

Language, Phonological Awareness, and Reading Test Directory

Denyse V. Hayward, G. Eleanor Stewart, Linda M. Phillips,
Stephen P. Norris, and Meridith A. Lovell

Canadian Centre for Research on Literacy
&
Canadian Language and Literacy Research Network



Réseau canadien de recherche
sur le langage et l'alphabétisation



CANADIAN LANGUAGE &
LITERACY RESEARCH NETWORK

Partager la science. Éveiller les esprits.

sharing the science. opening minds.

Introduction to the Test Reviews

List of Tests Reviewed	2
General Language Tests.....	2
Vocabulary and Grammar Tests	2
Narrative Tests	2
Phonological Awareness Tests	3
Reading Tests.....	3
Writing and Achievement Tests.....	3
Introduction to the Test Reviews	4
Rationale for Conducting the Reviews	4
Content of the Reviews	5
Using the Language, Phonological Awareness, and Reading Test Directory	9
References	10

List of Tests Reviewed

* Indicates that the review is not yet complete.

General Language Tests

<i>Clinical Evaluation of Language Fundamentals – 4 (CELF-4)</i>	At-A-Glance	Test Review
<i>Comprehensive Assessment of Spoken Language (CASL)</i>	At-A-Glance	Test Review
<i>Preschool Language Assessment Instrument – 2 (PLAI-2)</i>	At-A-Glance	Test Review
<i>Preschool Language Scale – 4 (PLS-4)</i>	At-A-Glance	Test Review
<i>Test of Language Development – Intermediate 3rd Edition (TOLD-I3)</i>	At-A-Glance	Test Review
<i>Test of Language Development – Primary 3rd Edition (TOLD-P3)</i>	At-A-Glance	Test Review

Vocabulary and Grammar Tests

<i>Expressive Vocabulary Test (EVT)</i>	At-A-Glance	Test Review
<i>Peabody Picture Vocabulary Test (PPVT-III)</i>	At-A-Glance	Test Review
<i>Structured Photographic Expressive Language Test – 3 (SPELT-3)</i>	At-A-Glance	Test Review

Narrative Tests

<i>The Renfrew Bus Story (RBS)</i>	At-A-Glance	Test Review
<i>The Test of Narrative Language (TNL)</i>	At-A-Glance	Test Review

Phonological Awareness Tests

<i>Comprehensive Test of Phonological Processing (CTOPP)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Lindamood Auditory Conceptualization Test – 3 (LAC-3)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Phonological Awareness Test – 2 (PAT-2)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Pre-Reading Inventory of Phonological Awareness (PIPA)</i>	<u>At-A-Glance</u>	<u>Test Review</u>

Reading Tests

<i>Emerging Literacy and Language Assessment (ELLA)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Gray Oral Reading Test – 4 (GORT-4)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Oral and Written Language Scales Listening Comprehension (OWLS)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Test of Early Reading Ability – 3 (TERA-3)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
* <i>Test of Preschool Early Literacy (TOPEL)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Test of Word Reading Efficiency (TOWRE)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Woodcock-Johnson Reading Mastery Tests – Revised (WRMT-R)</i>	<u>At-A-Glance</u>	<u>Test Review</u>

Writing and Achievement Tests

<i>Brigance Comprehensive Inventory of Basic Skills – Revised (CIBS-R)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Test of Early Written Language – 2 (TEWL-2)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Test of Written Language –3 (TOWL-3)</i>	<u>At-A-Glance</u>	<u>Test Review</u>
<i>Woodcock-Johnson Test of Achievement III (WJ-III)</i>	<u>At-A-Glance</u>	<u>Test Review</u>

Introduction to the Test Reviews

In researching the body of standardized tests of reading, writing, and other language skills currently available, the *Test of Early Language and Literacy (TELL)* group from the Canadian Centre for Research on Literacy (CCRL) produced a comprehensive review of tests which are commonly used by reading specialists, speech language pathologists, and special education teachers to assess the oral language, reading, writing, and other language skills of children. These tests are divided into six groups, based upon the main skills being assessed and include: (1) *General Language Tests*, (2) *Vocabulary and Grammar Tests*, (3) *Narrative Tests*, (4) *Phonological Awareness Tests*, (5) *Reading Tests*, and (6) *Writing and Achievement Tests*. The “At-a-Glance” summaries and full Test Reviews are presented on this site so that reading specialists, speech language pathologists, special education teachers, and researchers may use the reviews to make more informed decisions about which standardized tests they may wish to use with children.

Rationale for Conducting the Reviews

Standardized tests, especially in Language Arts, are frequently used for assessing children and planning intervention programs to address areas of need. Having access to a comprehensive test that accurately and efficiently assesses children’s areas of strength and need, will help reading specialists, speech language pathologists, and special education teachers make more informed decisions when planning programs.

A lack of available effective assessment instruments in either reading, writing, or other language areas often prompts the creation of a new standardized test. TELL creators recognized that no tests were available to assess both the language and literacy skills of children between three and eight years of age. In arriving at this conclusion, TELL creators undertook a review of

many current and widely used reading, writing, and language tests to determine their areas of focus, their strengths, and any areas of weakness. When embarking on any project, such as the creation of a new standardized test, it is important to survey existing tests for several reasons:

- 1) Test creators must determine what types of skills are measured by other standardized tests, how these skills are measured, and whether there are any gaps in what is being assessed.
- 2) Understanding what procedures were used for testing, norming, and evaluating other standardized tests helps test creators to establish their own procedures for these activities.
- 3) Often other tests which measure the same or similar skills as the new test are used for the purpose of validity testing, and having a thorough understanding of the content of the other tests allows the test creators to make informed decisions about which tests will be used and how validity testing will be undertaken.

Content of the Reviews

Several areas were considered for each test which was evaluated. Each of the major categories addressed in this review is briefly explained, and any unexpected findings are also presented. Determination of the sample size and characteristics, the purpose and theoretical basis, what the test will measure, how it will be administered, how results are to be interpreted, and whether the test is reliable and valid are each important considerations when assessing the design of any standardized test, and these considerations are outlined next.

First the Norming Sample was considered. The reviewers looked at sample size, location, and demographic information. It is interesting to note that tests had vastly different sample sizes, although most test creators attempted to stratify their samples to match U.S. census data as

closely as possible. From the TELL team's perspective, the failure of several tests to include or report data for students with various learning, language, or physical disabilities was a troubling omission.

After assessing the norming sample, reviewers considered the authors' stated purposes, theoretical models, and background research. Reviewers found that most, but not all, tests stated some sort of theoretical basis and background research from the literature. Several tests are revised editions of previous reading, writing, or other language tests, and in these cases, often the theoretical models remain unchanged from up to thirty years ago because the same test, although updated through the years, has remained in the same format with similar test items. In these instances, the theoretical basis for a test might not match current theories of language and literacy development.

Standardized tests of reading, writing, or other language skills are seldom comprehensive tests; they assess one or more aspects of their focus areas, but no test covers the entire spectrum of literacy skills. Each test in the six testing categories (General Language Tests, Vocabulary and Grammar Tests, Narrative Tests, Phonological Awareness Tests, Reading Tests, and Writing and Achievement Tests) addresses slightly different skills. A base list of skills was developed after reviewing the literature:

- 1) Oral language skills (tests of vocabulary, grammar, or narrative),
- 2) Print knowledge skills (tests of environmental print or alphabet awareness),
- 3) Phonological awareness skills (tests of segmenting, blending, elision, or rhyming),
- 4) Reading skills (tests of single word reading, decoding, or comprehension abilities),
- 5) Spelling skills,

- 6) Writing skills (tests of letter formation, capitalization, punctuation, conventional structures, word choice, or details), and
- 7) Listening skills (tests of lexical, syntactic, or supralinguistic skills).

This skill list addressed most of the skills listed by test authors, however, in some cases authors listed other types of skills and these were added.

The test administration information was also of interest to the TELL group. Most of the standardized tests of language that were reviewed may be administered by speech language pathologists (SLP), reading specialists, special education teachers, or other individuals with graduate level training in assessment procedures. Test administration times ranged from five minutes to over an hour, in some cases, depending upon which subtests were administered. Administration procedures, basal and ceiling rules (when applicable), materials, recording, and scoring procedures were fairly consistent among the various tests.

Of greater importance to the TELL group, or to those looking to administer various tests of reading, writing, or other language skills, is how different test manuals explain the interpretation of test results. In the interpretation, the scores obtained from all of the various subtests are translated into judgements about abilities, areas of strength, and areas of need, so that interventions can be planned if needed. Most of the reviewed tests have a chapter or two in their manuals discussing interpretation of results. Often, examples of completed scoring sheets and the accompanying interpretation are provided to explain what the results indicate so that the professional administering the test can understand what the results mean and then communicate results and recommendations to parents or guardians, or any other individuals involved. As all of the tests reviewed were standardized tests of reading, writing, or other language skills, the results were often in the form of age equivalent scores, grade equivalent scores, percentiles, a

standardized score, a stanine score, or a composite score or quotient. In several test manuals, authors caution against the use of age or grade equivalency scores, but still include these scores due to legislation mandating their use in many jurisdictions. When they are included in the results, age and grade equivalent scores should be used with caution, as they are easily misinterpreted. According to an International Reading Association (1981, p. 1) resolution, "... The International Reading Association strongly advocates that those who administer standardized reading tests abandon the practice of using grade equivalents to report performance of either individuals or groups of test-takers ... [and urges publishers] to eliminate grade equivalents from their tests". The resolution cites examples: a grade equivalent score of 5.0 does not necessarily mean that the reader can read fifth grade material, and a grade equivalent score of 10.0 awarded to a fourth grade student does not mean that child reads like a tenth grade student. That such scores are still mandated in many jurisdictions, although better methods of reporting and interpreting results exist and the IRA denounces grade equivalent scores, is troubling.

Reliability and validity testing were the final categories addressed by the TELL group reviewers. The importance of adequate reliability and validity testing cannot be overstated, as authors use them to ensure their tests measure what they are intended to measure. Reliability testing (internal consistency of items, test-retest results, and inter-rater reliability testing) ensures that the test items are suitable for testing what they are intended to test, that the same students taking the test will consistently get similar results, and that the scoring is fair and objective, to the extent possible. An interesting pattern emerged from the inter-rater reliability, which was not addressed in most of the manuals: When performing inter-rater reliability testing, most authors settled for having two or three trained raters score completed test protocols, rather than having the raters record the results while watching the assessment taking place. Having two raters record

and then score the same student's performance would seem to be a much more stringent measure of inter-rater reliability, however, few tests report reliability testing in this manner.

Validity testing (content validity, criterion prediction validity, construct validity, and differential item functioning) is also often, but not always, reported in test manuals as well. For validity testing, authors report the results of their literature searches, or other research strategies, to justify the content items they have used. Generally, authors test students with the test they are developing and a similar, well-known test and compare results for similar subtests using a stratified sample including typically developing students, students with disabilities, or other demographic characteristics chosen by the authors. Authors also provide analyses to support the validity of the areas (constructs) they have chosen to assess, and show that progression of skills within the area does occur and is measurable for the ages for which the test is designed. Although most authors report reliability and validity testing, some authors, surprisingly, do not report either their reliability or validity testing.

Using the Language, Phonological Awareness, and Reading Test Directory

The test directory is divided into six areas: (1) *General Language Tests*, (2) *Vocabulary and Grammar Tests*, (3) *Narrative Tests*, (4) *Phonological Awareness Tests*, (5) *Reading Tests*, and (6) *Writing and Achievement Tests*, with the corresponding test reviews and shorter "At-a-Glance" summaries provided alphabetically in each section. A full list of all tests is available on the main page of the directory, which contains all of the test titles which were reviewed. For each test, the reviewer provides background information on the test including (1) sample characteristics and theoretical background, (2) focus skills assessed, (3) administration procedures, (4) scoring and interpretation procedures, (5) reliability testing, and (6) validity testing. In addition to background information from the manuals, however, the reviewers'

comments and questions, as well as comments and concerns of other reviewers (most notably the Buross reviewers) are included.

References

International Reading Association (1981). *Misuse of Grade Equivalents*. Newark, DE: Author.

To cite this document:

Hayward, D. V., Stewart, G. E., Phillips, L. M., Norris, S. P., & Lovell, M. A. (2008).
Introduction to the test reviews. *Language, Phonological Awareness, and Reading Test
Directory* (pp. 1-10). Edmonton, AB: Canadian Centre for Research on Literacy.
Retrieved [insert date] from <http://www.uofaweb.ualberta.ca/elementaryed/ccrl.cfm>.

Last updated on October 30, 2008, by M. Lovell, CCRL.