	.94	.85	4	3	24	2
	90	−0.7	25	.67	.93		.88	4	2	25	2
	85	−1	16	.33	.96		.85	2	1	26	4
	78	−1.5	7	.17	1.00		.85	1	0	27	5
	74	−1.75	4	.17	1.00		.85	1	0	27	5
	70	−2	2	.00	1.00		.85	0	0	27	6

(continues)

Table 5 (continued)

Criterion	Cutoff index score	SD	Percentile rank	Sensitivity index	Specificity index	ROC/AUC	Classification accuracy	True positives	False positives	True negatives	False negatives
Language Impairment (<i>n</i> = 254)	92	−0.5	30	.75	.87	.90	.81	95	16	111	32
	90	−0.7	25	.68	.93		.80	86	9	118	41
	85	−1	16	.58	.99		.78	73	1	126	54
	78	−1.5	7	.32	1.00		.66	40	0	127	87
	74	−1.75	4	.24	1.00		.62	30	0	127	97
	70	−2	2	.15	1.00		.58	19	0	127	108
Summary											
Median	92	−0.5	30	.77	.85	.92	.80				
	90	−0.7	25	.68	.95		.81				
	85	−1	16	.56	1.00		.74				
	78	−1.5	7	.21	1.00		.51				
	74	−1.75	4	.12	1.00		.46				
	70	−2	2	.07	1.00		.43				

Note. *SD* = standard deviation; ROC/AUC = receiver operating characteristic/area under the curve. Children in Study 1 were administered either the CELF-4 or the CELF Preschool-2, depending on the age of the examinee. For the purposes of this study, the CELF-4 and the CELF Preschool-2 have been treated as one measure and will be referred to as the CELF. OWLS-II = *Oral and Written Language Scales, Second Edition* (Carrow-Woolfolk, 2011); CELF = *Clinical Evaluation of Language Fundamentals, Fourth Edition* (Semel, Wiig, & Secord, 2003) and *Clinical Evaluation of Language Fundamentals Preschool, Second Edition* (Wiig, Secord, & Semel 2006); DAB-4 = *Diagnostic Achievement Battery, Fourth Edition* (Newcomer, 2014).

disabilities, specific language impairment, deaf or hard of hearing, autism spectrum disorder, intellectual disability) would be lower (i.e., below 90).

The mean composite indexes for selected subgroups within the TACL-4/TEXL normative sample are listed in Table 6. In all, we have data on two “mainstream” subgroups (male and female), six racial and ethnic subgroups (White, Black/African American, Asian/Pacific Islander, American Indian, Hispanic, and two or more races), and eight “exceptionality” subgroups (gifted and talented, attention-deficit/hyperactivity disorder, articulation disorder, learning disability, language impairment, deaf and hard of hearing, autism spectrum disorder, and intellectual disability). We also had two groups of examinees from different linguistic backgrounds (fluent in English and another language and English as a second language). The size of the means for the subgroups conforms to expectations and provides strong support for the construct-identification validity of the TACL-4/TEXL composites.

Summary of Psychometric Properties

Based on the information provided in this section, one may conclude that the TACL-4/TEXL composite scores are reliable and valid measures of general oral language. Examiners can interpret these scores with confidence. We encourage professionals to continue to study the tests using different samples, statistical procedures, and related measures. We also encourage these researchers to share their results with us so that their findings can be included in subsequent editions of the tests. The accumulation of research data will help further clarify the reliability and validity of the TACL-4 and the TEXL and provide guidance for their future revisions.

Table 6
The TACL-4/TEXL Composite Standard Score Means for Selected Subgroups
(Decimals Omitted)

Subgroup	TACL-4/TEXL composite values			
	Vocabulary	Morphology	Syntax	Oral Language
Gender				
Male (<i>n</i> = 656)	99 (15)	99 (15)	98 (15)	98 (15)
Female (<i>n</i> = 637)	100 (15)	101 (15)	101 (15)	101 (14)
Race/ethnicity				
White (<i>n</i> = 1,058)	101 (15)	101 (15)	101 (15)	100 (14)
Black/African American (<i>n</i> = 151)	93 (13)	93 (14)	94 (14)	92 (14)
Asian/Pacific Islander (<i>n</i> = 8)	106 (12)	105 (10)	105 (8)	105 (10)
American Indian (<i>n</i> = 36)	104 (16)	98 (14)	98 (17)	100 (16)
Hispanic (<i>n</i> = 213)	94 (15)	95 (15)	97 (14)	95 (14)
Two or more races (<i>n</i> = 40)	96 (16)	96 (16)	96 (15)	95 (16)
Exceptionality				
Gifted and talented (<i>n</i> = 72)	113 (11)	112 (10)	113 (11)	113 (10)
Attention-deficit/hyperactivity disorder (<i>n</i> = 40)	95 (14)	94 (15)	92 (15)	93 (15)
Articulation disorder (<i>n</i> = 147)	93 (15)	93 (16)	93 (15)	92 (15)
Learning disability (<i>n</i> = 42)	88 (13)	90 (14)	87 (13)	87 (13)
Language impairment (<i>n</i> = 127)	86 (13)	84 (13)	83 (12)	83 (13)
Deaf and hard of hearing (<i>n</i> = 45)	81 (15)	80 (16)	84 (12)	80 (15)
Autism spectrum disorder (<i>n</i> = 38)	80 (18)	75 (15)	77 (17)	74 (17)
Intellectual disability (<i>n</i> = 51)	76 (15)	73 (12)	77 (14)	73 (14)
Linguistic background				
Fluent in English and another language (<i>n</i> = 91)	95 (15)	95 (14)	97 (14)	95 (13)
English as a second language (<i>n</i> = 4)	76 (6)	68 (6)	72 (7)	69 (4)

References

- Aiken, L. R., & Groth-Marnat, G. (2008). *Psychological testing and assessment* (12th ed.). Needham Heights, MA: Allyn & Bacon.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Anderson, J. D., Hess, R., & Richardson, K. (1980). Test-retest reliability of the test for auditory comprehension of language when it is used with mentally retarded children. *Journal of Speech and Hearing Disorders, 45*, 195–200.
- Bankson, N. (1989). Review of the Test for Auditory Comprehension of Language–Revised. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 822–824). Lincoln: The University of Nebraska Press.
- Bartel, N. R., Bryen, D., & Keehn, S. (1973). Language comprehension in the moderately retarded child. *Exceptional Children, 39*, 375–382.
- Boehm, A. E. (1971). *Boehm test of basic concepts*. New York, NY: Wiley.
- Burrows, E. H., & Neyland, D. (1978). Reading skills, auditory comprehension of language and academic achievement. *Journal of Speech and Hearing Disorders, 43*, 467–473.
- Carrow, E. (1971). Comprehension of English and Spanish by preschool Mexican-American children. *Modern Language Journal, 55*, 299–305.
- Carrow, E. (1972). Auditory comprehension of bilingual and monolingual preschool children. *Journal of Speech and Hearing Research, 15*, 407–412.
- Carrow, E. (1973). *Test for auditory comprehension of language*. Austin, TX: Learning Concepts.
- Carrow, E. (1974). *Carrow elicited language inventory*. Allen, TX: DLM.
- Carrow, E., & Lynch, J. I. (1973, November). *Comparison of auditory comprehension in three groups of linguistically deviant children*. Paper presented at the annual meeting of the American Speech and Hearing Association.
- Carrow, E., & Maudlin, M. (1973). Children's recall of approximations to English. *Journal of Speech and Hearing Research, 16*, June.
- Carrow, M. A. (1968). The development of auditory comprehension of language structure in children. *Journal of Speech and Hearing Disorders, 33*, 99–111.
- Carrow-Woolfolk, E. (1985). *Test for auditory comprehension of language* (Rev. ed.). Allen, TX: DLM.
- Carrow-Woolfolk, E. (1999). *Test for auditory comprehension of language* (3rd ed.). Austin, TX: PRO-ED.
- Carrow-Woolfolk, E. (2014). *Test for auditory comprehension of language* (4th ed.). Austin, TX: PRO-ED.
- Carrow-Woolfolk, E., & Allen, E. A. (2014). *Test of expressive language*. Austin, TX: PRO-ED.
- Carrow-Woolfolk, E., & Lynch, J. I. (1982). *An integrative approach to language disorders in children*. New York, NY: Grune & Stratton.
- Compton, C. (1980). *A guide to 65 tests for special education*. Belmont, CA: Pitman Learning.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.
- Davis, J. M. (1977). Reliability of hearing-impaired children's responses to oral and total presentations of the test for auditory comprehension of language. *Journal of Speech and Hearing Disorders, 42*, 520–527.
- Dean, E. A. (1975). *Language comprehension as a predictor of articulation skills in kindergarten-age children* (Unpublished master's thesis). University of Tennessee, Knoxville.
- DeRenzi, E., & Vignolo, L. A. (1962). The DeRenzi Token Test: A sensitive test to detect receptive disturbances in aphasia. *Brain, 85*, 556–678.
- DiSimoni, R. (1978). *The token test for children*. Allen, TX: DLM.
- Dunn, L. M., & Dunn, L. (1965). *Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.
- Emerick, L. L., & Hatten, J. J. (1979). *Diagnosis and evaluation in speech pathology*. Englewood Cliffs, NJ: Prentice-Hall.
- Evard, B. L., & Sabers, D. L. (1979). Speech and language testing with distinct ethnic-racial groups: A survey of procedures for improving test validity. *Journal of Speech and Hearing Disorders, 46*, 91–96.
- Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Mascolo, J. T. (2002). *The achievement test desk reference: Comprehensive assessment and learning disabilities*. Boston, MA: Allyn and Bacon.
- Foster, R., Giddan, J. J., & Stark, J. (1972). *Assessment of children's language comprehension*. Palo Alto, CA: Consulting Psychologists Press.
- Gates, A., & MacGinitie, W. (1965). *Gates-MacGinitie reading tests*. Chicago, IL: Riverside.
- Harris, L. G., & Shelton, I. S. (1993). *Desk reference for assessments in speech-language*. San Antonio, TX: Communication Skill Builders.

- Hatten, J. T. (1978). Test for Auditory Comprehension of Language. In O. E. Buros (Ed.), *The eighth mental measurements yearbook* (pp. 611–612). Highland Park, NJ: Gryphon Press.
- Haynes, W. D. (1989). Review of the Test for Auditory Comprehension of Language–Revised. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 824–826). Lincoln: The University of Nebraska Press.
- Hedrick, D. L., Prather, E. M., & Tobin, A. R. (1975). *Sequenced inventory of communication development*. Seattle: University of Washington Press.
- Jones, B. (1972). *A study of oral language comprehension of black and white, middle and lower class pre-school children using standard English and black dialect in Houston, Texas, 1972* (Unpublished doctoral dissertation). University of Houston, Texas.
- Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois test of psycholinguistic abilities*. Champaign: University of Illinois Press.
- Koenigsknecht, R. A., & Friedman, P. (1976). Syntax development in boys and girls. *Child Development*, 47, 1109–1115.
- Lambert, N. (Cons. Ed.). (1981). *Special education assessment matrix manual*. Special Education Assessment Coalition. Monterey, CA: CTB-McGraw Hill.
- Lee, L. (1969). *Northwestern syntax screening test*. Evanston, IL: Northwestern University Press.
- Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL: Northwestern University Press.
- Little, J. W., Moon, C. E., & Contraras, M. (1980). A comparative study of the language comprehension of rural American Indian and rural Caucasian kindergarten age children. *Early Child Development and Care*, 6(3-4), 201–206.
- Lloyd, L. L. (Ed.). (1976). *Communication assessment and intervention strategies*. Baltimore, MD: University Park Press.
- Madden, R., & Gardner, E. F. (1967). *Stanford early school achievement test*. Cleveland, OH: The Psychological Corporation.
- Manikam, R. (2001). Review of the Test for Auditory Comprehension of Language, Third Edition. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 1229–1230). Lincoln: The Buros Institute of Mental Measurements, University of Nebraska Press.
- Marquardt, T. P., & Saxman, J. H. (1972). Language comprehension and auditory discrimination in articulatory deficient kindergarten children. *Journal of Speech and Hearing Research*, 15, 382–389.
- Martin, R. F. (1977). A review of Test for Comprehension of Language, Fifth Edition. *Modern Language Journal*, 61, 427–428.
- McCarthy, D. (1972). *Manual: McCarthy scales of children's abilities*. New York, NY: The Psychological Corporation.
- McLean, J. E., & Snyder, L. K. (1978). *A transactional approach to early language training*. Columbus, OH: Charles E. Merrill.
- Mellenberg, G. J. (1983). *Conditional item bias methods*. New York, NY: Plenum.
- Menyuk, P. (1969). *Sentences children use*. Cambridge, MA: M.I.T. Press.
- Millen, C. E., & Prutting, C. A. (1979). Inconsistencies across three language comprehension tests for specific grammatical features. *Language, Speech, and Hearing Services in Schools*, 10, 162–170.
- Miller, J. F. (1979). Test for Auditory Comprehension of Language. In F. Darley (Ed.), *Evaluation of appraisal techniques in speech and language pathology* (p. 118). Boston, MA: Addison-Wesley.
- Musselwhite, C. R. (1983). Pluralistic assessment in speech-language pathology: Use of dual norms in the placement process. *Language, Speech, and Hearing Services in Schools*, 14, 29–37.
- Novak, C. (2001). Review of the Test for Auditory Comprehension of Language, Third Edition. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 1230–1231). Lincoln: The Buros Institute of Mental Measurements, University of Nebraska Press.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Nurse, J. R., & Day, D. E. (1971). Imitation, comprehension and production of grammatical structures. *Journal of Verbal Learning and Verbal Behavior*, 10, 68–74.
- Peterson, H. A., & Marquardt, T. P. (1981). *Appraisal and diagnosis of speech and language disorders*. Englewood Cliffs, NJ: Prentice-Hall.
- Pierangelo, R., & Giuliani, G. (1998). *Special educator's complete guide to 109 diagnostic tests*. West Nyack, NJ: The Center for Applied Research in Education.
- Pierangelo, R., & Giuliani, G. (2002). *Assessment in special education: A practical approach*. Boston, MA: Allyn and Bacon.
- ProQuest LLC. (2012). *ProQuest statistical abstracts of the United States, 2013*. Bethesda, MD: Bernan.
- Reynell, J. (1977). *Reynell developmental language scales* (Rev. ed.). Windsor, Berkshire, UK: NFER.
- Reynolds, C. R., Wright, D., & Wilkinson, W. A. (1980). Incremental validity of the Test for Auditory Comprehension of Language and the Developmental Test of Visual Motor Integration. *Educational and Psychological Measurement*, 40, 503–507.

- Rizzo, J. M., & Stephens, M. I. (1981). Performance of children with normal and impaired oral language production on a set of auditory comprehension tests. *Journal of Speech and Hearing Disorders, 46*, 150–160.
- Salvia, J., Ysseldyke, J. E., & Bolt, S. (2013). *Assessment in special and inclusive education* (12th ed.). Australia: Wadsworth Cengage Learning.
- Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler.
- Schmitt, J. F. (1987). Test for Auditory Comprehension of Language-Revised. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques: Volume VI* (pp. 586–593). Austin, TX: PRO-ED.
- Science Research Associates. (1978). *Science research associates achievement battery*. Monterey, CA: CTB Macmillan/McGraw-Hill.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 26*, 55–66.
- Terman, L. M., & Merrill, M. A. (1960). *Stanford–Binet intelligence scale* (3rd ed.; Form L-M). Boston, MA: Houghton-Mifflin.
- U.S. Bureau of the Census. (1997). *Statistical abstract of the United States*. Washington, DC: Author.
- Venn, J. J. (2004). *Assessing students with special needs* (3rd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Wallace, G., & Hammill, D. D. (1994). *Comprehensive receptive and expressive vocabulary test*. Austin, TX: PRO-ED.
- Weiner, P. S. (1972). The perceptual level functioning of dysphasic children: A follow-up study. *Journal of Psycholinguistic Research, 3*, 271–280.
- Wiig, E., & Semel, E. M. (1980). *Language assessment and intervention*. Columbus, OH: Merrill.
- Wright, D., Reynolds, C. R., & Dappen, L. (1980). Criterion-related validity of three common preschool assessment instruments for boys and girls. *Psychological Reports, 47*, 1291–1296.
- Zinkus, P. W., & Gottlieb, M. I. (1980). Patterns of perceptual and academic disorders related to early chronic otitis media. *Pediatrics, 66*, 246–253.

Appendix A

Converting Sums of TACL-4 and TEXL Subtest Scaled Scores to Percentile Ranks and Vocabulary, Morphology, Syntax, and Oral Language Composite Indexes

Table A.1
Converting Sums of TACL-4 and TEXL Scaled Scores to Percentile Ranks and Index Scores

Percentile rank	Sum of 2 Vocabulary subtests	Sum of 2 Grammatical Morpheme subtests	Sum of 2 Elaborated Phrases and Sentences subtests	Index score
<1	2			54
<1	3	2	2	56
<1			3	58
<1	4	3	4	61
<1	5	4		62
<1			5	63
1	6	5		64
1			6	66
1	7	6		67
2		7	7	69
2	8			70
3			8	71
3	9	8		72
4		9	9	74
5	10			75
5			10	76
6		10		77
7	11			78
8		11	11	79
9	12			80
10			12	81
12		12		82
13	13			83
14		13	13	84
18	14			86
19		14	14	87
21	15			88
23		15	15	89
27	16			91
30		16	16	92
35	17	17	17	94
39	18			96
42		18	18	97
45	19			98
47		19	19	99
50	20	20	20	100
58	21	21	21	103
68	22	22	22	107
75	23	23	23	110
79	24	24		112

Table A.1 (continued)

Percentile rank	Sum of 2 Vocabulary subtests	Sum of 2 Grammatical Morpheme subtests	Sum of 2 Elaborated Phrases and Sentences subtests	Index score
81			24	113
84	25	25	25	115
87		26		117
89	26		26	118
91	27	27	27	120
93		28		122
94	28		28	123
95		29		125
96	29		29	126
97		30		127
97	30		30	128
98		31		130
98	31		31	131
99		32	32	133
99	32			134
99		33		135
>99	33		33	136
>99		34	34	138
>99	34			139
>99		35		140
>99			35	141
>99	35			142
>99		36		143
>99	36		36	144
>99		37		145
>99			37	146
>99	37			147
>99		38		148
>99			38	149
>99	38	39		150
>99			39	151
>99	39			152
>99		40		153
>99			40	154
>99	40			155

Table A.2

Converting Sums of the TACL-4 and TEXL Scaled Scores to Percentile Ranks and Oral Language Index Scores

Percentile rank	Sum of 6 scaled scores	Oral Language Index	Percentile rank	Sum of 6 scaled scores	Oral Language Index
<1	6	50	19	44	87
<1	7	51	21	45	88
<1	8	52	23	46	89
<1	9	53	25	47	90
<1	10	54	27	48	91
<1	11	55	30	49	92
<1	12	56	32	50	93
<1	13	57	35	51	94
<1	14	58	37	52	95
<1	15	59	37	53	95
<1	16	60	39	54	96
<1	17	61	39	55	96
<1	18	62	42	56	97
<1	19	63	45	57	98
<1	20	64	45	58	99
1	21	65	47	59	99
1	22	66	50	60	100
1	23	67	53	61	101
1	24	68	58	62	103
2	25	69	61	63	104
2	26	69	63	64	105
2	27	70	65	65	106
3	28	71	68	66	107
3	29	72	70	67	108
3	30	73	73	68	109
4	31	74	75	69	110
5	32	75	77	70	111
5	33	76	79	71	112
6	34	77	81	72	113
7	35	78	82	73	114
8	36	79	84	74	115
9	37	80	86	75	116
10	38	81	87	76	117
12	39	82	89	77	118
13	40	83	90	78	119
14	41	84	91	79	120
16	42	85	92	80	121
18	43	86	92	81	121

Table A.2 (continued)

Percentile rank	Sum of 6 scaled scores	Oral Language Index
93	82	122
94	83	123
95	84	124
95	85	125
96	86	126
97	87	127
97	88	128
97	89	129
98	90	130
98	91	131
99	92	132
99	93	133
99	94	134
99	95	135
>99	96	136
>99	97	137
>99	98	138
>99	99	139
>99	100	140
>99	101	141
>99	102	142
>99	103	143
>99	104	144
>99	105	145
>99	106	146
>99	107	147
>99	108	148
>99	109	149
>99	110	150
>99	111	151
>99	112	152
>99	113	153
>99	114	154
>99	115	155
>99	116	156
>99	117	157
>99	118	158
>99	119	159
>99	120	160

Appendix B

Cumulative Percentages (Frequencies) of
Receptive Language–Expressive Language
Differences Across Ages

Table B.1
Cumulative Percentages (Frequencies) of Receptive Language–Expressive Language Differences Across Ages

Amount of difference (absolute value)	Cumulative percent
0	92.9
1	92.8
2	79.2
3	78.7
4	66.4
5	65.2
6	54.5
7	52.9
8	43.2
9	42.1
10	35.0
11	33.7
12	26.8
13	25.6
14	21.5
15	19.6
16	16.2
17	15.0
18	11.7
19	11.1
20	9.0
21	8.3
22	6.2
23	5.3
24	4.0
25	3.6
26	2.8
27	2.2
28	1.9
29	1.6
30	1.3
31	0.9
32	0.8
34	0.6
37	0.2
38	0.2
39	0.1
42	0.0

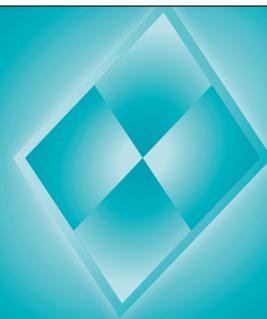
Appendix C

TACL-4/TEXL Summary Form

TACL-4/TEXL

Summary Form

Elizabeth Carrow-Woolfolk and Elizabeth A. Allen



Section 1. Identifying Information

Name _____ Female Male Grade _____

Year _____ Month _____ School _____

Age at TACL-4 Testing _____ Examiner's Name _____

Age at TEXL Testing _____ Examiner's Title _____

Reason for Testing _____

Section 2. Summary of Subtest Scores

TACL-4/TEXL Composites	Subtest Scaled Scores						Sum of Scaled Scores
	TACL-4			TEXL			
	V	GM	EPS	V	GM	EPS	
Vocabulary							
Morphology							
Syntax							
Receptive Language							
Expressive Language							
Oral Language							

Section 3. Summary of Composite Scores

	Composite Scores	Sum of Scaled Scores	%ile Rank	Index Score	SEM	Confidence Interval	Descriptive Term
Linguistic Features	Vocabulary Index (VI)	_____	_____	<input type="text"/>	3	____ - ____	_____
	Morphology Index (MI)	_____	_____	<input type="text"/>	3	____ - ____	_____
	Syntax Index (SI)	_____	_____	<input type="text"/>	3	____ - ____	_____
Linguistic Systems	Receptive Language Index (RLI)	_____	_____	<input type="text"/>	3	____ - ____	_____
	Expressive Language Index (ELI)	_____	_____	<input type="text"/>	3	____ - ____	_____
	Oral Language Index (OLI)	_____	_____	<input type="text"/>	3	____ - ____	_____

© 2014 by PRO-ED, Inc. PRO-ED, Inc. grants permission to reproduce this form for use with your clients. This license does not grant the right to reproduce these materials for resale, redistribution, or any other purposes. This Summary Form is available as a free PDF download on the TACL-4 and TEXL product pages at www.proedinc.com.

Section 4. Language Index Comparisons

	Index Score		Difference	Statistically Significant	Clinically Meaningful	% of Sample With This Difference
	RLI	ELI				
RLI vs. ELI Comparison	_____	- _____	= <input type="text"/>	>6	>20	_____

Section 5. Profile of Standard Scores

	Subtest Scaled Scores						Composite Indexes					OLI	
	TACL-4			TEXL			Linguistic Features			Linguistic Systems			
	V	GM	EPS	V	GM	EPS	VI	MI	SI	RLI	ELI		
													160
													155
20	•	•	•	•	•	•							150
19	•	•	•	•	•	•							145
18	•	•	•	•	•	•							140
17	•	•	•	•	•	•							135
16	•	•	•	•	•	•							130
15	•	•	•	•	•	•							125
14	•	•	•	•	•	•							120
13	•	•	•	•	•	•							115
12	•	•	•	•	•	•							110
11	•	•	•	•	•	•							105
10	•	•	•	•	•	•							100
9	•	•	•	•	•	•							95
8	•	•	•	•	•	•							90
7	•	•	•	•	•	•							85
6	•	•	•	•	•	•							80
5	•	•	•	•	•	•							75
4	•	•	•	•	•	•							70
3	•	•	•	•	•	•							65
2	•	•	•	•	•	•							60
1	•	•	•	•	•	•							55
													50

Section 6. Descriptive Terms

Scaled Score	1–3	4–5	6–7	8–12	13–14	15–16	17–20
Descriptive Term	Impaired or Delayed	Borderline Impaired or Delayed	Below Average	Average	Above Average	Superior	Gifted or Very Advanced
Index Score	<70	70–79	80–89	90–109	110–119	120–129	>129

© 2014 by PRO-ED, Inc. PRO-ED, Inc. grants permission to reproduce this form for use with your clients. This license does not grant the right to reproduce these materials for resale, redistribution, or any other purposes. This Summary Form is available as a free PDF download on the TACL-4 and TEXL product pages at www.proedinc.com.