ASSESSMENT WITH THE WOODCOCK-JOHNSON III

The Woodcock-Johnson Tests of Cognitive Abilities (WJ III COG) and the Woodcock-Johnson Tests of Achievement (WJ III ACH; Woodcock, McGrew, & Mather, 2001) are revised versions of the WJ-R Tests of Cognitive Abilities and Achievement (Woodcock & Johnson, 1989). These two co-normed instruments form the Woodcock-Johnson III (WJ III), a comprehensive battery of individually administered tests designed to measure various intellectual and academic abilities. The WJ III covers a wide age range (preschool through mature adulthood). Depending on the purpose of the assessment, the WJ III COG and WJ III ACH may be used independently or in conjunction with each other.

Historical Foundation

The original Woodcock-Johnson Psycho-Educational Battery (WJ) provided the first comprehensive, co-normed battery of cognitive abilities, achievement, and interest (Woodcock & Johnson, 1977). The Tests of Cognitive Ability presented a multifactor approach to test interpretation by providing four interpretive factors: Comprehension-Knowledge (Gc), Fluid Reasoning (Gf), Short-Term Memory (Gsm), and Processing Speed (Gs). The Tests of Achievement
consisted of 10 tests organized into four curricular areas: reading, mathematics, written language, and knowledge.

The Woodcock-Johnson—Revised (WJ-R; Woodcock & Johnson, 1989) was designed to expand and increase the diagnostic capabilities of the WJ. Similar to the organization of the WJ, the tests were divided into two main batteries: the Tests of Cognitive Ability (WJ-R COG) and the Tests of Achievement (WJ-R ACH). Both the WJ-R COG and WJ-R ACH have two easel test books: the Standard Battery and the Supplemental Battery.

For the 1989 WJ-R COG, interpretation was enhanced by the measurement of seven factors that represent major components of human intellectual abilities (McGrew, 1994; Reschly, 1990; Ysseldyke, 1990). Two additional factors (Grw, a reading/writing factor, and Gq, a quantitative ability factor) were measured by the WJ-R ACH. Thus, nine Gf-Gc abilities were measured across the WJ-R COG and WJ-R ACH.

The WJ-R Tests of Achievement consisted of 14 tests organized into four curricular areas: reading, mathematics, written language, and knowledge. Several new tests were added to the reading and written language areas. To facilitate pre- and posttesting, parallel alternate forms of the Tests of Achievement, Forms A and B, were available. In addition, both the WJ-R COG and WJ-R ACH have direct Spanish-language counterparts, the Bateria-R COG and the Bateria-R ACH (Woodcock & Munoz-Sandoval, 1996a, 1996b), that contain all of the same tests and interpretive features. Although the basic features of the WJ-R have been retained in the third edition, the extensive re-norming and the new tests, clusters, and interpretive procedures improve and increase the diagnostic power. In addition, two empirically derived theories guided development of the WJ III.

**Theoretical Model**

The WJ III COG is based on the Cattell-Horn-Carroll (CHC) theory of cognitive abilities, a merging of two theories of intellectual abilities: Gf-Gc theory and Carroll's three-stratum theory. Gf-Gc theory is a model that has been developed and refined over the last 60 years. Carroll’s (1993) three-stratum theory is based upon his extensive investigation of the structure of human cognitive abilities. Interpretation of these theories is enhanced by the Cognitive Performance Model (CPM; Woodcock, 1993, 1998).

**Gf-Gc Theory**

The first theory that guided the development of the WJ III stems primarily from the factor analytic students of Raymond Cattell and John Horn. At an American Psychological Association conference, Cattell (1941) proposed that human abilities consisted of two types: Fluid intelligence (Gf) and crystal-
lized intelligence (Gc). As a result of this early conceptualization, the framework for this evolving theory is often referred to as Gf-Gc theory (Cattell, 1941; Horn, 1965, 1991; Horn & Noll, 1997). Based on the research conducted in the last 40 years, as well as the work of Cattell, Horn, and Carroll, this theory has been expanded to include 9 or 10 broad factors. McGrew (1997) provided a Gf-Gc taxonomy that integrated research from Carroll (1993, 1998) as well as from Horn and Noll (1997). This taxonomy included three additional broad abilities: quantitative knowledge (Gq), decision/reaction time or speed (CDS), and a reading/writing factor (Grw). This purpose of this framework was to provide a bridge between the theoretical and empirical research on the factors of intelligence and the interpretation of assessment batteries (McGrew, 1997).

Three-Stratum Theory

The second theory that guided test development was three-stratum theory (Carroll, 1993). Using the results from exploratory factor analysis, Carroll developed a similar theory describing the content and structure of cognitive abilities. Within this theory, Stratum III represents a general factor (g); Stratum II represents the broad abilities; and Stratum I encompasses the many narrow abilities (Carroll, 1993, 1998). In the WJ III, the General Intellectual Ability score represents Stratum III; the broad abilities of Gf-Gc theory represent Stratum II; and the tests within the battery represent Stratum I.

The combination of these two similar theories into CHC theory provides a comprehensive and empirically supported psychometric framework for understanding the structure of human cognitive abilities (McGrew & Flanagan, 1998). CHC theory is viewed as dynamic, rather than static, and in the future will most likely include more factors (Woodcock, 1990). Although the WJ III is based upon this theory, Flanagan and McGrew (1997) remind us that no current measurement instrument taps all of the broad and narrow abilities identified by Carroll (1993, 1998). Carroll (1993) also noted that although Gf-Gc theory covers the major domains of intellectual functioning, numerous details about human cognitive abilities need to be filled in through future research.

Cognitive Performance Model

The application of CHC theory, as the basis for interpreting the meaning and implications of test scores, is enhanced by a simple dynamic model called the Cognitive Performance Model (CPM; Woodcock, 1993, 1998). The CPM, shown in Figure 5.1, implies that the various Gf-Gc abilities are not autonomous, but fall into several functional categories. The level and quality of an individual’s cognitive performance result from the interaction among three
types of cognitive factors: (a) stores of acquired knowledge, (b) thinking abilities, and (c) cognitive efficiency. Noncognitive factors, described as facilitator-inhibitors, represent personality attributes such as motivation, as well as extrinsic factors, such as family support.

The acquired knowledge factor represents domains of specific information, as well as recall of procedures used to perform certain tasks. The thinking abilities factor draws upon information that cannot be processed automatically and must be processed with some level of intent. The cognitive efficiency factor includes speed of processing and short-term memory. These abilities are important prerequisites for smooth, automatic, cognitive processing. Facilitator-inhibitors represent the noncognitive factors that impact cognitive performance for better or for worse, often overriding or mediating strengths and weaknesses among cognitive abilities. Facilitator-inhibitors include internal factors (e.g., health, emotional status, or persistence), as well as external factors (e.g., distractions in the environment or type of instruction).

Experienced clinicians know that scores obtained from cognitive tests must be interpreted with caution because the observed performance may be affected by environmental and test situation variables. The CPM emphasizes the concept that both cognitive and noncognitive variables interact to pro-

FIGURE 5.1
Cognitive performance model.
duce performance. In other words, cognitive performance is rarely the result of a single influence.

**DESCRIPTION OF THE WJ III**

The WJ III COG has two separate easel books that contain the tests of the Standard and Extended batteries. Table 5.1 depicts the organizational format and the clusters and tests included in the WJ III COG. The notations indicate which tests are timed (T) and which are recorded (R). Table 5.2 provides a brief description of the WJ III COG tests grouped under the CPM model. Supplemental tests are described at the end of the section.

Like the WJ III COG, the WJ III ACH has two easel books that contain the tests of the Standard and Extended batteries. Table 5.3 depicts the organization and tests of the WJ III ACH. Table 5.4 provides a brief description of the

### TABLE 5.1

**WJ III COG Organization**

<table>
<thead>
<tr>
<th>Test</th>
<th>Standard Battery</th>
<th>Extended Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Ability:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Comprehension-Knowledge (Gc)</td>
<td>1. Verbal Comprehension (Picture Vocabulary, Synonyms, Antonyms, and Verbal Analogies)</td>
<td>11. General Information</td>
</tr>
<tr>
<td>Thinking Abilities:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Term Retrieval (Glr)</td>
<td>2. Visual-Auditory Learning</td>
<td>12. Retrieval Fluency (T)</td>
</tr>
<tr>
<td>Visual-Spatial Thinking (Gv)</td>
<td>3. Spatial Relations</td>
<td>13. Picture Recognition</td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>4. Sound Blending (R)</td>
<td>14. Auditory Attention (R)</td>
</tr>
<tr>
<td>Fluid Reasoning (Gf)</td>
<td>5. Concept Formation</td>
<td>15. Analysis-Synthesis</td>
</tr>
<tr>
<td>Cognitive Efficiency:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing Speed (Gs)</td>
<td>6. Visual Matching (T)</td>
<td>16. Decision Speed (T)</td>
</tr>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>7. Numbers Reversed (R)</td>
<td>17. Memory for Words (R)</td>
</tr>
<tr>
<td>Supplemental (Ga, Gsm, Glr, Gs, Gf)</td>
<td>8. Incomplete Words (R) (Ga)</td>
<td>18. Rapid Picture Naming (T) (Gs)</td>
</tr>
<tr>
<td></td>
<td>9. Auditory Working Memory (R) (Gsm)</td>
<td>19. Planning (Gf)</td>
</tr>
<tr>
<td></td>
<td>10. Visual-Auditory Learning—Delayed (Glr)</td>
<td>20. Pair Cancellation (T) (Gs)</td>
</tr>
</tbody>
</table>

(T) = Timed, (R) = Recorded
TABLE 5.2
Description of WJ III COG Tests

**Verbal Ability (Gc)**
Test 1: Verbal Comprehension measures word knowledge through four tasks. The first task requires naming pictured objects, the second requires providing antonyms, the third requires providing synonyms for orally presented words, and the fourth requires providing analogies.
Test 10: General Information measures aspects of acquired knowledge. The task measures the ability to identify common or typical characteristics of certain objects.

**Thinking Abilities/Processes (Glr, Gv, Ga, Gf)**

**Long-Term Retrieval (Glr)**
Test 2: Visual-Auditory Learning measures an aspect of associative and meaningful memory. The task requires pairing novel visual symbols (rebuses) with familiar words and then translating the symbols into verbal phrases and sentences (visual-auditory association task).
Test 12: Retrieval Fluency measures an aspect of ideational fluency. The task measures the ability to list orally as many items as possible in several categories (things to eat or drink, first names of people, and animals) in 1 minute.

**Visual-Spatial Thinking (Gv)**
Test 3: Spatial Relations measures an aspect of visualization. The task measures the ability to select the component parts from a series of visual shapes that are needed to form a whole shape.
Test 13: Picture Recognition measures an aspect of recognition memory. The task requires the identification of a subset of previously presented pictures within a larger set of pictures.

**Auditory Processing (Ga)**
Test 4: Sound Blending measures an aspect of phonemic processing. The task measures the ability to synthesize sounds to make a whole word (auditory blending).
Test 14: Auditory Attention measures an aspect of speech discrimination and selective attention. The task measures the ability to differentiate among similar sounding words with increasing levels of background noise.

**Fluid Reasoning (Gf)**
Test 5: Concept Formation measures an aspect of inductive reasoning. The task measures the ability to examine a set of geometric figures and identify the rules when shown instances and noninstances of the concept.
Test 15: Analysis-Synthesis primarily measures general sequential reasoning (deductive logic). The task measures the ability to analyze the presented components of an incomplete logic puzzle and then determine the missing components.

**Cognitive Efficiency**

**Short-Term Memory (Gsm)**
Test 7: Numbers Reversed measures aspects of short-term auditory memory span and working memory. The task requires rearranging a series of numbers presented orally in reverse order.
Test 17: Memory for Words measures short-term auditory memory span. The task measures the ability to repeat lists of unrelated words in the correct sequences.

**Processing Speed (Gs)**
Test 6: Visual Matching measures an aspect of visual perceptual speed. The task has an early development component that requires pointing to the two identical pictures in a row.
Table 5.2—Continued

assorted colored shapes. The next component requires locating quickly and circling two identical numbers ranging from one to three digits, in a row of six numbers.

Test 16: Decision Speed measures an aspect of conceptual reasoning speed. The task measures the ability to rapidly scan a row of pictures and then circle the two drawings that are most related.

Supplemental (Glr, Gf, Gs)

Test 8: Visual-Auditory Learning—Delayed (Glr) measures an aspect of associative and meaningful memory using previously presented symbols. The task requires the subject to recall and relearn, after 30 minutes to 8 hours or 1 to 8 days, the visual-auditory associations from the Visual-Auditory Learning test.

Test 9: Auditory Working Memory (Gsm) measures aspects of short-term auditory memory span and working memory. The task involves retaining two types of information (words and numbers) that are presented orally in a specified random order and then reordering that information sequentially.

Test 10: Incomplete Words (Ga) measures one aspect of phonemic processing. The task requires the subject to identify words with missing phonemes (auditory analysis).

Test 18: Rapid Picture Naming (Glr/Gs) measures aspects of lexical retrieval and fluency. The task requires the subject to name common objects rapidly.

Test 19: Planning (Gv) measures an aspect of spatial scanning and planning. The task requires the subject to use forward thinking by planning a tracing route that covers as many segments of a visual pattern as possible without retracing or lifting the pencil.

Test 20: Pair Cancellation (Gs) measures an aspect of sustained attention. The task measures the ability to rapidly scan and circle a repeated pattern in a row of pictures.

WJ III ACH tests that are grouped under related achievement areas. Supplemental tests are described at the end of the section. To increase the reliability of decisions, cluster scores, rather than single test scores, form the basis for interpretation of the WJ III.

Norms

To date, the WJ III has the largest standardization sample of any individually administered ability or achievement test. The total normative sample included 8818 individuals living in more than 100 geographically and economically diverse communities in the United States. The examinees were randomly selected within a stratified sampling design that controlled for 10 specific variables, including socioeconomic status. The preschool sample includes 1143 children from 2 to 5 years of age (not enrolled in kindergarten). The kindergarten to 12th grade sample is composed of 4784 students, the college/university sample is based on 1165 students, and the adult sample includes 1843 individuals.
### TABLE 5.3
Organization of WJ III ACH

<table>
<thead>
<tr>
<th>Test</th>
<th>Standard Battery</th>
<th>Extended Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading (Grw)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Fluency</td>
<td>2. Reading Fluency (T)</td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>9. Passage Comprehension</td>
<td>17. Reading Vocabulary</td>
</tr>
<tr>
<td><strong>Oral Language (Gc)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>4. Understanding Directions (R)</td>
<td>15. Oral Comprehension (R)</td>
</tr>
<tr>
<td><strong>Mathematics (Gq)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>5. Calculation</td>
<td></td>
</tr>
<tr>
<td>Math Fluency</td>
<td>6. Math Fluency (T)</td>
<td></td>
</tr>
<tr>
<td><strong>Written Language (Grw)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>7. Spelling</td>
<td>16. Editing</td>
</tr>
<tr>
<td>Writing Fluency</td>
<td>8. Writing Fluency (T)</td>
<td></td>
</tr>
<tr>
<td>Written Expression</td>
<td>11. Writing Samples</td>
<td></td>
</tr>
<tr>
<td><strong>Knowledge (Gc)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supplemental (Glr, Grw, Gq)</td>
<td>12. Story Recall—Delayed (Glr)</td>
<td>20. Spelling of Sounds (R) (Gq)</td>
</tr>
<tr>
<td>H Handwriting Legibility Scale</td>
<td></td>
<td>21. Sound Awareness (R) (Gq)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22. Punctuation &amp; Capitalization (Grw)</td>
</tr>
</tbody>
</table>

(T) = Timed, (R) = Recorded

**Brief Overview of the Psychometric Properties of the Test**

For the tests, the number of items and average item density was set so that all tests have a reliability of .80 or higher. Most cluster score reliabilities are at .90 or higher. The WJ III Technical Manual (McGrew & Woodcock, 2001) contains extensive information on the reliability and validity of the WJ III. The precision of each test and cluster score is reported in terms of the Standard Error of Measurement. The reliability of the various WJ III ability/achievement discrepancy scores is reported, along with the intra-ability discrepancy scores (intra-cognitive, intra-achievement, and intra-individual). Odd-even correlations, corrected by the Spearman-Brown formula, were used to estimate reliability for each untimed test.

The Rasch single-parameter logistic test model was used to guide test
### TABLE 5.4
**Description of the WJ III ACH Tests**

#### Reading

Test 1: Letter-Word Identification measures an aspect of reading decoding. The task requires identifying and pronouncing isolated letters and words.

Test 2: Reading Fluency measures reading speed. The task requires reading and comprehending simple sentences rapidly.

Test 9: Passage Comprehension measures an aspect of reading comprehension. The task requires reading a short passage and then supplying a key missing word.

Test 13: Word Attack measures aspects of phonological and orthographic coding. The task requires applying phonic and structural analysis skills to the pronunciation of phonically regular nonsense words.

Test 17: Reading Vocabulary measures reading vocabulary through three tasks: reading words, providing synonyms and antonyms, and completing analogies.

#### Mathematics

Test 5: Calculation measures the ability to perform mathematical computations, from simple addition facts to calculus.

Test 6: Math Fluency measures aspects of number facility and math achievement. The task requires rapid calculation of simple, single-digit addition, subtraction, and multiplication facts.

Test 10: Applied Problems measures the ability to analyze and solve practical mathematical problems.

Test 18: Quantitative Concepts measures aspects of quantitative reasoning and math knowledge. The task requires applying mathematical concepts and analyzing numerical relationships.

#### Written Language

Test 7: Spelling measures the ability to write the correct spellings of words presented orally.

Test 8: Writing Fluency measures aspects of writing ability and expressive fluency. The task requires formulating and writing simple sentences quickly.

Test 11: Writing Samples measures the ability to write sentences in response to a variety of demands that are then evaluated based on the quality of expression.

Test 16: Editing measures the ability to identify and correct mistakes in spelling, punctuation, capitalization, or word usage in short typewritten passages.

Test 20: Spelling of Sounds measures aspects of phonological and orthographic coding. The task requires spelling nonsense words that conform to conventional phonics and spelling rules.

Test 22: Punctuation and Capitalization measures the ability to apply punctuation/capitalization rules.

Handwriting (H) provides a rating of writing legibility. The samples of writing evaluated may come from the Writing Samples test or another source.

#### Oral Language

Test 3: Story Recall measures aspects of language development, listening ability, and meaningful memory. The task requires listening to passages of gradually increasing length and complexity and then recalling the story elements.

Test 4: Understanding Directions measures aspects of language development and listening ability. The task requires pointing to objects in a picture after listening to instructions that increase in linguistic complexity.

(continues)
Table 5.4—Continued

Test 12: Story Recall—Delayed measures an aspect of meaningful memory using previously presented stories. The task requires the subject to recall, after 1 to 8 hours or 1 to 8 days, the story elements presented in the Oral Recall test.

Test 14: Picture Vocabulary measures word knowledge. The task requires naming common to less familiar pictured objects.

Test 15: Oral Comprehension measures aspects of listening ability and language development. The task requires listening to short passages and then supplying the missing final word. (Taped)

Test 21: Sound Awareness measures various aspects of phonological awareness, including rhyming and phonemic manipulation tasks.

Academic Knowledge

Test 19: Academic Knowledge measures aspects of general information and acquired content or curricular knowledge in various areas of the biological and physical sciences, history, geography, government, economics, art, music, and literature.

devision, item calibration, scaling, and cluster composition. The technical criteria for item selection were stringent; all selected items had to fit the Rasch measurement model as well as other criteria, including bias and sensitivity reviews. The evidence of concurrent validity comes from studies using a broad age range of individuals. For the WJ III COG, scores were compared with performances on other intellectual measures appropriate for individuals at the ages tested, such as the Wechsler Intelligence Scale for Children—Third Edition (WISC-III), the Differential Ability Scale (DAS), the Universal Nonverbal Intelligence Test (UNIT), and the Leiter-R. The correlations between the WJ III General Intellectual Ability (GIA) score and the WISC-III Full Scale IQ range from .69 to .73. Because these correlations are based on a sample of restricted range, they most likely underestimate the relationship. For the WJ III ACH, scores were compared with other appropriate achievement measures, including the Wechsler Individual Achievement Tests, the Kaufman Tests of Educational Achievement, and the Wide Range Achievement Test—3. The magnitude of the correlations suggests that the WJ III ACH is measuring skills similar to those measured by other achievement tests.

Supported by factor-analytic research, the WJ III provides valid scores for general intellectual ability (g), seven Gf-Gc factors of cognitive ability, and several areas of academic performance. Factor loadings are all high, and the relationships between the tests and the factors are supported by good fit statistics. Further construct validity evidence is apparent by examining the intercorrelations among the tests within each battery. As is to be expected, tests assessing the same broad cognitive ability or achievement area are more highly correlated with each other than with tests that assess different abilities.
Administration

Each test requires about 5 to 10 minutes to administer. The WJ III COG Standard Battery takes about 40 minutes to administer, whereas the WJ III ACH Standard Battery takes about 1 hour. As with the WJ-R, a guiding principle of the WJ III is selective testing. Depending upon the purpose of the referral, an examiner can conduct a comprehensive or focused assessment. For example, for a student referred for a reading evaluation, an evaluator may wish to administer only the cognitive/linguistic tests related to reading performance, as well as the specific academic measures of reading performance. It would rarely be necessary to administer all of the tests of the WJ III to one person.

The batteries require a trained examiner for administration. In contrast to the administration procedures, interpretation of the WJ III requires a higher level of knowledge and skill. Because of the relative complexity in administration and interpretation of the WJ III COG, this battery requires more advanced training in clinical assessment than does the WJ III ACH.

Test Session Observation Checklist

The WJ III also includes a new procedure for gathering qualitative information that can be used to help interpret behavior during testing. A seven-category, criterion-referenced Test Session Observation Checklist is included on each Test Record that was field-tested on the normative sample from the WJ III standardization. The checklist provides a method to help categorize and describe the individual's behavior under the standardized test administration conditions. Each of the seven items uses a wide range of descriptors to help identify typical and atypical behaviors during testing. The seven areas include: the individual's levels of conversational proficiency, cooperation, activity, attention and concentration, self-confidence, care in responding, and response to difficult tasks.

Scoring

With the exception of the estimated age and grade equivalents that are provided on the Test Record, the WJ III, unlike the WJ and the WJ-R, is scored only by computer. The computer scoring requires the examiner to calculate the number correct for each subtest or test and then enter these scores into the WJ III Compuscore and Profiles Program (Schrank & Woodcock, 2001). Tables 5.5 and 5.6 illustrate sample test reports from the scoring program.¹

¹The results in these tables were generated from the pilot version of the Compuscore and Profiles program.
**TABLE 5.5**  
Case One: Jovita, WJ III COG—Score Summary

COMPUSCORE VERSION 1.1  
SCORE REPORT

Name: School Age, Jovita  
School: Little Rock  
Date of Birth: 02/25/1992  
Age: 8 years, 8 months  
Grade: 3.2  
Sex: Female  
Dates of Testing: 10/30/2000 (COG)  
11/02/2000 (ACH)  
Examiner: Nancy Mather  

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>(68% BAND)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIA (Ext)</td>
<td>2.5</td>
<td>1.3</td>
<td>4.6</td>
<td>85/90</td>
<td>35</td>
<td>94 (92-96)</td>
</tr>
<tr>
<td>Verbal Ability (Ext)</td>
<td>4.3</td>
<td>2.9</td>
<td>6.0</td>
<td>95/90</td>
<td>74</td>
<td>110 (105-114)</td>
</tr>
<tr>
<td>Thinking Ability (Ext)</td>
<td>3.2</td>
<td>1.4</td>
<td>7.7</td>
<td>90/90</td>
<td>50</td>
<td>100 (97-103)</td>
</tr>
<tr>
<td>Cognitive Efficiency (Ext)</td>
<td>1.8</td>
<td>1.2</td>
<td>2.6</td>
<td>60/90</td>
<td>13</td>
<td>83 (80-87)</td>
</tr>
<tr>
<td>Comprehension-Knowledge (Gc)</td>
<td>4.3</td>
<td>2.9</td>
<td>6.0</td>
<td>95/90</td>
<td>74</td>
<td>110 (105-114)</td>
</tr>
<tr>
<td>Long-Term Retrieval (Gh)</td>
<td>1.9</td>
<td>K.3</td>
<td>8.0</td>
<td>85/90</td>
<td>21</td>
<td>88 (83-93)</td>
</tr>
<tr>
<td>Visual-Spatial Thinking (Gv)</td>
<td>9.8</td>
<td>3.7</td>
<td>&gt;18.0</td>
<td>97/90</td>
<td>92</td>
<td>121 (116-127)</td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>1.0</td>
<td>K.1</td>
<td>3.2</td>
<td>75/90</td>
<td>12</td>
<td>83 (77-88)</td>
</tr>
<tr>
<td>Fluid Reasoning (Gf)</td>
<td>3.7</td>
<td>2.4</td>
<td>6.0</td>
<td>93/90</td>
<td>60</td>
<td>104 (99-109)</td>
</tr>
<tr>
<td>Processing Speed (Gs)</td>
<td>2.1</td>
<td>1.5</td>
<td>2.8</td>
<td>64/90</td>
<td>17</td>
<td>86 (82-89)</td>
</tr>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>1.4</td>
<td>K.7</td>
<td>2.3</td>
<td>56/90</td>
<td>16</td>
<td>85 (80-90)</td>
</tr>
<tr>
<td>Phonemic Awareness</td>
<td>K.4</td>
<td>K.0</td>
<td>2.1</td>
<td>67/90</td>
<td>7</td>
<td>78 (72-83)</td>
</tr>
<tr>
<td>Phonemic Awareness III</td>
<td>K.5</td>
<td>K.1</td>
<td>1.4</td>
<td>49/90</td>
<td>2</td>
<td>69 (65-73)</td>
</tr>
<tr>
<td>Working Memory</td>
<td>1.3</td>
<td>K.5</td>
<td>2.1</td>
<td>47/90</td>
<td>10</td>
<td>81 (77-85)</td>
</tr>
<tr>
<td>Broad Attention</td>
<td>1.8</td>
<td>K.9</td>
<td>2.9</td>
<td>71/90</td>
<td>16</td>
<td>85 (81-88)</td>
</tr>
<tr>
<td>Cognitive Fluency</td>
<td>2.1</td>
<td>1.0</td>
<td>3.6</td>
<td>80/90</td>
<td>27</td>
<td>91 (88-93)</td>
</tr>
<tr>
<td>Executive Processes</td>
<td>3.2</td>
<td>1.8</td>
<td>5.5</td>
<td>90/90</td>
<td>51</td>
<td>100 (96-105)</td>
</tr>
<tr>
<td>Knowledge</td>
<td>3.2</td>
<td>2.1</td>
<td>4.6</td>
<td>90/90</td>
<td>51</td>
<td>100 (96-105)</td>
</tr>
<tr>
<td>Oral Language (Ext)</td>
<td>3.3</td>
<td>1.7</td>
<td>5.8</td>
<td>91/90</td>
<td>53</td>
<td>101 (97-105)</td>
</tr>
<tr>
<td>Oral Expression</td>
<td>4.0</td>
<td>1.7</td>
<td>7.2</td>
<td>93/90</td>
<td>62</td>
<td>105 (99-110)</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>3.0</td>
<td>1.7</td>
<td>4.7</td>
<td>88/90</td>
<td>45</td>
<td>98 (94-103)</td>
</tr>
<tr>
<td>Total Achievement</td>
<td>1.5</td>
<td>1.1</td>
<td>1.9</td>
<td>19/90</td>
<td>2</td>
<td>70 (68-72)</td>
</tr>
<tr>
<td>Broad Reading</td>
<td>1.1</td>
<td>K.9</td>
<td>1.3</td>
<td>1/90</td>
<td>0.2</td>
<td>58 (56-60)</td>
</tr>
<tr>
<td>Broad Mathematics</td>
<td>3.1</td>
<td>2.3</td>
<td>4.0</td>
<td>88/90</td>
<td>46</td>
<td>98 (95-101)</td>
</tr>
<tr>
<td>Broad Written Language</td>
<td>1.1</td>
<td>K.7</td>
<td>1.5</td>
<td>18/90</td>
<td>1</td>
<td>64 (59-68)</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>1.1</td>
<td>K.9</td>
<td>1.3</td>
<td>0/90</td>
<td>&lt;0.1</td>
<td>54 (50-57)</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>1.2</td>
<td>K.9</td>
<td>1.5</td>
<td>7/90</td>
<td>2</td>
<td>68 (65-71)</td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>3.0</td>
<td>2.1</td>
<td>4.1</td>
<td>88/90</td>
<td>44</td>
<td>98 (94-101)</td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>3.2</td>
<td>2.6</td>
<td>4.0</td>
<td>90/90</td>
<td>51</td>
<td>101 (96-105)</td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>1.2</td>
<td>K.9</td>
<td>1.6</td>
<td>11/90</td>
<td>2</td>
<td>68 (63-72)</td>
</tr>
<tr>
<td>Written Expression</td>
<td>1.3</td>
<td>K.8</td>
<td>1.8</td>
<td>38/90</td>
<td>2</td>
<td>68 (61-74)</td>
</tr>
</tbody>
</table>
### Table 5.5 — Continued

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>SS (68% BAND)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Skills</td>
<td>—</td>
<td>1.4</td>
<td>1.1</td>
<td>1.7</td>
<td>4/90</td>
<td>0.2</td>
</tr>
<tr>
<td>Academic Fluency</td>
<td>—</td>
<td>1.6</td>
<td>1.0</td>
<td>2.2</td>
<td>43/90</td>
<td>4</td>
</tr>
<tr>
<td>Academic Applications</td>
<td>—</td>
<td>1.5</td>
<td>1.2</td>
<td>1.9</td>
<td>29/90</td>
<td>5</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>—</td>
<td>3.2</td>
<td>2.0</td>
<td>4.6</td>
<td>90/90</td>
<td>50</td>
</tr>
<tr>
<td>Phoneme/Grapheme Knowledge</td>
<td>—</td>
<td>1.3</td>
<td>1.0</td>
<td>1.6</td>
<td>14/90</td>
<td>3</td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>—</td>
<td>5.6</td>
<td>3.9</td>
<td>7.7</td>
<td>98/90</td>
<td>87</td>
</tr>
<tr>
<td>Visual-Auditory Learning</td>
<td>18-E</td>
<td>2.7</td>
<td>1.2</td>
<td>10.3</td>
<td>88/90</td>
<td>43</td>
</tr>
<tr>
<td>Spatial Relations</td>
<td>52-C</td>
<td>&gt;18.0</td>
<td>9.5</td>
<td>&gt;18.0</td>
<td>99/90</td>
<td>98</td>
</tr>
<tr>
<td>Sound Blending</td>
<td>12</td>
<td>K.4</td>
<td>&lt;K.0</td>
<td>1.5</td>
<td>58/90</td>
<td>8</td>
</tr>
<tr>
<td>Concept Formation</td>
<td>18-C</td>
<td>3.6</td>
<td>2.5</td>
<td>5.5</td>
<td>93/90</td>
<td>57</td>
</tr>
<tr>
<td>Visual Matching</td>
<td>24-2</td>
<td>1.7</td>
<td>1.2</td>
<td>2.2</td>
<td>29/90</td>
<td>6</td>
</tr>
<tr>
<td>Numbers Reversed</td>
<td>7</td>
<td>1.5</td>
<td>K.9</td>
<td>2.2</td>
<td>46/90</td>
<td>17</td>
</tr>
<tr>
<td>Incomplete Words</td>
<td>14</td>
<td>K.3</td>
<td>&lt;K.0</td>
<td>3.4</td>
<td>76/90</td>
<td>19</td>
</tr>
<tr>
<td>Auditory Work Memory</td>
<td>7</td>
<td>K.9</td>
<td>&lt;K.0</td>
<td>1.9</td>
<td>47/90</td>
<td>11</td>
</tr>
<tr>
<td>General Information</td>
<td>—</td>
<td>3.3</td>
<td>2.1</td>
<td>4.6</td>
<td>90/90</td>
<td>51</td>
</tr>
<tr>
<td>Retrieval Fluency</td>
<td>29</td>
<td>K.7</td>
<td>&lt;K.0</td>
<td>5.7</td>
<td>81/90</td>
<td>5</td>
</tr>
<tr>
<td>Picture Recognition</td>
<td>38-D</td>
<td>3.7</td>
<td>1.2</td>
<td>10.2</td>
<td>91/90</td>
<td>55</td>
</tr>
<tr>
<td>Auditory Attention</td>
<td>33</td>
<td>2.1</td>
<td>K.6</td>
<td>9.6</td>
<td>87/90</td>
<td>41</td>
</tr>
<tr>
<td>Analysis-Synthesis</td>
<td>21-D</td>
<td>3.8</td>
<td>2.3</td>
<td>6.5</td>
<td>93/90</td>
<td>59</td>
</tr>
<tr>
<td>Decision Speed</td>
<td>23</td>
<td>3.0</td>
<td>2.1</td>
<td>4.2</td>
<td>88/90</td>
<td>46</td>
</tr>
<tr>
<td>Memory for Words</td>
<td>14</td>
<td>1.3</td>
<td>K.3</td>
<td>2.6</td>
<td>65/90</td>
<td>24</td>
</tr>
<tr>
<td>Rapid Picture Naming</td>
<td>82</td>
<td>1.7</td>
<td>K.9</td>
<td>2.7</td>
<td>65/90</td>
<td>28</td>
</tr>
<tr>
<td>Planning</td>
<td>—</td>
<td>2.8</td>
<td>&lt;K.0</td>
<td>&gt;18.0</td>
<td>90/90</td>
<td>47</td>
</tr>
<tr>
<td>Pair Cancellation</td>
<td>43</td>
<td>3.0</td>
<td>1.9</td>
<td>4.2</td>
<td>88/90</td>
<td>45</td>
</tr>
</tbody>
</table>

Form A of the following achievement tests was administered:

- **Letter-Word Identification**: 17 | K.9 | K.7 | 1.1 | 0/90 | <0.1 | 52 (49–55)
- **Reading Fluency**: 7 | 1.5 | 1.2 | 1.9 | 11/90 | 8 | 79 (76–81)
- **Story Recall**: — | 3.2 | K.2 | >18.0 | 90/90 | 49 | 100 (90–110)
- **Understanding Directions**: — | 1.8 | K.7 | 3.5 | 78/90 | 26 | 90 (85–96)
- **Calculation**: 14 | 3.5 | 2.8 | 4.4 | 93/90 | 63 | 105 (98–113)
- **Math Fluency**: 22 | 1.7 | <K.0 | 3.6 | 79/90 | 9 | 80 (76–84)
- **Spelling**: 15 | K.9 | K.6 | 1.2 | 3/90 | 1 | 64 (59–69)
- **Writing Fluency**: 4 | 1.6 | 1.0 | 2.3 | 47/90 | 9 | 80 (74–85)
- **Passage Comprehension**: 10 | 1.0 | K.8 | 1.3 | 2/90 | 1 | 65 (61–69)
- **Applied Problems**: 29 | 3.1 | 2.5 | 3.9 | 89/90 | 49 | 99 (95–104)
- **Writing Samples**: 4-A | 1.1 | K.7 | 1.5 | 29/90 | <0.1 | 54 (41–66)
- **Word Attack**: 4 | 1.4 | 1.2 | 1.6 | 5/90 | 4 | 75 (69–80)
- **Picture Vocabulary**: 24 | 4.3 | 2.6 | 6.2 | 95/90 | 65 | 106 (100–111)
- **Oral Comprehension**: 18 | 3.9 | 2.7 | 5.9 | 94/90 | 63 | 105 (100–110)
- **Editing**: 2 | 1.7 | 1.3 | 2.2 | 31/90 | 11 | 81 (76–87)
- **Reading Vocabulary**: — | 1.4 | 1.1 | 1.7 | 24/90 | 7 | 78 (75–82)
- **Quantitative Concepts**: — | 3.4 | 2.6 | 4.3 | 92/90 | 56 | 102 (96–109)
- **Academic Knowledge**: — | 3.2 | 2.0 | 4.6 | 90/90 | 50 | 100 (94–106)

(continues)
Table 5.5—Continued

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>SS (68% BAND)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spelling of Sounds</td>
<td>9</td>
<td>1.1</td>
<td>K.6</td>
<td>1.4</td>
<td>32/90</td>
<td>71 (65–76)</td>
</tr>
<tr>
<td>Sound Awareness</td>
<td>12</td>
<td>K.5</td>
<td>K.1</td>
<td>1.1</td>
<td>17/90</td>
<td>67 (63–70)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISCREPANCIES</th>
<th>Actual</th>
<th>Predicted</th>
<th>Difference</th>
<th>PR</th>
<th>SD</th>
<th>Significant at + or –1.50 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-Individual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension-Knowledge (Gc)</td>
<td>110</td>
<td>86</td>
<td>+24</td>
<td>99</td>
<td>+2.40</td>
<td>Yes</td>
</tr>
<tr>
<td>Long-Term Retrieval (Glr)</td>
<td>88</td>
<td>89</td>
<td>–1</td>
<td>46</td>
<td>–0.10</td>
<td>No</td>
</tr>
<tr>
<td>Visual-Spatial Thinking (Gv)</td>
<td>121</td>
<td>93</td>
<td>+28</td>
<td>98</td>
<td>+2.00</td>
<td>Yes</td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>83</td>
<td>93</td>
<td>–10</td>
<td>23</td>
<td>–0.72</td>
<td>No</td>
</tr>
<tr>
<td>Fluid Reasoning (Gf)</td>
<td>104</td>
<td>89</td>
<td>+15</td>
<td>88</td>
<td>+1.18</td>
<td>No</td>
</tr>
<tr>
<td>Processing Speed (Gs)</td>
<td>86</td>
<td>93</td>
<td>–7</td>
<td>31</td>
<td>–0.49</td>
<td>No</td>
</tr>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>85</td>
<td>92</td>
<td>–7</td>
<td>31</td>
<td>–0.49</td>
<td>No</td>
</tr>
<tr>
<td>Phonemic Awareness</td>
<td>78</td>
<td>92</td>
<td>–14</td>
<td>14</td>
<td>–1.06</td>
<td>No</td>
</tr>
<tr>
<td>Working Memory</td>
<td>81</td>
<td>91</td>
<td>–10</td>
<td>21</td>
<td>–0.82</td>
<td>No</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>54</td>
<td>91</td>
<td>–37</td>
<td>&lt;0.1</td>
<td>–4.00</td>
<td>Yes</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>68</td>
<td>91</td>
<td>–23</td>
<td>1</td>
<td>–2.24</td>
<td>Yes</td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>98</td>
<td>91</td>
<td>+7</td>
<td>70</td>
<td>+0.53</td>
<td>No</td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>101</td>
<td>88</td>
<td>+13</td>
<td>90</td>
<td>+1.27</td>
<td>No</td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>68</td>
<td>92</td>
<td>–24</td>
<td>1</td>
<td>–2.20</td>
<td>Yes</td>
</tr>
<tr>
<td>Written Expression</td>
<td>68</td>
<td>92</td>
<td>–24</td>
<td>2</td>
<td>–2.13</td>
<td>Yes</td>
</tr>
<tr>
<td>Oral Expression</td>
<td>105</td>
<td>89</td>
<td>+16</td>
<td>90</td>
<td>+1.29</td>
<td>No</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>98</td>
<td>88</td>
<td>+10</td>
<td>79</td>
<td>+0.82</td>
<td>No</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>100</td>
<td>90</td>
<td>+10</td>
<td>82</td>
<td>+0.93</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISCREPANCIES</th>
<th>Actual</th>
<th>Predicted</th>
<th>Difference</th>
<th>PR</th>
<th>SD</th>
<th>Significant at + or –1.50 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intellectual Ability/Achievement Discrepancies*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broad Reading</td>
<td>58</td>
<td>96</td>
<td>–38</td>
<td>&lt;0.1</td>
<td>–3.37</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>54</td>
<td>97</td>
<td>–43</td>
<td>&lt;0.1</td>
<td>–3.77</td>
<td>Yes</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>68</td>
<td>96</td>
<td>–28</td>
<td>0.4</td>
<td>–2.62</td>
<td>Yes</td>
</tr>
<tr>
<td>Broad Mathematics</td>
<td>98</td>
<td>97</td>
<td>+1</td>
<td>55</td>
<td>+0.12</td>
<td>No</td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>98</td>
<td>98</td>
<td>0</td>
<td>50</td>
<td>0.00</td>
<td>No</td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>101</td>
<td>97</td>
<td>+4</td>
<td>64</td>
<td>+0.35</td>
<td>No</td>
</tr>
<tr>
<td>Broad Written Language</td>
<td>64</td>
<td>97</td>
<td>–33</td>
<td>0.1</td>
<td>–3.01</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>68</td>
<td>97</td>
<td>–29</td>
<td>1</td>
<td>–2.51</td>
<td>Yes</td>
</tr>
<tr>
<td>Written Expression</td>
<td>68</td>
<td>97</td>
<td>–29</td>
<td>1</td>
<td>–2.52</td>
<td>Yes</td>
</tr>
<tr>
<td>Oral Language (Ext)</td>
<td>101</td>
<td>96</td>
<td>+5</td>
<td>66</td>
<td>+0.41</td>
<td>No</td>
</tr>
<tr>
<td>Oral Expression</td>
<td>105</td>
<td>97</td>
<td>+8</td>
<td>74</td>
<td>+0.66</td>
<td>No</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>98</td>
<td>97</td>
<td>+1</td>
<td>55</td>
<td>+0.11</td>
<td>No</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>100</td>
<td>96</td>
<td>+4</td>
<td>63</td>
<td>+0.34</td>
<td>No</td>
</tr>
</tbody>
</table>

*These discrepancies based on GIA (Ext) with ACH Broad, Basic, and Applied clusters.
### TABLE 5.6
Case Two: Tom, WJ III COG—Score Summary

**COMPUSCORE VERSION 1.1**

**SCORE REPORT**

<table>
<thead>
<tr>
<th>Name: College, Tom</th>
<th>Organization: University of Georgia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Birth: 10/02/1979</td>
<td>Department: LDC</td>
</tr>
<tr>
<td>Age: 21 years, 1 month</td>
<td>Grade: 13.2</td>
</tr>
<tr>
<td>Sex: Male</td>
<td>Examiner: Noel Gregg</td>
</tr>
<tr>
<td>Dates of Testing: 10/25/2000 (COG)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10/30/2000 (ACH)</td>
</tr>
</tbody>
</table>

**TABLE OF SCORES: Woodcock-Johnson III Tests of Cognitive Abilities and Tests of Achievement**
Norms based on grade 13.2 (4-year college/university)

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>(68% BAND)</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIA (Ext)</td>
<td>11.9</td>
<td>7.5</td>
<td>17.3</td>
<td>81/90</td>
<td>28</td>
<td>91 (89–93)</td>
<td></td>
</tr>
<tr>
<td>Verbal Ability (Ext)</td>
<td>&gt;18.0 &gt;18.0 &gt;18.0</td>
<td>99/90</td>
<td>98</td>
<td>130 (124–135)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thinking Ability (Ext)</td>
<td>6.9</td>
<td>2.9</td>
<td>13.1</td>
<td>75/90</td>
<td>10</td>
<td>81 (78–83)</td>
<td></td>
</tr>
<tr>
<td>Cognitive Efficiency (Ext)</td>
<td>6.7</td>
<td>5.0</td>
<td>8.9</td>
<td>35/90</td>
<td>9</td>
<td>80 (77–82)</td>
<td></td>
</tr>
<tr>
<td>Comprehension-Knowledge (Gc)</td>
<td>&gt;18.0 &gt;18.0 &gt;18.0</td>
<td>99/90</td>
<td>98</td>
<td>130 (124–135)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Term Retrieval (Glr)</td>
<td>6.1</td>
<td>1.6</td>
<td>&gt;18.0</td>
<td>84/90</td>
<td>18</td>
<td>86 (82–90)</td>
<td></td>
</tr>
<tr>
<td>Visual-Spatial Thinking (Gv)</td>
<td>1.2 &lt;K.0 4.0</td>
<td>45/90</td>
<td>1</td>
<td>63 (59–67)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>11.1</td>
<td>4.5</td>
<td>14.3</td>
<td>77/90</td>
<td>16</td>
<td>85 (79–91)</td>
<td></td>
</tr>
<tr>
<td>Fluid Reasoning (Gf)</td>
<td>10.8</td>
<td>6.9</td>
<td>&gt;18.0</td>
<td>84/90</td>
<td>37</td>
<td>95 (91–99)</td>
<td></td>
</tr>
<tr>
<td>Processing Speed (Gs)</td>
<td>4.1</td>
<td>3.2</td>
<td>5.1</td>
<td>3/90</td>
<td>1</td>
<td>64 (61–66)</td>
<td></td>
</tr>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>13.0</td>
<td>10.7</td>
<td>&gt;18.0</td>
<td>90/90</td>
<td>50</td>
<td>100 (95–105)</td>
<td></td>
</tr>
<tr>
<td>Phonemic Awareness</td>
<td>7.3</td>
<td>2.5</td>
<td>13.0</td>
<td>72/90</td>
<td>17</td>
<td>86 (81–90)</td>
<td></td>
</tr>
<tr>
<td>Phonemic Awareness III</td>
<td>6.2</td>
<td>2.7</td>
<td>12.9</td>
<td>70/90</td>
<td>11</td>
<td>81 (77–86)</td>
<td></td>
</tr>
<tr>
<td>Working Memory</td>
<td>9.8</td>
<td>6.9</td>
<td>13.0</td>
<td>70/90</td>
<td>26</td>
<td>90 (87–94)</td>
<td></td>
</tr>
<tr>
<td>Broad Attention</td>
<td>9.4</td>
<td>6.3</td>
<td>13.0</td>
<td>71/90</td>
<td>10</td>
<td>81 (77–85)</td>
<td></td>
</tr>
<tr>
<td>Cognitive Fluency</td>
<td>3.9</td>
<td>2.4</td>
<td>5.9</td>
<td>16/90</td>
<td>2</td>
<td>68 (67–70)</td>
<td></td>
</tr>
<tr>
<td>Executive Processes</td>
<td>8.1</td>
<td>4.8</td>
<td>13.0</td>
<td>74/90</td>
<td>11</td>
<td>82 (79–85)</td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>&gt;18.0</td>
<td>17.5</td>
<td>&gt;18.0</td>
<td>99/90</td>
<td>94</td>
<td>124 (119–128)</td>
<td></td>
</tr>
<tr>
<td>Oral Language (Ext)</td>
<td>&gt;18.0</td>
<td>13.3</td>
<td>&gt;18.0</td>
<td>96/90</td>
<td>88</td>
<td>118 (113–123)</td>
<td></td>
</tr>
<tr>
<td>Oral Expression</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>98/90</td>
<td>96</td>
<td>125 (120–131)</td>
<td></td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>14.5</td>
<td>11.1</td>
<td>&gt;18.0</td>
<td>92/90</td>
<td>60</td>
<td>104 (98–109)</td>
<td></td>
</tr>
<tr>
<td>Total Achievement</td>
<td>10.6</td>
<td>8.1</td>
<td>13.0</td>
<td>72/90</td>
<td>20</td>
<td>87 (86–89)</td>
<td></td>
</tr>
<tr>
<td>Broad Reading</td>
<td>6.9</td>
<td>5.5</td>
<td>8.4</td>
<td>11/90</td>
<td>5</td>
<td>75 (73–77)</td>
<td></td>
</tr>
<tr>
<td>Broad Math</td>
<td>13.0</td>
<td>10.0</td>
<td>&gt;18.0</td>
<td>89/90</td>
<td>47</td>
<td>99 (97–101)</td>
<td></td>
</tr>
<tr>
<td>Broad Written Language</td>
<td>&gt;18.0</td>
<td>12.9</td>
<td>&gt;18.0</td>
<td>95/90</td>
<td>80</td>
<td>113 (107–119)</td>
<td></td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>9.0</td>
<td>6.8</td>
<td>12.8</td>
<td>72/90</td>
<td>25</td>
<td>90 (87–93)</td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>16.4</td>
<td>12.3</td>
<td>&gt;18.0</td>
<td>94/90</td>
<td>68</td>
<td>107 (102–112)</td>
<td></td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>13.4</td>
<td>10.3</td>
<td>&gt;18.0</td>
<td>90/90</td>
<td>51</td>
<td>100 (96–104)</td>
<td></td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>13.5</td>
<td>10.1</td>
<td>&gt;18.0</td>
<td>91/90</td>
<td>53</td>
<td>101 (97–105)</td>
<td></td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>13.0</td>
<td>10.6</td>
<td>16.6</td>
<td>88/90</td>
<td>46</td>
<td>99 (95–102)</td>
<td></td>
</tr>
<tr>
<td>Written Expression</td>
<td>&gt;18.0</td>
<td>12.9</td>
<td>&gt;18.0</td>
<td>96/90</td>
<td>83</td>
<td>115 (108–121)</td>
<td></td>
</tr>
</tbody>
</table>

(continues)
Table 5.6—Continued

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>SS (68% BAND)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Skills</td>
<td>17.4</td>
<td>13.0</td>
<td>&gt;18.0</td>
<td>95/90</td>
<td>73</td>
<td>109 (105-113)</td>
</tr>
<tr>
<td>Academic Fluency</td>
<td>6.4</td>
<td>5.1</td>
<td>7.9</td>
<td>9/90</td>
<td>3</td>
<td>73 (71-74)</td>
</tr>
<tr>
<td>Academic Applications</td>
<td>13.1</td>
<td>9.6</td>
<td>&gt;18.0</td>
<td>90/90</td>
<td>49</td>
<td>100 (96-103)</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>&gt;18.0</td>
<td>16.5</td>
<td>&gt;18.0</td>
<td>98/90</td>
<td>86</td>
<td>116 (111-122)</td>
</tr>
<tr>
<td>Phoneme/Grapheme Knowledge</td>
<td>3.9</td>
<td>2.5</td>
<td>6.2</td>
<td>34/90</td>
<td>2</td>
<td>70 (67-74)</td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>&gt;18.0</td>
<td>17.7</td>
<td>&gt;18.0</td>
<td>99/90</td>
<td>97</td>
<td>128 (121-135)</td>
</tr>
<tr>
<td>General Information</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>99/90</td>
<td>97</td>
<td>129 (122-136)</td>
</tr>
<tr>
<td>Retrieval Fluency</td>
<td>94</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>91/90</td>
<td>62</td>
<td>105 (98-111)</td>
</tr>
<tr>
<td>Picture Recognition</td>
<td>42-D</td>
<td>6.4</td>
<td>2.3</td>
<td>&gt;18.0</td>
<td>79/90</td>
<td>28</td>
</tr>
<tr>
<td>Auditory Attention</td>
<td>44</td>
<td>13.0</td>
<td>&gt;18.0</td>
<td>89/90</td>
<td>45</td>
<td>98 (88-108)</td>
</tr>
<tr>
<td>Analysis-Synthesis</td>
<td>28-E</td>
<td>13.4</td>
<td>8.5</td>
<td>&gt;18.0</td>
<td>90/90</td>
<td>52</td>
</tr>
<tr>
<td>Decision Speed</td>
<td>26</td>
<td>4.1</td>
<td>3.0</td>
<td>5/90</td>
<td>3</td>
<td>72 (69-76)</td>
</tr>
<tr>
<td>Memory for Words</td>
<td>20</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>&gt;18.0</td>
<td>98/90</td>
<td>74</td>
</tr>
<tr>
<td>Rapid Picture Naming</td>
<td>90</td>
<td>2.8</td>
<td>1.8</td>
<td>4/90</td>
<td>2</td>
<td>71 (69-72)</td>
</tr>
<tr>
<td>Planning</td>
<td>&gt;18.0</td>
<td>K.9</td>
<td>&gt;18.0</td>
<td>92/90</td>
<td>74</td>
<td>109 (103-116)</td>
</tr>
<tr>
<td>Pair Cancellation</td>
<td>68</td>
<td>6.9</td>
<td>5.1</td>
<td>9/90</td>
<td>11</td>
<td>82 (80-84)</td>
</tr>
</tbody>
</table>

Form A of the following achievement tests was administered:

- Letter-Word Identification
- Reading Fluency
- Story Recall
- Understanding Directions
- Calculation
- Math Fluency
- Spelling
- Writing Fluency
- Passage Comprehension
- Applied Problems
- Writing Samples
- Story Recall—Delayed
- Word Attack
- Picture Vocabulary
- Oral Comprehension
- Editing
- Reading Vocabulary
- Quantitative Concepts
- Academic Knowledge
### Table 5.6 — Continued

<table>
<thead>
<tr>
<th>CLUSTER/Test</th>
<th>RAW</th>
<th>GE</th>
<th>EASY to DIFF</th>
<th>RPI</th>
<th>PR</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spelling of Sounds</td>
<td>22</td>
<td>3.3</td>
<td>1.8</td>
<td>6.4</td>
<td>44/90</td>
<td>7</td>
</tr>
<tr>
<td>Sound Awareness</td>
<td>38</td>
<td>4.7</td>
<td>2.8</td>
<td>10.3</td>
<td>64/90</td>
<td>17</td>
</tr>
<tr>
<td>Punctuation &amp; Capitals</td>
<td>29</td>
<td>12.9</td>
<td>9.2</td>
<td>&gt;18.0</td>
<td>86/90</td>
<td>37</td>
</tr>
</tbody>
</table>

**STANDARD SCORES**

<table>
<thead>
<tr>
<th>DISCREPANCIES</th>
<th>Actual</th>
<th>Predicted</th>
<th>Difference</th>
<th>PR</th>
<th>SD</th>
<th>Significant at + or −1.50 SD (SEE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-individual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension-Knowledge (Gc)</td>
<td>130</td>
<td>96</td>
<td>+34</td>
<td>&gt;99.9</td>
<td>+4.01</td>
<td>Yes</td>
</tr>
<tr>
<td>Long-Term Retrieval (Glr)</td>
<td>86</td>
<td>100</td>
<td>−14</td>
<td>11</td>
<td>−1.23</td>
<td>No</td>
</tr>
<tr>
<td>Visual-Spatial Thinking (Gv)</td>
<td>63</td>
<td>101</td>
<td>−38</td>
<td>0.3</td>
<td>−2.80</td>
<td>Yes</td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>85</td>
<td>100</td>
<td>−15</td>
<td>13</td>
<td>−1.13</td>
<td>No</td>
</tr>
<tr>
<td>Fluid Reasoning (Gs)</td>
<td>95</td>
<td>99</td>
<td>−4</td>
<td>35</td>
<td>−0.39</td>
<td>No</td>
</tr>
<tr>
<td>Process Speed (Gs)</td>
<td>64</td>
<td>101</td>
<td>−37</td>
<td>0.3</td>
<td>−2.71</td>
<td>Yes</td>
</tr>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>100</td>
<td>99</td>
<td>+1</td>
<td>53</td>
<td>+0.09</td>
<td>No</td>
</tr>
<tr>
<td>Phonemic Awareness</td>
<td>86</td>
<td>100</td>
<td>−14</td>
<td>13</td>
<td>−1.15</td>
<td>No</td>
</tr>
<tr>
<td>Working Memory</td>
<td>90</td>
<td>99</td>
<td>−9</td>
<td>24</td>
<td>−0.72</td>
<td>No</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>90</td>
<td>99</td>
<td>−9</td>
<td>16</td>
<td>−0.99</td>
<td>No</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>107</td>
<td>98</td>
<td>+9</td>
<td>83</td>
<td>+0.97</td>
<td>No</td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>100</td>
<td>99</td>
<td>+1</td>
<td>53</td>
<td>+0.07</td>
<td>No</td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>101</td>
<td>99</td>
<td>+2</td>
<td>58</td>
<td>+0.21</td>
<td>No</td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>99</td>
<td>99</td>
<td>0</td>
<td>49</td>
<td>−0.03</td>
<td>No</td>
</tr>
<tr>
<td>Written Expression</td>
<td>115</td>
<td>98</td>
<td>+17</td>
<td>93</td>
<td>+1.49</td>
<td>No</td>
</tr>
<tr>
<td>Oral Expression</td>
<td>125</td>
<td>97</td>
<td>+28</td>
<td>99.6</td>
<td>+2.69</td>
<td>Yes</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>104</td>
<td>98</td>
<td>+6</td>
<td>68</td>
<td>+0.46</td>
<td>No</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>116</td>
<td>98</td>
<td>+18</td>
<td>97</td>
<td>+1.83</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Intellectual Ability/Achievement Discrepancies**

<table>
<thead>
<tr>
<th>DISCREPANCIES</th>
<th>Actual</th>
<th>Predicted</th>
<th>Difference</th>
<th>PR</th>
<th>SD</th>
<th>Significant at + or −1.50 SD (SEE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Reading</td>
<td>75</td>
<td>94</td>
<td>−19</td>
<td>3</td>
<td>−1.94</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>90</td>
<td>94</td>
<td>−4</td>
<td>36</td>
<td>−0.36</td>
<td>No</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>107</td>
<td>94</td>
<td>+13</td>
<td>88</td>
<td>+1.19</td>
<td>No</td>
</tr>
<tr>
<td>Broad Mathematics</td>
<td>99</td>
<td>97</td>
<td>+2</td>
<td>57</td>
<td>+0.17</td>
<td>No</td>
</tr>
<tr>
<td>Math Calculation Skills</td>
<td>100</td>
<td>98</td>
<td>+2</td>
<td>57</td>
<td>+0.18</td>
<td>No</td>
</tr>
<tr>
<td>Math Reasoning</td>
<td>101</td>
<td>96</td>
<td>+5</td>
<td>66</td>
<td>+0.42</td>
<td>No</td>
</tr>
<tr>
<td>Broad Written Language</td>
<td>113</td>
<td>94</td>
<td>+19</td>
<td>96</td>
<td>+1.79</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic Writing Skills</td>
<td>99</td>
<td>96</td>
<td>+3</td>
<td>58</td>
<td>+0.20</td>
<td>No</td>
</tr>
<tr>
<td>Written Expression</td>
<td>115</td>
<td>95</td>
<td>+20</td>
<td>97</td>
<td>+1.84</td>
<td>Yes</td>
</tr>
<tr>
<td>Oral Language (Ext)</td>
<td>118</td>
<td>93</td>
<td>+25</td>
<td>99</td>
<td>+2.54</td>
<td>Yes</td>
</tr>
<tr>
<td>Oral Expression</td>
<td>125</td>
<td>95</td>
<td>+30</td>
<td>99.5</td>
<td>+2.59</td>
<td>Yes</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>104</td>
<td>94</td>
<td>+10</td>
<td>81</td>
<td>+0.88</td>
<td>No</td>
</tr>
<tr>
<td>Academic Knowledge</td>
<td>116</td>
<td>95</td>
<td>+21</td>
<td>96</td>
<td>+1.81</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*These discrepancies based on GIA (Ext) with ACH Broad, Basic, and Applied clusters.*
Derived Scores

A variety of derived scores are reported on the scoring program's Table of Scores. In addition to age norms (age 2 to 90+), grade norms are also provided (from kindergarten through first year graduate school). The age- and grade-equivalent scores used in the WJ III do not possess the limitations associated with age- or grade-equivalents from many other tests; group-administered instruments are particularly limited, because they represent the median performance of individuals at the defined age- or grade-equivalent. Because the WJ III grade norms are continuous (rather than extrapolated), a third-grade individual who obtains a grade equivalent of 8.0 obtained the same score, using the same items, that was obtained by the median norming student in grade 8.0.

A criterion-referenced Relative Proficiency Index (RPI) is reported that describes the relative ease or difficulty with which the individual will find similar grade- (or age-) level tasks. This RPI is then represented as a developmental or instructional zone that has implications in regard to the difficulty level of school assignments. For example, a student in the first month of sixth grade has an easy level of reading at beginning first grade and a difficult level of reading at the end of second grade. Clearly, sixth-grade reading materials would be too difficult for this student to read independently.

For the percentile ranks and standard scores, examiners may select to report a 68%, 90%, or 95% confidence band around the standard score. Additional scores are provided for interpretation of the discrepancy procedures. The two most useful scores for interpreting the presence and severity of any discrepancy are the discrepancy percentile rank (DISCREPANCY PR) and the discrepancy standard deviation (DISCREPANCY SD). The DISCREPANCY PR displays the percentage of the population with a similar positive or negative discrepancy (such as 95% or 5%). The DISCREPANCY SD is a standardized z score that transforms the same discrepancy into standard deviation units, such as using a criterion of + or −1.5 standard deviations.

CLINICAL SIGNIFICANCE

Due to the recent publication of the WJ III, our clinical experiences, as well as the number of studies in the published literature, are limited. Therefore, we will discuss several of the new features of the test and explain how these features will be useful to those engaged in clinical practice. In addition, we will present case studies of two individuals who were administered the normative version of the WJ III. Because the final norm tables were not yet available when this chapter was prepared, a pilot version of the WJ III Compuscore and Profiles program was used.
Discrepancy Procedures

The WJ III provides several diagnostic procedures for helping an evaluator describe a person's abilities. The main options for analyzing within-individual variability are the intra-ability discrepancies. The WJ III has three intra-ability discrepancy procedures: (a) intra-individual, (b) intra-cognitive, and (c) intra-achievement. The purpose of the intra-ability discrepancy procedure is diagnosis, or to help determine an individual's strengths and weaknesses (Mather & Schrank, 2001). One notable feature of the WJ III is the provision of a new intra-individual discrepancy procedure. This intra-individual discrepancy procedure allows examiners to analyze an individual's cognitive and academic performance across the cluster scores of the WJ III COG and WJ III ACH.

The WJ III also provides procedures for evaluating three types of ability/achievement discrepancies. The purpose of the ability/achievement discrepancies is to predict achievement (Mather & Schrank, 2001). In addition to use of the GIA-Standard or GIA-Extended scores and the Predicted Achievement scores in the WJ III COG, the WJ III ACH contains an ability/achievement discrepancy procedure. For this procedure, the Oral Language cluster may be used as the measure of ability to predict academic achievement. The Oral Language ability/achievement procedure has particular relevance for helping clinicians distinguish between individuals with adequate oral language capabilities, but poor reading and writing abilities (i.e., specific reading disabilities), and individuals whose oral language abilities are commensurate with their reading and writing performance. In the first case, intervention would focus on reading and writing development; in the second case, intervention would be directed to all aspects of language.

Special Clinical Clusters

The WJ III contains several special clinical clusters that will help evaluators to make informed decisions regarding instructional programming and the need for accommodations. Although additional research is needed on the clinical efficacy of these clusters, a substantial body of research documents the importance of these abilities to scholastic and vocational success. The WJ III clinical clusters provide a psychometrically sound score based upon extensive exploratory and confirmatory factor analyses (McGrew & Woodcock, 2001). It is important to keep in mind the multidimensional nature of any task or factor when attempting to measure a cognitive processing area (e.g., working memory, attention). For instance, a cancellation task, like the Pair Cancellation test of the WJ III, involves sustained attention, processing speed, and motoric speed, as well as executive functioning. Careful consideration of an individual's performance across the WJ III (COG as well as ACH) and behavioral observations will provide evidence of whether good or poor
performance can be attributed to the cognitive factor under investigation or to something else.

**Broad Attention**

The WJ III Broad Attention cluster measures four different cognitive aspects of attention: The Numbers Reversed Test measures attentional capacity; The Auditory Working Memory test measures the ability to divide information in short-term memory; The Auditory Attention test measures the ability to attend to speech sounds; and the Pair Cancellation test measures the ability to sustain attention. This broad sampling of tasks may assist evaluators in the identification of individuals with attentional difficulties, particularly when accompanied with systematic behavioral observations, both within the test setting and the natural environment. Attention is a superordinate construct with multifactorial attributes. Barkley (1996) suggests that individuals with learning disabilities have trouble sustaining attention, whereas individuals with attention-deficit/hyperactivity disorder (ADHD) appear deficient in response inhibition and the capacity to delay response. Furthermore, internal factors, such as motivation, may influence attentional performance as much as the format of a task (Barkley, 1996). Therefore, practitioners are cautioned to consider multiple factors when interpreting the results of this cluster.

**Phonological Awareness and Phoneme/Grapheme Knowledge**

A powerful feature of the WJ III is inclusion of new measures of phonological awareness, as well as measures of phoneme/grapheme knowledge. Phonological awareness measures on the WJ III include a variety of tasks, such as rhyming, deletion, closure, and speech discrimination. Phoneme/grapheme knowledge is measured through the abilities to read and spell nonwords with regular English spelling patterns. A substantial body of research has documented the relationship between poor phonological awareness and delayed literacy development. Deficits in phonological skill have been identified as a major cause of severe reading problems (Ehri, 1994; Morris et al., 1998; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993). These clusters will be useful as early intervention measures to identify children at risk for reading failure, as well as to document the cause of reading failure in older students. In addition, a comparison of performance on phonological awareness tasks to performance on tasks involving knowledge of sound-symbol relationships will help evaluators determine appropriate treatment plans. For example, a child with poor phonological awareness may initially need attention directed toward the development of these oral language abilities. A child with good phonological awareness, but poor reading, will require direct instruction in phoneme/grapheme relationships.
Cognitive and Academic Fluency

The WJ III contains several measures of fluency, or speed of processing. Automaticity refers to the ability to perform tasks rapidly with ease. The ability to work easily and efficiently is an important factor for both scholastic and vocational success. The fluency tests should help evaluators document weaknesses in the speed of performance that will be particularly relevant to a need for accommodations, such as extended time. The Cognitive Fluency Cluster consists of the following tests: Retrieval Fluency, Decision Speed, and Rapid Picture Naming. The Academic Fluency Cluster consists of Reading Fluency, Math Fluency, and Writing Fluency.

Recent research has confirmed the importance of cognitive processing speed to the process of reading, particularly decoding (Denckla & Cutting, 1999; Wolf, 1999; Wolf & Segal, 1999). While debate is ongoing as to which associated cognitive abilities impact most on the automatic processing of reading and spelling words, the importance of fluency and speed to the acquisition of literacy is well supported. As Ehri (1998) noted, the key to skilled reading is processing words in text quickly without having to pay conscious attention to the words.

The relationship between the Cognitive and Academic Fluency Clusters and reading comprehension and written expression is unclear, particularly with the adolescent and adult populations. Considerable controversy exists regarding the nature of the relationship between reading rate and comprehension and production of text (Carver, 1997; Kintsch, 1998; Perfetti, 1985). Reading rate may reflect the strength of retrieval from long-term memory (Kintsch, 1998) or the efficiency in which an individual is able to code information (Perfetti, 1985). Reading rate and reading comprehension appear to be dissociable factors (Palmer, MacLeod, Hunt, & Davidson, 1985; Rupley, Willson, & Nichols, 1998). The Cognitive and Academic Fluency Clusters of the WJ III will provide a strong psychometric tool for exploring this relationship across the life span.

Working Memory

The construct of working memory and its relationship to academic performance has been a topic of recent investigation. Generally, the conceptualization of working memory is described as a "mental workspace consisting of activated memory representations that are available in a temporary buffer for manipulation during cognitive processing" (Stoltzfus, Hasher, & Zacks, 1996, p. 66). Although researchers appear to adhere to the same definition, controversy exists about the best way to measure the components. Therefore, many different types of measures have been designed to investigate the construct of working memory (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980; Wagner, Torgesen, & Rashotte, 1999).

The Working Memory Cluster is composed of two tests: Numbers Reversed
and Auditory Working Memory. This cluster will provide researchers and practitioners a psychometrically strong tool to investigate the relationships among an individual's working memory and achievement (e.g., reading decoding, reading comprehension, calculation, spelling). Clinically, it will be interesting to observe a student's performance on this cluster in comparison to the Wechsler Memory Scale—III (WMS-III; Wechsler, 1997) Working Memory Index and/or the Phonological Memory Comprehensive Composite Score on the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999). Each of the working memory composite scores across these different measures (WJ III, WMS-III, and the CTOPP) are composed of different types of tasks. Therefore, a student might show varied performance across these different composite scores of working memory due to the difference in the type of tasks. For instance, on the WJ III Auditory Working Memory test, the individual is required to recall real words, not just letters or numbers. For students with specific word recall problems, such a task might present significant problems that would not be observed on the Letter-Numbers Subtest (WMS-III) or on a numbers reversed task. On the CTOPP, the Phonological Memory Cluster contains a test that requires repetition of nonwords, as well as a digit-reversed task. The recall of nonwords could prove more difficult for students with significant phonological processing deficits as compared to tasks involving letters, numbers, or real words. Practitioners should, therefore, look at the performance of an individual across different types of working memory tests, because significant discrepancies may provide an important piece of diagnostic information.

**Executive Processing**

The Executive Processing Cluster consists of the following tests: Concept Formation, Planning, and Pair Cancellation. As with the constructs of processing speed, working memory, and attention, definition and measurement problems have surrounded the construct of executive functioning. The Executive Processing Cluster attempts to provide professionals with a broad measure of cognitive processes (planning, self-regulation, and reasoning) that have been identified in the literature as components of executive functioning. Recent research investigating the academic performance (reading, written expression, and mathematics) of individuals with learning disabilities has identified a strong relationship between difficulties with academic tasks and executive functioning (Meltzer, 1994). The Executive Processing Cluster will provide a means for researchers and practitioners to add to the body of literature investigating the relationships between executive functioning and school performance.

**Academic Skills, Fluency, and Application**

The WJ III ACH provides clusters of three different facets of academic performance: basic skills, fluency and rate, and problem solving. A comparison
among these three areas will help evaluators describe an individual's academic strengths and weaknesses, as well as make recommendations for an effective treatment plan. For example, one individual may perform poorly on tasks that are of a lower order, such as those involved in decoding, encoding, and calculating, but then do well on tasks that involve the application of information, logical thinking, and problem solving. Another individual may demonstrate the opposite pattern and may have mastered the coding aspects of the symbol system, but then struggles with higher-order tasks involving language and reasoning. Still another individual may have strengths in both the coding and symbolic aspects of language, but may comprehend and complete tasks at a very slow rate. Clearly, the implications differ.

ADVANTAGES AND DISADVANTAGES

The WJ III is based upon a comprehensive theory of intellectual functioning. The combination of Carroll's three-stratum theory with Gf-Gc theory provides a comprehensive and empirically supported framework for understanding the structure of human cognitive abilities. The instrument is a valuable tool for helping an examiner develop diagnostic and clinical hypotheses. The WJ III is unique in that it provides a wide age range, comprehensive system for measuring: (a) general intellectual ability, (b) specific cognitive abilities, (c) predicted achievement, (d) oral language, and (e) achievement. The co-normed batteries facilitate comparisons among these areas.

Discrepancy Procedures

A major advantage of this instrument is the variety of discrepancy procedures. Because all norms for the WJ III COG and the WJ III ACH are based on data from the same sample of subjects, examiners can report discrepancies between and among an individual's WJ III scores without applying a correction for the statistical phenomenon of regression to the mean. This correction procedure would require use of a regression equation or a table based upon a regression equation. The WJ III discrepancy scores do not contain error from any unknown differences that would exist when using two tests based on different norming samples.

The intra-ability discrepancies allow the examiner to analyze patterns of variability and generate hypotheses regarding an individual's learning abilities. On the intra-ability discrepancies, identification of specific cognitive and academic strengths, as well as any cognitive inefficiencies and academic weaknesses, can help the examiner formulate ideas for intervention, leading to appropriate treatment interventions. For example, test results may indicate that an individual is low in both phonological awareness and phoneme/grapheme knowledge. The examiner may hypothesize that poor phonological awareness has affected the development of word identification and spelling
skill. Therefore, the treatment recommendation would be to develop phonological awareness, coupled with direct instruction in word identification and spelling.

The intra-individual discrepancy procedure is in line with current thinking about the identification of specific learning disabilities. Identification of patterns of strengths and weaknesses among these abilities may help professionals more accurately identify a specific reading, mathematics, or writing disability. Many individuals with these types of disabilities do not demonstrate a significant discrepancy between traditional IQ and achievement measures because their various cognitive processing weaknesses are reflected in the intelligence test score. Fletcher and colleagues (1998) recommended that a more appropriate procedure than aptitude-achievement discrepancy would be to evaluate “domain-specific achievement skills and abilities correlated with those skills” (p. 186). In the intra-individual discrepancy procedure, covarying cognitive and achievement weaknesses can be viewed as observable symptoms of the same underlying disability.

Interpretation of the Predicted Achievement Clusters

Some users may not understand the purpose of the WJ III COG Predicted Achievement clusters and equate this procedure as the sole method for documenting the existence of a learning disability. On the WJ III COG, the Predicted Achievement clusters are used to predict academic performance in each curricular area using a mix of the cognitive tasks statistically associated with performance in the particular area, that is, they include the cognitive tests most relevant to the specific achievement domain.

In the field of learning disabilities, the main criterion used for identification of a learning disability is a discrepancy between aptitude (equated with potential for school success) and achievement. In other words, a specific learning disability is characterized as “unexpected” poor performance based upon a person's other capabilities. The WJ III Predicted Achievement clusters were not designed to predict future performance, to document unexpected poor performance, or to estimate one's “potential” for school success (Mather & Schrank, 2001). They merely reflect the relationship between the cognitive variables most related to an area of academic performance and the actual present level of achievement.

In some cases, students with specific learning disabilities may not exhibit Predicted Achievement/Achievement discrepancies, because a weak ability or weak abilities are included in the predictive score. For example, the tests of Sound Blending and Visual Matching are part of the Reading Aptitude cluster. A student with poor basic reading skills will often obtain low scores on measures of phonological awareness as well as on measures of processing speed. Thus, the Predicted Achievement cluster for reading will be lowered, and the individual will not evidence an Predicted Achievement/Achievement
discrepancy. The cluster predicts that the student will struggle with reading, and he or she does. This lack of discrepancy does not rule out the existence of a specific reading disability. It shows that the person's cognitive abilities related to reading, as well as his or her reading skills, are weak.

In many cases, the most sensitive procedures for substantiating the existence of specific learning disabilities on the WJ III are the intra-ability discrepancies, as well as the GIA/Achievement and Oral Language Ability/Achievement Discrepancies. The student with the profile described above would most likely evidence weaknesses in Auditory Processing and Processing Speed, as well as in Basic Reading and Writing Skills. This joint evaluation of an individual's cognitive and achievement abilities is a promising avenue for helping professionals to identify domain-specific learning disabilities.

**Interpretation of Gf-Gc Abilities**

Another advantage of the WJ III is the application of Gf-Gc theory to the interpretation of cognitive abilities. Each of the factors measures a broad intellectual ability and each test measures a narrow aspect of the broad ability. By analyzing the varied patterns of strengths and weaknesses exhibited by an individual, an evaluator can deduce the types of tasks that will be easy and the types of tasks that will be difficult for the person within an educational or vocational setting. Table 5.7 provides a few examples of possible performance implications.

**Wide Range of Measurement**

Another major advantage of the WJ III is that it is designed to provide a range of measurement from the age of 2 years through mature adulthood. Separate college norms (two-year college, four-year college, and first-year graduate levels) are provided. This comprehensive age range allows an examiner to follow an individual's performance at different periods of development.

**Complexity of the Instrument**

The WJ III is a complex, multifaceted instrument. The complexity of this instrument may be viewed as both an advantage and a disadvantage. Although computer scoring simplifies the scoring process, learning to administer and then interpret this instrument require extensive study and clinical experience. Although an examiner may use selective testing and just administer certain portions of the instrument, mastery of the entire scope of interpretive options may seem overwhelming to beginning and less experienced examiners.
TABLE 5.7
Sample Performance Implications of the Gf-Gc Broad Abilities in the WJ III COG

<table>
<thead>
<tr>
<th>Gf-Gc Ability</th>
<th>Sample Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Term Memory (Gsm)</td>
<td>Ease in remembering just-imparted instructions or information and following complex or multistep verbal directions.</td>
</tr>
<tr>
<td>Processing Speed (Gs)</td>
<td>Efficiency in executing easy cognitive tasks; amount of time needed for responding to well-practiced tasks.</td>
</tr>
<tr>
<td>Comprehension-Knowledge (Gc)</td>
<td>Ease in using and comprehending general information and vocabulary; knowledge of procedures.</td>
</tr>
<tr>
<td>Visual Processing (Gv)</td>
<td>Facility on tasks involving spatial orientation, object-space relationships, awareness of visual detail, and visual imagery.</td>
</tr>
<tr>
<td>Auditory Processing (Ga)</td>
<td>Facility on tasks involving speech discrimination, phonological awareness, and sound manipulation. Ability to apply this knowledge to the interpretation of phoneme-grapheme relationships.</td>
</tr>
<tr>
<td>Long-Term Retrieval (Glr)</td>
<td>Ease in recalling relevant information and in learning and retrieving names; amount of practice and repetition needed to learn information; ability to remember previously learned material.</td>
</tr>
<tr>
<td>Fluid Reasoning (Gf)</td>
<td>Ease in grasping abstract concepts, generalizing rules, and seeing implications; ability to change strategies if first approach does not work.</td>
</tr>
</tbody>
</table>

Clinical Clusters

Although the clinical clusters will assist skilled evaluators in diagnosis, caution must be exercised in interpreting the results. As noted, the abilities subsumed under "executive functioning" or "attention" are broad and multifaceted and do not represent factorially pure constructs. One would not want to assume that because an individual has a low score on Broad Attention that he or she has attention-deficit/hyperactivity disorder. This type of diagnosis must be coupled with qualitative information. Future validity research will help clarify the application and interpretation of the clinical clusters.

Interpretation of Cluster Scores

Cluster scores form the basis for interpretation of the WJ III. On occasion, a cluster score may mask individual differences in test performance that have clinical implications. Each test within a cluster is composed of two qualitatively different narrow abilities. In some instances, the tests within a cluster will differ significantly from each other. For example, in the WJ III ACH, the Broad Reading cluster includes one measure of word reading ability (Letter-Word Identification), one measure of reading rate (Reading Fluency), and one measure of reading comprehension (Passage Comprehension). An individual may have mastered the decoding aspects of reading, but have trouble under-
standing what he or she reads. When the test scores within a cluster differ significantly, the meaning of the cluster score must be interpreted with caution. Additional testing may be needed to document a specific difficulty.

**CASE STUDIES**

Two brief case studies are presented to illustrate how the WJ III can provide diagnostic information to substantiate an individual's strengths and weaknesses, as well as to help an evaluator derive information for intervention and accommodation purposes. For these examples, the scores are described using the WJ III verbal classification for standard scores and percentile ranks. More in-depth case reports would include additional qualitative and quantitative information. One example represents an elementary-level student and the other represents a postsecondary/college student.

**Jovita**

Jovita, an 8-year-old third grader, was referred due to difficulties with reading and written expression. Her first-grade teacher reported that Jovita had a very difficult time with phonological awareness and letter-identification activities. She was referred at the end of first grade for special education placement but did not have a large enough score discrepancy between ability and achievement measures to qualify for services. Jovita's third-grade teacher observed that she tried to read very fast, but did not recognize many words. The teacher noted that on occasion Jovita would refuse to write and put her head down on her desk during writing times. In contrast, Jovita was successful on classroom tasks involving listening and speaking. Because Jovita continued to be frustrated on all tasks involving any type of reading or writing, her third-grade teacher initiated another referral to try and discover the nature of her difficulties. The WJ III was administered as a means to identify whether or not Jovita had specific processing deficits that were affecting reading and writing performance as well as her present levels of academic functioning.

**Background Information**

Jovita lives with her mother and two older brothers. Her mother works full-time at a local convenience store. Jovita's mother had an uneventful pregnancy and delivery. All developmental milestones were met with normal limits. Jovita did experience recurring ear infections, which were initially treated with antibiotics, and then a myringotomy was performed to insert tubes. Since the myringotomy, Jovita has been free of ear infections and has been an otherwise healthy child.

Jovita's mother reported reading to her children every night. She reported
that Jovita has always enjoyed being read to since she was an infant. Because of Jovita's early reading experiences, her mother was surprised when Jovita had difficulty learning to read in school.

**Summary—Cognitive WJ III**

Jovita's verbal ability (acquired knowledge and language comprehension) and thinking ability (intentional cognitive processing) were in the average range. However, her cognitive efficiency (automatic cognitive processing) was in the low-average range. (See Table 5.5 for an overview of Jovita's scores.)

When compared with others at her grade level, Jovita's performance was superior in visual-spatial thinking; average in comprehension-knowledge and fluid reasoning; and low-average in long-term retrieval, auditory processing, processing speed, and short-term memory. Jovita's cognitive performance may be further influenced by her cognitive fluency, executive processes, and knowledge (average), her working memory capacity and attentional resources (low-average), and her phonemic awareness (low to very low).

**Summary—Achievement WJ III**

When compared to others at her grade level, Jovita's overall level of achievement was low. Her math calculation skills, math reasoning, and academic knowledge all were within the average range. Her oral expression and listening comprehension were also average. In contrast, her scores on basic reading, reading comprehension, written language, basic writing skills, and written expression were very low. Additionally, her academic skills, academic applications, and academic fluency and knowledge of phoneme-grapheme relationships were all significantly low.

Jovita attempted to pronounce words, but she appeared to rely mainly on initial and final consonant sounds. Consequently, she mispronounced many medial vowel sounds. Even though many of her responses were close approximations, her use of partial alphabetic cues resulted in a very low score in letter/word identification. Similar error patterns were apparent in spelling. When asked to write short responses and simple sentences, Jovita said that her hand was too tired from the other writing test. When asked later to write a few more answers, she replied that her hand was always tired. Therefore, her obtained score on the Writing Samples test is likely an underestimate of writing performance and more a reflection of her frustration with writing.

**Discrepancy Analysis**

When her abilities across the WJ III COG and ACH were compared, Jovita demonstrated significant strengths in comprehension-knowledge and visual-
spatial thinking. She had significant weaknesses in basic reading and writing skills, reading comprehension, and written expression. When her General Intellectual Ability—Extended (GIA-EXT) was compared to her academic performance, Jovita's performance was significantly lower than predicted in all measured areas of reading and writing.

**Recommendations**

Assessment results suggest that Jovita should be eligible for learning disability services. Treatment recommendations would focus on provision of age-appropriate oral language activities, an adaptation of reading and writing materials to her present instructional levels (approximately beginning first grade), and direct instruction in phonological awareness and phoneme/grapheme relationships. In addition, efforts must be made to address her negative feelings toward reading and writing and her occasional refusals to engage in these activities.

**Tom (Postsecondary/College)**

Tom, a 21-year-old right-handed college student was selected due to his unusual profile. He requested an evaluation to support a request for appropriate accommodations. Tom reported that it takes him an extraordinary amount of time to read and write. He reports that he sees words in mirror images and finds it easier to write backward than forward. Because a complete reversal of letters is very unusual in the adult population, the WJ III was chosen to help identify any specific cognitive processes that might be impacting his academic performance.

**Background Information**

Tom reached developmental milestones within normal age ranges. All visual and auditory acuity screening was within normal limits. He received speech therapy for articulation in second grade. He remembers his second-grade teacher rubbing his hand with sandpaper because she thought he was a smart aleck for writing backward. After his mother contacted school officials about the matter, he was referred to special education. At the beginning of the third grade, he was diagnosed with dyslexia and attended resource classes for 1 year.

Tom reported that he did not like classes in special education because they moved too slowly for him and he wanted to take more challenging courses. By high school, Tom was taking advanced placement classes. He noted that he often missed questions on exams, particularly when required to use "bubble" sheets that included "b's" and "d's." Tom is now attending a
4-year college and believes that, to be successful, he needs accommodations in the academic setting.

**Summary—Cognitive WJ III**

Tom's verbal ability (acquired knowledge and language comprehension) was in the superior range when compared to others in his grade group. His thinking ability (intentional cognitive processing) was in the low-average range. It is important to note the impact of Tom's low score on the Spatial Relations test on his Thinking Ability cluster score. This score should not be interpreted as an inability to reason, but rather as reflecting his difficulty with visual-spatial thinking. Tom's cognitive efficiency (automatic cognitive processing) was low-average. (See Table 5.6 for a summary of Tom's results.)

When compared to others in his grade group, Tom's performance was superior in comprehension-knowledge; average in fluid reasoning and short-term memory; low-average in long-term retrieval and auditory processing; and very low in visual-spatial thinking and processing speed. His knowledge was in the superior range; his working memory in the average range; his phonemic awareness, broad attention, and executive processes in the low-average range; and his cognitive fluency in the very low range.

**Summary—Achievement WJ III**

When compared to others at his grade level, Tom's overall level of achievement was in the low-average range. He demonstrated a significant strength in oral expression and oral language. His written expression and academic knowledge were within the high-average range. Tom was not penalized for “backward writing” on any of the writing tests. His reading, writing, and math performance fell within the average range. His academic skills and academic applications were both within the average range, but his academic fluency was significantly low. Tom's phoneme/grapheme knowledge was also low.

**Discrepancy Analysis**

When his abilities across the WJ III COG and ACH clusters are compared, Tom demonstrated a significant strength in comprehension-knowledge, knowledge, and oral expression. He demonstrated significant weaknesses in visual-spatial thinking and processing speed. These findings suggest that Tom has good verbal abilities, including a rich vocabulary and considerable knowledge. In contrast, he has severe weaknesses in visual-spatial thinking and speed of processing.

**Recommendations for Accommodations**

Based upon his history of difficulties, his slow processing speed and his unusual writing style, Tom was deemed eligible for an accommodation of ex-
tended time on all examinations. In addition, he received copies of all of his textbooks on tape. He was further allowed to tape-record any lectures that required extensive note-taking. Additionally, results supported the use of a proofreader, scribe, and note taker. Tom was also encouraged to seek a substitution for any foreign language requirement that would be part of his program of study. Application of these accommodations will increase the chances of Tom's success in a postsecondary setting.

CONCLUSION

Cognition, thinking, intelligence, and ability are all terms used throughout the literature to refer to an individual's abilities to process information and solve problems. The term “problem solving” provides an accurate description of the dynamic nature of cognitive and academic development (Hoy & Gregg, 1994). Cognitive and academic abilities are dynamic processes, rather than the end products of development. As Rogoff (1990) clearly stated: “The thinking organism is active in participating in an event, exploring a situation, directing attention, attempting solution” (p. 31).

The WJ III provides a unique means by which to observe the interactive and dynamic patterns produced during cognitive and academic problem solving, in tasks requiring lower-level and/or higher-level processing. New tests have been added to the WJ III that measure important clinical factors, such as working memory, planning, cognitive fluency, and attention. The procedures for investigating ability/achievement and intra-ability discrepancies have been expanded. While the WJ and the WJ-R provided a means to observe the dynamic interaction between intra-cognitive and intra-achievement abilities, the WJ III provides a new intra-individual discrepancy procedure that allows for a more in-depth examination of the multidimensional interaction among cognitive and achievement abilities, allowing evaluators to examine an individual's abilities across cognitive and achievement clusters.

The WJ III provides a sophisticated and comprehensive tool by which to observe the cognitive and academic abilities of individuals across the life span and to explore the dynamic nature of intellectual and academic development. As noted by Woodcock (1997), the primary purpose of an assessment instrument such as the WJ III is not to determine an IQ score, but to find out more about the problem. We feel certain that this instrument will prove to be a valuable diagnostic tool for many practitioners, because it will help them accomplish the primary goal of many assessments: to identify, address, and attempt to remedy the problems experienced by an individual.

References


Carver, R. P. (1997). Reading for one second, one minute, or one year from the perspective of reading theory. Scientific Studies of Reading, 1, 453–471.


INTRODUCTION

The measurement of academic achievement using standardized test instruments is a routine activity in today's schools. In fact, with renewed interest in student performance and increased emphasis on teacher and school accountability, the time spent testing school children is actually increasing. This phenomenon is true not only for large-scale, high-stakes group assessment, but also for the use of individually administered achievement instruments. For school psychologists, the passage of the Education for All Handicapped Children Act (1975) became the driving force behind their assessment practices (Fagan & Wise, 1994), and numerous achievement instruments have been published in the last two decades.

Student achievement is typically evaluated for four possible reasons using an individually administered tool:

- To determine where the student falls along a continuum of skill acquisition
- To identify those students who score at the lower and upper ends of the continuum for purposes of intervention (e.g., remediation or acceleration/enrichment)
- To determine eligibility for special programs
- To measure effectiveness of instruction or intervention
In practice, however, an individual achievement test is most often administered with an intelligence or ability test to determine if there is a significant discrepancy between the two scores. For some time, this discrepancy model has been the cornerstone for the process of determining whether a referred student has a learning disability. In recent years, there has been considerable discussion about this model; some researchers are promoting models that do not rely on an ability test score as part of the decision-making process. Others redefine a learning disability as a failure to respond to intervention/treatment regardless of ability level (Abbott, Reed, Abbott, & Berninger, 1997). However, the controversy surrounding the continued use of ability testing in the ability-achievement paradigm does not negate the use of achievement testing. In fact, most researchers as well as educators continue to rely heavily on individually administered achievement batteries to evaluate student progress and to guide instruction.

**WECHSLER INDIVIDUAL ACHIEVEMENT TEST (WIAT)**

The WIAT, introduced in 1992 by The Psychological Corporation, is a comprehensive norm-referenced battery designed to assess the achievement of academic skills in students in kindergarten through grade 12 (ages 5.0 to 19.11).

Its use is recommended (a) to evaluate individual performance in relationship to grade- or agemates, (b) to identify individual strengths and weaknesses across a broad range of academic skills, (c) to compare a student's general ability level with his or her level of achievement using two different methods of discrepancy analysis—the simple-difference method and the predicted-achievement method, and (d) to guide instructional intervention planning. The WIAT is currently in revision and differences between the two editions will be discussed later in this chapter.

**Unique Features**

Two unique features have been the hallmark of the WIAT. First, although a few other achievement tests are co-normed with ability tests (e.g., the Woodcock-Johnson Psycho-Educational Battery—Revised (WJPB-R); Woodcock & Johnson, 1989), the WIAT is the only individual achievement battery directly linked with the Wechsler scales. Measurement experts (Reynolds, 1990) stress the value of using tests normed on the same standardization sample when employing a discrepancy model to determine eligibility for a specific learning disability. Second, the WIAT is distinct for its comprehensive coverage in assessing all areas of learning disability as specified in U.S. federal law (Education for All Handicapped Children Act, 1975; Individuals with Disabilities Education Act, 1997).
Subtest Content

WIAT items encompass a wide range of curriculum objectives. Item development began with a review of educational research, a search through representative curriculum guides, and consultation with curriculum specialists; it produced the scope and sequence of curriculum objectives outlined in the manual (see p. 123, Wechsler Individual Achievement Test Manual; The Psychological Corporation, 1992). Final test items were selected after extensive field testing and item analysis using both item response theory (Lord, 1980) and conventional methods. The record forms delineate specific skills as they are evaluated. Scores from each of the eight subtests contribute to one of the four composite scores.

Reading Composite Subtests

**Basic Reading** is designed to assess the student’s ability to identify beginning or ending sounds by letter name, to match a written word with the appropriate picture, and to read words orally from a word list. Students in grade 3 or above begin with the word list. Items are scored dichotomously and the examiner is able to record the student’s behavior when presented with unfamiliar words. It is recommended that only those students with Basic Reading raw scores of 8 or higher be given Reading Comprehension.

**Reading Comprehension** is composed of short passages (carefully constructed and analyzed for appropriate grade-level readability) that the student must read and orally presented comprehension questions that must be answered verbally. Standardization procedure requires the continued presentation of the passage as the student responds to the questions. Early items offer visual cues to assist the beginning reader, but standard scores are not available for 5-year-olds or students in kindergarten.

Mathematics Composite Subtests

**Mathematics Reasoning** was developed to measure problem solving when a student is presented with word problems that require determination and application of the appropriate processes necessary for solution (e.g., addition and subtraction), comparison and ordering of fractions and decimals, interpretation of graphs and statistics, and computation of practical problems related to money, time, and measurement. Many items are presented visually, and text is provided for most items so that the student can read the question along with the examiner.

Unlike other subtests, **Numerical Operations** contains item sets, and each item must be given within the specific item set. The first set measures the ability to write dictated numerals; the following sets ask the student to
write answers to increasingly difficult computational problems that require an understanding of whole numbers, fractions, decimals, and the solution of basic algebraic equations.

**Language Composite Subtests**

Two developmental levels of listening comprehension are represented in the *Listening Comprehension* subtest: (a) nine receptive vocabulary items and (b) 26 items similar to Reading Comprehension in that the student listens to a short passage, then responds verbally to questions about it.

*Oral Expression* measures various language skills. The beginning items parallel the receptive vocabulary tasks on Listening Comprehension with expressive vocabulary items that require a verbal response rather than a pointing one. The balance of the subtest items measure the child's ability to communicate effectively by requiring him or her to provide detailed descriptions, accurate directions, and logical sequences of information in response to visual cues, with responses evaluated using a scoring rubric.

**Writing Composite Subtests**

The final two subtests, *Spelling* and *Written Expression*, contribute to the Writing Composite score. Spelling is a dictation test in which a word is pronounced, used in a sentence, then repeated for the student to write. About 10% of the words are homonyms, requiring that the word be spelled in context; the balance was selected from seven different basal spelling series.

*Written Expression* consists of a single (with an alternate) verbal prompt to which the student (only grades 3–12) must respond by composing a descriptive or narrative discourse. Scoring methods include holistic scoring and analytic scoring—the latter focusing on specific elements of writing such as ideas and development, organization, vocabulary, sentence structure, and mechanics. Because it provides a profile of strengths and weaknesses, only the analytic scoring method is used to derive standard scores.

**TECHNICAL EVALUATION**

**Standardization Sample**

The WIAT norms were derived from a standardization sample of 4252 children in 13 age groups ranging from 5.0 to 19.11 years and enrolled in grades K–12. The sample was representative of the U.S. population in 1988 based on stratification of age, grade, gender, race/ethnicity, geographic region, and parent education level. A small degree of case weighting, described on page 131 in the WIAT manual, was used to adjust the race/ethnicity proportions.
Further, the sample was drawn from both public and private school settings, and students receiving mainstream special services were not excluded. As a result, 6% of the standardization sample consisted of students with learning disabilities, speech/language disabilities, emotional disturbance, or physical impairments. In a 1999 review of four widely used achievement test batteries (Alfonso & Tarnofsky, 1999), the WIAT standardization sample was rated as *adequate* based on the number of variables representing U.S. population, and *good* based on size of normative sample and recency of data.

**Reliability**

WIAT internal consistency has been rated from adequate to very good by reviewers Nicholson (1992), Thompson (1993), and Alfonso and Tarnofsky (1999). Split-half reliability coefficients, corrected by the Spearman-Brown formula, range from .83 to .92 for all subtests across ages and from .77 to .92 for subtests across grades. Reliability coefficients for the composite scores, computed using a formula suggested by Nunnally (1978), range from .90 to .97 across ages and from .88 to .97 across grades. Test-retest stability for five grade groups (*n* = 367) with a median retest interval of 17 days produced corrected *r* coefficients ranging from .61 to .95 on subtests and from .65 to .97 on composites. The majority of the reported composite coefficients were .90 or higher. Standard error of measurement (SEM) was calculated for each standard score at each grade and age and is reported for both the fall and spring standardization samples. For composite scores the average SEMs for the fall ranged from 4.35 (Reading) to 5.19 (Language); for the spring they ranged from 3.87 (Reading) to 4.91 (Writing). Interscorer reliability is reported for Reading Comprehension and Listening Comprehension as an average correlation of .98; for Oral Expression as an average coefficient of .93; and as average coefficients of .89 for Written Expression (prompt 1) and .79 (prompt 2).

**Validity**

Evidence of three traditional types of validity—content, construct, and criterion evidence—were evaluated in order to demonstrate that the WIAT measures what it is intended to measure. Even though the content validity was assessed by a panel of national curriculum experts and deemed representative, school curricula vary. Users should review achievement test items to determine how closely the items match what is taught in their school.

Evidence of construct validity includes the expected intercorrelations among subtests reported by age, intercorrelations with the Wechsler scales (see Table D.6 of the WIAT manual), studies of group differences between the standardization sample and various clinical groups as well as differences between the various age/grade groups, and a multitrait-multimethod study
of the WIAT and other achievement tests (Roid, Twing, O'Brien, & Williams, 1992). In summary, there was a striking consistency in the correlations among scores on the reading, mathematics, and spelling subtests of the WIAT and those of the corresponding subtests on the other achievement measures.

Since the majority of school psychologists' assessment time is spent with students with learning disabilities (Smith, Clifford, Hesley, & Leifgren, 1992), WIAT scores were correlated with school grades, group-administered achievement tests, and clinical study groups. Flanagan (1997) notes that a strength of the WIAT is the demonstrated treatment validity because "data are reported that indicate that the WIAT effectively aids in diagnosis of educational/clinical concerns" (p. 84). Special study groups included children classified as gifted and children with mental retardation, emotional disturbance, learning disabilities, attention-deficit/hyperactivity disorder (ADHD), or hearing impairment. Mean composite scores ranged from 112.1 (SD = 9.9) to 117.8 (SD = 9.5) for gifted children, from 58.0 (SD = 10.2) to 66.3 (SD = 10.3) for children with mental retardation, and from 74.6 (SD = 12.0) to 92.8 (SD = 12.6) for children with learning disabilities. These results confirmed predicted expectations for achievement scores in each group.

Independent studies (Slate, 1994; Martelle & Smith, 1994; Saklofske, Schwean, & O'Donnell, 1996; Michalko & Saklofske, 1996) have provided additional evidence of WIAT validity. Saklofske, Schwean, and O'Donnell (1996) studied a sample of 21 children on Ritalin and diagnosed with ADHD and obtained subtest and composite means quite similar to those reported in the WIAT manual. Gentry, Sapp, and Daw (1995) compared subtest scores on the WIAT and the Kaufman Test of Educational Achievement (K-TEA) for 27 emotionally conflicted adolescents and found higher correlations between pairs of subtests (range of .79 to .91) than those reported in the WIAT manual. Because comparisons are often made between the WIAT and the WJPB-R, Martelle and Smith (1994) compared composite and cluster scores for the two tests in a sample of 48 students referred for evaluation of learning disabilities. WIAT composite score means were reported as 83.38 (SD = 10.31) on Reading, 89.32 (SD = 10.60) on Mathematics, 99.24 (SD = 11.84) on Language, and 80.32 on Writing. WJPB-R cluster score means were 87.67 (SD = 11.80) on Broad Reading, 92.09 (SD = 11.62) on Broad Mathematics, and 83.88 (SD = 8.24) on Broad Written Language. Although global scales of the WIAT and WJPB-R relate strongly to each other, mean WIAT composites were 3 to 6 points lower than mean scores on the WJPB-R clusters. Subtest analysis indicated some differences in the way skills are measured; for example, the two reading comprehension subtests (WIAT Reading Comprehension and WJPB-R Passage Comprehension) are essentially unrelated (r = .06). Study authors suggest that "for students with learning disabilities, the two subtests measure reading comprehension in different ways, resulting in scores that may vary greatly from test to test" (p. 7). In addition to a strong relationship
between the WIAT Mathematics Composite and the WJPB-R Broad Mathematics cluster, WIAT Numerical Operations correlated equally well with Applied Problems ($r = .63$) and with Calculation ($r = .57$) on the WJPB-R, suggesting that WIAT Numerical Operations incorporates into one subtest those skills measured by the two WJPB-R subtests. At the same time, the WJPB-R Quantitative Concepts subtest does not have a counterpart on the WIAT. The Language Composite of the WIAT, however, is a unique feature of that test.

**ADMINISTRATION AND SCORING**

WIAT is administered in both school and clinical settings. School-based examiners (generally school psychologists, educational diagnosticians, reading specialists, or special education teachers), like their more clinically-based colleagues (psychologists, neuropsychologists, and psychometricians), typically use the WIAT as part of a comprehensive psychoeducational evaluation. Examiners must have graduate-level training in the use of individually administered assessment instruments to be eligible to administer and interpret WIAT results. Administration time is approximately 30–50 minutes for children in kindergarten through grade 2 and about 55–60 minutes, excluding the time for Written Expression, for grades 3–12. It can take as long as 15 additional minutes to administer Written Expression. Although a student’s start point should be determined by current grade placement, beginning at a start point for an earlier grade is permissible if the student is low-functioning. Likewise, starting at the next-higher grade level is acceptable when a student is known to function above grade placement. All subtests except Written Expression are untimed. Testing materials include two stimulus booklets (on easel frames) which contain all administration directions, a record form, and a response booklet.

**Types of WIAT Scores**

Raw scores are computed for each subtest, but cannot be compared accurately to each other because each subtest has a different number of items and a different range of possible total scores. Comparisons across subtests or comparison of an individual student’s performance to age- or grade-peers is best made by using the derived scores (e.g., standard scores or percentile ranks) included in the test manual appendix.

**Standard Scores**

The distribution of WIAT standard scores forms a normal curve ($M = 100, SD = 15$). Approximately 68% of students in the standardization sample
scored between 85 and 115 (within 1 standard deviation of the mean), and
95% scored between 70 and 130 (within 2 standard deviations of the mean).
WIAT standard scores can also be transformed into stanines, normal curve
equivalents, and percentile ranks. The WIAT manual cautions against inap-
propriate use of grade- and age-equivalent scores as they are easily misun-
derstood by parents and teachers.

Score Limitations

WIAT scores cannot be interpreted by reporting standard scores if the stu-
dent's age or grade falls outside the WIAT age/grade range. Some examiners
have used the WIAT with older students and adults in spite of the fact that no
age or grade norms are currently available, but results from such administra-
tion must be used cautiously and out-of-range testing should be reported.
Because of ceiling effects, caution should also be used when testing gifted
adolescents.

INTERPRETATION

Qualitative Analysis

The IDEA Amendments of 1997 emphasize that a major purpose of evaluat-
ing a student is to "determine the educational needs of such child." Experi-
enced examiners have learned that test results (e.g., scores) provide only a
piece of the puzzle and that valuable information is available to the astute
examiner who knows how to supplement subtest and composite standard
scores with qualitative analysis of the student's performance while being
tested. Error analysis, behavioral observation, and testing the limits provide
abundant information about a student's strengths and weaknesses and how
to optimize the learning experience. Nonetheless, an examiner may wish to
administer criterion-referenced assessment tools such as the Stanford Diag-
nostic Reading Test, Fourth Edition (Karlsen & Gardner, 1995) and the Stanford Di-
gnostic Mathematics Test, Fourth Edition (The Psychological Corporation, 1995)
for additional diagnostic assessment of specific skills. The new Process Assess-
ment of the Learner (PAL) Test Battery for Reading and Writing (Berninger, in press),
a diagnostic reading/writing assessment, provides a focused evaluation of
the underlying processes that contribute to the acquisition of reading and
writing skills. Using the Wechsler Individual Achievement Test—Second Edi-
tion (WIAT II; The Psychological Corporation, in press) and the PAL Test Bat-
tery together yields powerful information that can lead to educational deci-
sion making based on empirical evidence. The relationship between the two
tests will be more closely examined later in this chapter.
USE WITH OTHER POPULATIONS

Research investigating the psychometric properties of the WIAT has in large part been limited to U.S. samples, although a 1996 study by Michalko and Saklofske considered its use with Canadian children. They found preliminary evidence supporting the reliability and validity of the instrument with their Saskatchewan sample (n = 90). Correlations in the expected moderate range resulted between subtests of similar content. Correlations between WIAT reading and math composites and the group-administered Canadian Achievement Test/2 (CAT/2; Canadian Test Centre, 1992)—considered to be one of the most psychometrically sound and widely used English measures of academic achievement in Canada (Bachor & Summers, 1985)—range from .65 (grade 4) to .81 (grade 6) in reading, and from .49 (grade 6) to .61 (grade 4) in math. Each of the three grade levels studied (grades 4, 5, and 6) scored significantly higher on several of the WIAT subtests and composite scores in comparison to the U.S. standardization sample. This could be the result of the small Canadian sample size or of a sample that was not representative of the Canadian student population. Study authors concluded that larger, more nationally representative studies should be conducted to determine if the development of Canadian norms is warranted.

USING THE WIAT WITH THE WECHSLER INTELLIGENCE SCALES

Although there is considerable variability among professionals regarding the procedures necessary to establish that a referred student has a learning disability, the most common practice is the inclusion of a statistically significant ability-achievement discrepancy. For this reason the WIAT was linked to the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991) through a sample of 1118 children aged 5–19; to the Wechsler Preschool and Primary Scale of Intelligence—Revised (WPPSI-R; Wechsler, 1989) through a sample of 84 5-year-olds; and to the Wechsler Adult Intelligence Scale—Revised (WAIS-R; Wechsler, 1981) with a sample of 82 adolescents aged 17–19 who were administered both tests during WIAT standardization. The demographic characteristics of the linking samples are reported in the WIAT manual (p. 130). This linking feature is especially important to the many clinical and school psychologists who continue to favor WISC-III as their ability measure of choice (Hegstrom, 1999).

Two controversial issues related to the use of the ability-achievement paradigm warrant further discussion. The first issue concerns the examiner’s decision to use either the simple-difference or the predicted-achievement method. Braden and Weiss (1988) caution that using the simple-difference
method can be problematic because the relationship between achievement and ability scores is ignored. Further, error of measurement is omitted if the simple difference is not further tested for statistical significance. They recommend that when the simple-difference method is employed the user should (a) determine if the difference is statistically significant and (b) establish how frequently the statistically significant difference occurred in the standardization sample. WIAT Table C.8 (Wechsler, 1992, p. 355) presents the subtest and composite score differences required between the WIAT and the Wechsler Full Scale IQ (FSIQ) for statistical significance. Tables C.9, C.10, and C.11 (Wechsler, p. 356) report the percentage of children in the respective standardization samples whose achievement standard score was below FSIQ by a specified amount.

To illustrate, if a 9-year-old youngster scored an FSIQ 102 on WISC-III with standard scores of 89 (simple difference of 13) on the WIAT reading composite and 73 (simple difference of 29) on the math composite, both scores would meet the required difference for significance, which would be 8.82 \( (p = .05) \) for reading and 9.75 \( (p = .05) \) for math. Further, while 10% of the standardization sample had a similar difference in reading, less than 1% had such a large difference in math.

However, Thorndike (1963) argued that underachievement should be defined as the discrepancy of actual achievement from the predicted value based on a regression equation between aptitude and achievement. WIAT Table C.4 (Wechsler, 1992, p. 352) provides differences required for statistical significance using the predicted-achievement method and Tables C.5 and C.6 (Wechsler, p. 353) chart the differences obtained by children in the WISC-III and WPPSI-R standardization linking samples.

In comparison, the same 9-year-old student with an FSIQ of 102 has a predicted reading composite score of 101 (predicted difference is 12) and a predicted math composite score of 101 (predicted difference is 28). The difference required for statistical significance at \( p = .05 \) is 17.04 for reading and 12.46 for math. In the WISC-III linking study, 10–15% of students had a similar difference in reading, and again, less than 1% had such a significant score in math.

The second controversial issue concerns which Wechsler IQ score should be employed to calculate the differences. The WIAT manual reports predicted subtest and composite standard scores only from the Full Scale IQ. Flanagan and Alfonso (1993a, 1993b) provide additional tables reporting the differences using Wechsler Verbal IQ (VIQ) or Performance IQ (PIQ), which they computed using the predicted-achievement formula in the WIAT manual (see pp. 188–189 of the manual). They employed the predicted-achievement method because it takes into account regression to the mean and measurement error, but note that the FSIQ may be misleading. In cases where VIQ and PIQ scores differ significantly from each other, using one of those scores may
represent a more accurate estimate of ability than using the Full Scale IQ score.

Conversely, Glutting, Youngstrom, Ward, Ward, and Hale (1997) proposed that using FSIQ is preferable to using WISC-III VIQ, PIQ, or Index scores. They concluded that FSIQ is the "most parsimonious and powerful predictor of academic achievement obtainable from the WISC-III," and that using VIQ, PIQ, or Index scores as predictors leads to more "laborious calculations with meager dividends" (p. 300).

Another important issue surrounding the use of WISC-III and WIAT in the ability-achievement paradigm has been to evaluate the fairness of using WISC-III IQ scores to predict WIAT scores across groups. In a 1995 study, Weiss and Prifitera analyzed 1000 cases of students aged 6–16 (Mdn = 10.5) from the WIAT standardization data that were considered representative of the 1988 U.S. population. They examined differential prediction of the four WIAT composite standard scores across ethnic groups and genders using regression equations developed with the WISC-III Full Scale IQ. Differential prediction was observed in 4 of the 12 comparisons, but in each case the magnitude of the effect size was small. Results suggested that there is no compelling reason to not use the WISC-III/WIAT regression equations. In fact, the study provides considerable evidence in support of the hypothesis that WISC-III FSIQ adequately predicts achievement scores on the WIAT across racial/ethnic groups and genders using a common regression equation calculated from the total sample.

Although some examiners assume that the same process can be used by plugging in the numbers from a non-Wechsler ability test, the WIAT manual clearly warns that "alternative ability measures should be used with extreme caution" when applying the formula (p. 189).

**REVIEWS OF WIAT**

The WIAT is an achievement battery that has been widely used by both school psychologists and special educators. Reviewers have noted both positive and negative attributes of the instrument.

Cohen (1993), Nicholson (1992), Thompson (1993), and Treloar (1994) consider the following to be advantages of the WIAT:

1. The overall standardization sample, reliability, and validity
2. Ease of administration
3. The link between the WIAT and the Wechsler IQ scales
4. The ability-achievement discrepancy analysis with both simple-difference and predicated-difference data
5. Coverage of all areas mandated by federal legislation to assess for learning disabilities
6. The close association to educational curricula

At the same time, Thompson (1993) voiced concern related to complexity of task and/or directions when testing very young or low-functioning children. Sharp (1992), Riccio (1992), Woitaszewski and Gridley (1995) and Alfonso and Tarnofsky (1999) noted problems with inadequate floor items (or, the number of appropriate items for kindergarten or older low-functioning students). Woitaszewski and Gridley, along with Hishinuma and Tadaki (1997), noted lower reliability with very young children.

These issues have been addressed in the WIAT manual with a precaution about testing very young children. The new WIAT II should alleviate these concerns with the downward extension of the age range, the inclusion of a large number of at-risk 4-year-olds in the standardization sample, the addition of several new pre-reading, pre-writing and early-math concept tasks, and the inclusion of age-appropriate directions.

Reviewers have also noted problems related to the scoring procedures for Oral Expression and Written Expression. Specifically, the scoring is too subjective and verbatim responses are difficult to record (Thompson, 1993), scores for low-scoring students appear to be inflated (Sharp, 1992; Woitaszewski & Gridley, 1995), and there is an overemphasis on spelling on the Written Expression composite score (Sharp). On the WIAT II, the scoring rubrics for both Oral Language and Written Expression were reworked in consultation with leading researchers in the field of writing instruction and assessment, and spelling carries less weight when calculating the Written Language composite score.

Sharp (1992) and Woitaszewski and Gridley (1995) have observed depressed reliability coefficients for Listening Comprehension, Oral Expression, and Written Expression. When revising these subtests for the WIAT II, additional items with higher reliabilities were included and problematic items were either dropped or reworked.

Riccio (1992) noted a lack of evidence of construct validity for Oral Expression with no correlational data presented with other measures of language. He also noted that Written Expression correlational studies were limited only to other spelling tests rather than writing sample measures. WIAT II standardization added validity studies with the Peabody Picture Vocabulary Test—III (PPVT-III; Dunn & Dunn, 1997) and the Oral and Written Language Scales (OWLS; Carrow-Woolfolk, 1995). Riccio also mentions the problem of limited content of subtests. Specifically, Written Expression contained only a single writing prompt and Listening Comprehension had items that were limited because they did not require an oral response. WIAT II focuses on the writing process rather than just the product by assessing writing skills at the
word, sentence, and discourse levels. A new scoring rubric is in place for both Written Expression and Oral Expression tasks. Listening Comprehension as well as Oral Expression items require the student to respond in both verbal and non-verbal ways. One Oral Expression task, for example, asks the student to generate a story based on sequential pictorial cues.

Finally, Cohen (1993) believes there is a lack of evidence that the WIAT is useful in program planning, program evaluation, or placement decision making. WIAT II development required input from curriculum specialists, classroom teachers, and researchers in both the U.S. and Canada to ensure appropriate item selection. The new WIAT II scoring software helps the classroom teacher use the qualitative and quantitative information from the test to develop a focused intervention plan. One of the primary goals of the WIAT revision was to more closely link assessment to intervention. This is evident with the inclusion of an entire chapter in the WIAT II manual dedicated to intervention.

**WIAT II: A STRONGER LINK BETWEEN ASSESSMENT AND INTERVENTION**

Although the WIAT II retains the basic features of the original, it incorporates several changes and additions that enhance the diagnostic utility of the instrument, addresses the criticisms leveled at the WIAT, and extends its use to new populations. In addition, a primary goal of the test development team was to form a closer link between the WIAT II and intervention planning and implementation. Discussion of the most noteworthy differences between the two versions of the instrument follows.

**An Expanded Age Range**

The age range of WIAT II extends down to 4 years and stretches to include adults enrolled in higher education (e.g., 2- or 4-year degree-seeking programs) as well as non-student adults. Student norms continue to be reported by age and grade, whereas adult norms are reported by age bands similar to those in the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III; Wechsler, 1997). This age extension has necessitated the addition of more floor and ceiling items (the former of which also addresses the need for easier items for low-functioning students). As a result, there are more items that can be categorized as measuring pre-reading or beginning math concepts. A greater number of difficult items lets administrators adequately assess college students and adults with learning disabilities for the purposes of remediation, modification, and accommodation.
Testing All Levels of Language Skills in Reading

Three reading subtests contribute to the Reading Composite score on the new instrument. They include Word Reading (Basic Reading renamed), Reading Comprehension, and a new subtest—Pseudoword Decoding. Based on the research of Berninger (1998), the WIAT II measures reading skills at all levels of language—namely, the subword, word, and text levels—for both diagnostic and intervention-planning purposes.

Subword Level

Because a measure of phonological awareness is the best single predictor of reading achievement in young children (Stanovich, Cunningham, & Cramer, 1984; Berninger, 1998), Word Reading includes pre-reading items such as letter naming and requires both application of the alphabet principle with sound-symbol relationships and identification of beginning or ending sounds and of rhyming words. Pseudoword Decoding measures the student's ability to identify the appropriate sounds of letters or groups of letters and blend them into a new word by applying acquired phonetic and structural analysis skills. Nonsense words are best suited for this task so that it mimics the act of reading a new, unfamiliar word and taps word-attack skills instead of stored reading vocabulary. It was therefore essential that WIAT II pseudowords be linguistically correct and phonemically representative of the developmental sequence of grapheme-phoneme relationships. According to Joshi (1995), many tests utilizing nonsense words have been developed without knowledge of these requirements.

Word Level

School psychologists and diagnosticians may relate to the experience of being confronted by a classroom teacher who disagrees with the results of a word-reading test with the comment, “That is not how the student reads in my class!” This discrepancy is often based on the teacher's expectation of automaticity (the ability to recognize and say a word quickly). Most word-reading tests do not penalize the student who laboriously decodes a new word. By introducing a measure of automaticity to Word Reading, the WIAT II user is able to compare reading accuracy to reading automaticity, providing a more functional measure of non-contextual word reading.

Text Level

The end result of reading, however, is not the pronunciation of familiar or new words, but the application of that skill to comprehension of text. One of the major goals in WIAT II Reading Comprehension was to develop a subtest that is more representative of reading as it occurs in the classroom or in the “real world.” To this end, passages were lengthened and varied to include narrative, descriptive, and functional messages. Colorful illustrations and more
FIGURE 6.1
WIAT II Reading subtests measure at the subword, word, and text levels.
Copyright 1999 by The Psychological Corporation. All rights reserved.
Adapted with permission of the author.

culturally diverse passages were designed to engage readers across a wide age span. Examiners have voiced concern about the practice of allowing the student to continue to view the WIAT passage while answering comprehension questions. The WIAT II addresses this concern by providing items that tap more higher-order comprehension skills such as identifying or summarizing the main idea, drawing conclusions, making inferences or generalizations, and defining unfamiliar words by using context cues. Two additional goals of WIAT II development were to create a measure of reading rate and a means to assess word reading in context.

Figure 6.1 shows how WIAT II reading subtests measure all levels of language.
Evaluating All Levels of Written Language

The WIAT II Written Language subtests measure writing at the subword, word, and text level in order to have a better understanding of the writing process as it relates to the writing product.

Subword and Word Levels

Spelling performance is a good indicator of phonological skills, and spelling acquisition, like grapheme-phoneme conversion skill, progresses through distinct stages. The new Spelling subtest is familiar, as it retains the dictation format, but additional error analysis based on developmental trends and word morphology is introduced. New spelling items include letter production and more high-frequency words. Some homonyms have been replaced by more discriminating, frequently-used homonyms (e.g., they're, their, and there).

Timed Alphabet Writing is a new task based on current research (Berninger et al., 1992; Berninger, Cartwright, Yates, Swanson, & Abbott, 1994) indicating that this activity is a strong predictor of reading, spelling, and writing skill acquisition for primary-grade students. Word fluency on the new Written Expression subtest is tapped by asking the student to write the name of as many things as he/she can in a specific category within a given time frame. Written Expression word fluency can also be compared to Oral Expression word fluency.

Text Level

Rather than offering a single writing prompt, the WIAT II provides multiple opportunities for a student to demonstrate writing skills at the sentence as well as the discourse level. Further, the new writing prompts, which replace the current descriptive/narrative essay requirement, ask the student to produce a persuasive essay stating and supporting a position. Because persuasive writing is more demanding than descriptive/narrative writing, the examiner can more readily identify problems with higher-order writing skills such as organizing ideas, presenting both sides of an argument, and defending a position. Essays are evaluated using a new rubric that allows both holistic and analytical scoring.

Figure 6.2 shows how the WIAT II writing subtests measure all levels of language.

A New Oral Expression Subtest

Significant revisions are also evident in the new Oral Language subtests that include Listening Comprehension and Oral Expression. These subtests are
FIGURE 6.2
WIAT II Written Language subtests measure at the subword, word, and text levels. Copyright 1999 by The Psychological Corporation. All rights reserved. Adapted with permission of the author.

intended to help the examiner identify students who would benefit from a more comprehensive language evaluation by assessing samples of expressive and receptive language skills similar to those required in the classroom. The student is provided multiple opportunities to demonstrate language facility at the word and text levels when he/she is asked to generate words or discourse in response to visual cues, match verbal and visual information, repeat sentences, give detailed directions, and explain the steps in a familiar process.

Math Subtest Changes
The major changes in Math Reasoning on the WIAT II are the addition of floor and ceiling items; an increase in the number of more demanding, multi-step problems; and more items related to time, money, measurement, and the interpretation of graphically represented data. Numerical Operations
begins with number identification items and ends with algebraic equations and is presented in a more traditional format rather than in item sets.

Table 6.1 summarizes the changes from the WIAT to the WIAT II subtest by subtest.

<table>
<thead>
<tr>
<th>WIAT Subtest</th>
<th>Measures</th>
<th>WIAT II Subtest</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Reading</td>
<td>• Accuracy of word recognition</td>
<td>Word Reading</td>
<td>• Letter identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Phonological awareness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Alphabet principle (letter-sound awareness)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Accuracy of word recognition</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Automaticity of word recognition</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>• Literal comprehension</td>
<td>Pseudoword Decoding</td>
<td>• Phonological decoding</td>
</tr>
<tr>
<td></td>
<td>• Inferential comprehension</td>
<td></td>
<td>• Accuracy of word attack</td>
</tr>
<tr>
<td>Spelling</td>
<td>• Alphabet principle (sound-letter awareness)</td>
<td>Spelling</td>
<td>• Literal comprehension</td>
</tr>
<tr>
<td></td>
<td>• Written spelling of regular and irregular words</td>
<td></td>
<td>• Inferential comprehension</td>
</tr>
<tr>
<td></td>
<td>• Written spelling of homonyms (integration of spelling and lexical comprehension)</td>
<td></td>
<td>• Lexical comprehension</td>
</tr>
<tr>
<td></td>
<td>• Written responses to verbal and visual cues</td>
<td></td>
<td>• Reading rate</td>
</tr>
<tr>
<td>Written Expression</td>
<td>• Descriptive writing (evaluated on extension and elaboration, grammar and usage, ideas and development, organization, unity and coherence, and</td>
<td>Written Expression</td>
<td>• Timed alphabet writing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Word fluency (written)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Sentence combining</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Sentence generation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Written responses to verbal and visual cues</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Descriptive writing (evaluated on organization, vocabulary, and mechanics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Persuasive writing (evaluated on organization, vocabulary, theme)</td>
</tr>
</tbody>
</table>
### Table 6.1 — Continued

<table>
<thead>
<tr>
<th>Sentence structure and variety</th>
<th>Numerical Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Narrative writing (evaluated on the same criteria as descriptive)</td>
<td>• Numerical writing</td>
</tr>
<tr>
<td>• Calculation (addition, subtraction, multiplication, and division)</td>
<td>• Calculation (addition, subtraction, multiplication, and division)</td>
</tr>
<tr>
<td>• Fractions, decimals, and algebra</td>
<td>• Fractions, decimals, and algebra</td>
</tr>
<tr>
<td>• Problem solving</td>
<td>• Problem solving</td>
</tr>
<tr>
<td>• Money, time, and measurement</td>
<td>• Money, time, and measurement</td>
</tr>
<tr>
<td>• Geometry</td>
<td>• Geometry</td>
</tr>
<tr>
<td>• Reading and interpreting charts and graphs</td>
<td>• Reading and interpreting charts and graphs</td>
</tr>
<tr>
<td>• Statistics</td>
<td>• Statistics</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mathematics Reasoning</th>
<th>Listening Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Quantitative concepts</td>
<td>• Receptive vocabulary</td>
</tr>
<tr>
<td>• Problem solving</td>
<td>• Expressive vocabulary</td>
</tr>
<tr>
<td>• Money, time, and measurement</td>
<td>• Listening—inferential comprehension</td>
</tr>
<tr>
<td>• Geometry</td>
<td>• Auditory short-term recall for contextual information</td>
</tr>
<tr>
<td>• Reading and interpreting charts and graphs</td>
<td>• Story generation</td>
</tr>
<tr>
<td>• Statistics</td>
<td>• Giving directions</td>
</tr>
<tr>
<td></td>
<td>• Explaining steps in sequential tasks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oral Expression</th>
<th>Oral Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Expressive vocabulary</td>
<td>• Word fluency (oral)</td>
</tr>
<tr>
<td>• Giving directions</td>
<td>• Auditory short-term recall for contextual information</td>
</tr>
<tr>
<td>• Explaining steps in sequential tasks</td>
<td>• Story generation</td>
</tr>
<tr>
<td></td>
<td>• Giving directions</td>
</tr>
<tr>
<td></td>
<td>• Explaining steps in sequential tasks</td>
</tr>
</tbody>
</table>

### Canadian Standardization

The WIAT II has simultaneously gone through the entire development process in the U.S. and Canada. For the Canadian standardization, expert curriculum reviews were conducted by representatives of each province, *p*-values for
each item were compared to those achieved by U.S. examinees, and about a
dozen unique Canadian items (e.g., money and measurement) were devel-
oped. Analysis of standardization data will determine if differences between
the two samples necessitate separate norms. WIAT and WIAT II equating
studies are also being conducted in Australia, where items have undergone a
similar review.

**A WIAT–WIAT II CASE STUDY**

One way to illustrate the changes from the WIAT to the WIAT II is to inves-
tigate how one student performs on each. Abby is an African-American, 11-
year-old (11.9), sixth-grade student at a small rural school in the southern
United States. Both of her parents are high school graduates, and her teacher
is concerned about her inconsistent classroom performance, especially in
reading. Her IQ scores, as measured on the WISC-III, are at the lower end of
the average range. Abby’s reading standard scores on the WIAT (given 2 weeks
after the WISC-III) are provided below along with the predicted scores in
parentheses, which are based on her WISC-III performance as reported in
Table C.1 in the WIAT manual (p. 346):

- Basic Reading 90 (87)
- Reading Composite 89 (85)
- Reading Comprehension 93 (85)

Abby’s reading achievement scores are slightly higher than predicted and
within the same range as her ability scores. Her teacher’s concerns do not ap-
pear warranted based on these scores, and it is unlikely that Abby would
qualify for special education services as a student with a reading disability.

Since WIAT II norms are still under development, comparisons can only
be made at the item level. WIAT II was administered between the WISC-III
and the WIAT. Abby makes similar errors on the WIAT Basic Reading and the
WIAT II Word Reading, but overall, her responses to the two word lists are only
slightly below the mean for her grade/age. On the WIAT II, Abby demonstrates
good automaticity, with no self-corrections, but tends to read unfamiliar
words based on beginning letter(s) and word configuration on both lists. She
does not attend to visual detail nor take time to sound out the words she
does not immediately recognize, thereby making an error on both subtests
when she misreads “phonograph” as “photograph.” On WIAT II Pseudoword
Decoding, she makes numerous errors on both vowel sounds and consonant
blends. When presented with an unfamiliar nonsense word she quickly reads
a “real” word that resembles the configuration of the unknown word and
matches its beginning and/or ending sound. As a result, a word she should
decode easily like “dreep” is read as “drip,” and “clurt” becomes “shirt.”

Abby’s performance on the more demanding WIAT II Reading Comprehen-
sion tasks reveal several significant problem areas even though she reads the passages quickly. When asked to read a sentence or short passage aloud, she makes several word-reading errors and fails to use context cues to self-correct. As a result, Abby misses many comprehension questions dealing with specific details, main idea, drawing conclusions, or making predictions. Her ability to read grade-level text is suspect, as she relies heavily on pictorial clues to respond. Since she subvocalizes as she reads, her errors are also detected and reported by the examiner. This performance matches Abby’s teacher’s report of her inconsistent ability to understand what she has just read in spite of the fact that she appears to be actively engaged in the reading process.

Abby’s performance on the WIAT II reading tasks may not be significantly discrepant from her IQ-based predicted performance, but considerable information has been gleaned about her reading strengths (e.g., reliance on visual cues) and weaknesses (e.g., poor attention to visual detail, poor phonological coding, and underdeveloped comprehension strategies), and she could definitely benefit from intervention in the classroom. Certain recommendations can assist the teacher in meeting Abby’s educational needs. These include spending a few minutes daily improving her phonemic awareness (Pseudoword Decoding provides a list of unmastered letter-sounds), giving her opportunities to read aloud and discuss comprehension questions with a partner, and teaching her specific reading strategies such as using context cues to improve understanding. In addition, assessment with the PAL Test Battery for Reading and Writing can identify which specific underlying processes are deficient, and results from this instrument can direct remediation efforts.

THE WIAT II AND THE PAL TEST BATTERY FOR READING AND WRITING

A single test cannot measure all things for all people. The primary focus of the WIAT II will continue to be related to eligibility issues and the identification of student strengths and weaknesses; however, its contribution to intervention planning and the evaluation of program effectiveness should not be overlooked. This role has been significantly enhanced by linking WIAT II to the PAL Test Battery for Reading and Writing (Berninger, in press). For students who demonstrate deficits or unexplained performance on any of the WIAT II reading or writing subtests, the PAL Test Battery is the diagnostic tool that helps the examiner establish the developmental level of the underlying processes that are necessary to acquire reading and writing skills. In addition, using the WIAT II with the PAL Test Battery for Reading and Writing yields an enhanced means of establishing learning disability eligibility—namely, the identification of specific processing deficits.
For example, when a student's WIAT II Reading Composite score is significantly lower than predicted based on his/her IQ (or is statistically significantly below the IQ score), the next step would be to look at performance on the Word Reading, Pseudoword Decoding, and Reading Comprehension subtests to pinpoint problem areas. By following up with the PAL Test Battery for Reading and Writing, the exact process (e.g., phonemic awareness, phoneme discrimination, orthographic coding, and rapid automatized naming) is eval-
uated and the student's performance can be placed on the developmental continuum.

Figure 6.3 describes the PAL Test Battery for Reading and Writing subtests that measure processes at the subword, word, and text levels.

Once this diagnostic information is obtained, specific intervention planning can occur. The Individual Education Plan (IEP) team can move beyond the referral question, "Does this student qualify for services because of a specific learning disability?" to "What specifically needs to occur in this student's instructional program so that identified skill gaps can be closed?" With this approach, assessment drives intervention; likewise, intervention directs ongoing assessment because a target instructional plan has been outlined (a true IEP), along with a timeline and course of action to evaluate intervention effectiveness and response to treatment of the student.

Even though we have been measuring student achievement for close to a hundred years, the primary goal of the next generation of test instruments, like the WIAT II, should be to evaluate student performance in a fair manner and to be able to produce both quantitative and qualitative test results to guide instruction, thereby enhancing student success. Only by shifting emphasis from "the numbers" (e.g., the student's test scores) to the student's specific instructional needs will the value of achievement assessment be fully appreciated.

References


6. Wechsler Individual Achievement Test

In this chapter we describe an "assessment for intervention" model and its scientific support. The model is designed to translate research on prevention and treatment of reading and writing disabilities into practice. In the first tier (also called tier 1) of assessment for intervention—screening for early intervention—all K–2 students in a school are screened to identify those who are at-risk for reading and writing problems. The screening measures are brief, but research-based. Children who are identified as being at-risk receive early intervention. Intervention should be science-based, which means that a theory-driven experiment in which competing hypotheses are tested provides empirical evidence that an intervention is effective in improving student learning outcome.

In the second tier (also called tier 2) of assessment for intervention—modifying the regular instructional program and monitoring the progress of students—the classroom program is modified for students who fail to respond or who respond very slowly to tier 1 early intervention. The modification may consist of adding necessary curriculum components, changing pedagogical practices,
revising instructional materials, or providing additional practice of skills. The goal of tier 2 is to assess whether all the essential curriculum components are in place, add them if needed, and modify how they are delivered if necessary. Curriculum-based measurement can be used to monitor progress. A multidisciplinary collaborative team uses a problem-solving approach in which further curriculum modifications are made and monitored if the student does not make progress after the initial modification.

In the third tier (also called tier 3) of assessment for intervention — diagnosis and treatment of referred children — students who have failed to respond or who responded very slowly to early intervention (first tier) and/or curriculum modification, as indicated by progress monitoring (second tier), are given a thorough assessment. The purpose of this assessment is not merely to make a placement decision about whether the student qualifies for special education services. Rather, the assessment should serve two additional purposes: (a) to diagnose, based on current scientific knowledge, why the student is having difficulty learning to read or write and (b) to design and implement a systematic and coordinated treatment plan and evaluate student response to it. Tier 3 intervention is usually more intensive than Tier 1 or Tier 2 intervention.

There are three noteworthy features of the assessment for intervention model. First, systematic application of the three-tier model can reduce the amount of time school psychologists spend doing comprehensive evaluations for special education placement. Because at-risk children are identified and given intervention early in their schooling, many severe problems are prevented and there are fewer students in the upper grades who require time-consuming assessment or pull-out special education services. Consequently, school psychologists have more time for their balanced role of assessment, consultation, and intervention specialist for mental health and behavioral as well as academic problems. Second, assessment and intervention are linked bidirectionally. In all tiers, assessment (of the student — tiers 1 and 3, or of the curriculum — tier 2) precedes intervention, but response to intervention is also assessed routinely and intervention is modified depending on the student’s response to it. Third, in all tiers assessment and intervention should be science-based.

As we make the transition to the 21st century, education is evolving from a philosophical enterprise based on the latest fad and frequent pendulum swings to a scientific enterprise based on assessment and instructional practices that have been validated in research. For example, in the United States two recent developments are promoting science-based educational practice. The first development is the Workforce Investment Act of 1998, which established the National Institute for Literacy to provide a national resource for literacy programs. This institute offers the most up-to-date research information available — for example, from the National Institute of Child Health and Human Development (NICHD) and other sources — on phonemic awareness,
systematic phonics, fluency, and reading comprehension. All schools that receive federal funding for reading (e.g., from Titles I and VII of the Elementary and Secondary Education Act of 1965, the Head Start Act, and the Individuals with Disabilities Education Act) are required to consult the available information. The Workforce Investment Act also supports the creation of new ways to offer services of proven effectiveness. The second development—the National Reading Excellence Act—empowers the National Institute for Literacy to provide competitive grants to states that implement science-based reading programs for the purpose of increasing the odds that all children will be readers by the end of third grade.

In this chapter we provide a brief overview of the scientific support for each of the three tiers of the assessment for intervention model and introduce an assessment instrument—Process Assessment of the Learner Test Battery for Reading and Writing (PAL-RW) (Berninger, 2001). The PAL-RW is a science-based instrument that can be used for assessment at each tier of the model, as we explain. The Learning Triangle in Figure 7.1 captures the complexities of the assessment for intervention process and we will reference it throughout the chapter to assist our readers in integrating the dynamic interrelationships among individual differences in learners' processing abilities (what the PAL-RW is designed to assess), pedagogical approaches for intervention (what teachers do), and curriculum/materials (what teachers use). We will also refer to the Process Assessment of the Learner (PAL): Guides for Reading and Writing Intervention (PAL Intervention Guides; Berninger, 1998a), which is an instructional resource for pedagogical approaches and curriculum/instructional materials linked to the PAL-RW.

![FIGURE 7.1](Image)  
The Learning Triangle.
FIRST TIER: SCREENING FOR EARLY INTERVENTION

Research-Validation of Screening Measures

Reading

Numerous studies have documented the critical role of phonological skills in learning to read (e.g., Adams, 1990; Ball & Blachman, 1991; Bradley & Bryant, 1983; Liberman, Shankweiler, Fischer, & Carter, 1974; Wagner & Torgesen, 1987). The research evidence for the important role of orthographic skills (e.g., Adams, 1990; Barker, Torgesen, & Wagner, 1992; Berninger & Abbott, 1994; Berninger, Yates, & Lester, 1991; Cunningham & Stanovich, 1990; Olsen, Forsberg, & Wise, 1994) and rapid naming skills (e.g., Denckla & Rudel, 1976; Felton, Wood, Brown, Campbell, & Harter, 1987; Wolf, 1984, 1986, 1991; Wolf, Bally, & Morris, 1986) in learning to read is also growing.

In our earlier assessment studies with unreferred primary-grade children (first through third grade), both orthographic and phonological measures contributed unique variance over and beyond their shared covariance to the prediction of real word reading and pseudoword reading (e.g., Berninger & Abbott, 1994). In our more recent assessment studies with a referred sample of elementary students (first through sixth grade) who qualified as probands in our family genetics study for reading and writing disability (Berninger, Abbott, Thomson, & Raskind, 2001), four language factors—Verbal IQ, phonological, orthographic, and rapid naming—were used in structural models for predicting three reading factors: reading accuracy, reading rate, and reading comprehension. For reading accuracy, only the direct paths from the orthographic factor and the phonological factor were significant, indicating that each contributed unique variance. For reading rate, only the direct paths from the orthographic factor and the rapid naming factor were significant, indicating that each contributed unique variance. For reading comprehension, only the direct paths from Verbal IQ and the phonological factor were significant, indicating that each contributed unique variance. Thus, all four predictor factors are valid for screening for reading problems but each factor predicts different components of the functional reading system in school-aged children with reading problems. The orthographic factor predicts reading accuracy and reading rate. The phonological factor predicts reading accuracy and reading comprehension. Rapid naming predicts reading rate. Verbal IQ predicts reading comprehension.

Reading real words and reading pseudowords tap different mechanisms that may not develop at the same rate (see Berninger, 1998a). Reading real words taps the word-specific mechanism that activates orthographic, phonological, and semantic codes for words in the mental dictionary, whereas reading pseudowords taps the phonological decoding mechanism that activates orthographic and phonological but not semantic codes. Reading real words may also activate phonological decoding in addition to the word-specific
mechanism. Children who grow significantly in both mechanisms achieve at a higher level in reading than children who grow in only one mechanism, and children who grow in only one mechanism achieve at a higher rate than children who grow in neither mechanism (Berninger, 1994; Berninger, Abbott, & Stage, 1999).

**Writing**

Our research over the past decade has shown that orthographic, fine motor, oral language (especially phonological), and reading skills are important for learning to write (for a review, see Berninger, 1994, 1998a). Orthographic coding is the rapid representation and analysis of written words in short-term memory, whereas phonological coding is the representation and analysis of spoken words in short-term memory. Our earlier assessment studies with unreferred primary-grade students showed that orthographic coding, fine-motor skills, and orthographic-motor integration contributed uniquely to handwriting and compositional fluency (Berninger, Yates, Cartwright, Rumberg, Remy, & Abbott, 1992); orthographic and phonological coding contributed uniquely to spelling (Berninger & Abbott, 1994); and Verbal IQ contributed uniquely to compositional quality along with the skills that also contributed to compositional fluency (Berninger et al.). More recently we studied the phenotype for writing disability from a systems perspective (see PAL Intervention Guides for discussion of the functional reading and writing systems) with a referred sample (Berninger, Abbott, Thomson et al., 2001). The specific language factors that were the best predictors depended on the component writing skill in the structural model. For example, in children with reading and writing disabilities, the orthographic factor contributed uniquely to handwriting, but the orthographic and phonological factors contributed uniquely to spelling and composition.

**Screening Battery**

Based on these studies with unreferred and referred students, we propose the screening battery for beginning reading in Table 7.1 and for beginning writing in Table 7.2. The PAL-RW in Tables 7.1 and 7.2 has the advantage that all measures are normed on the same sample, which is representative of the U.S. population. Some of the measures in these two tables are group-administered, but some are individually administered. In most cases the screening should require 20 minutes or less.

**Research-Validated Early Intervention**

The purpose of the screening is to identify students who would benefit from early intervention. The exact criterion used for identifying students for early
### TABLE 7.1
Screening Battery for First Tier Early Intervention for Reading

<table>
<thead>
<tr>
<th>Grade/Skill</th>
<th>Subskill</th>
<th>Tests&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kindergarten</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthographic</td>
<td>Naming alphabet letters</td>
<td>WIAT II</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rate</td>
<td>PAL-RW</td>
</tr>
<tr>
<td>Phonological</td>
<td>Nonword memory</td>
<td>CTOPP</td>
</tr>
<tr>
<td></td>
<td>Rhyming</td>
<td>PAL-RW</td>
</tr>
<tr>
<td>Rapid Naming</td>
<td>Colors and/or Objects</td>
<td>NEPSY, CTOPP</td>
</tr>
<tr>
<td></td>
<td>Syllable Segmentation</td>
<td>PAL-RW</td>
</tr>
<tr>
<td><strong>First grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthographic&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Alphabet Task</td>
<td>PAL-RW, WIAT II</td>
</tr>
<tr>
<td>Phonological&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Phoneme Segmentation</td>
<td>PAL, CTOPP</td>
</tr>
<tr>
<td>Rapid Naming&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Letters</td>
<td>PAL, CTOPP</td>
</tr>
<tr>
<td></td>
<td>Letters and Numbers</td>
<td>CTOPP</td>
</tr>
<tr>
<td>Verbal Intelligence (estimated)</td>
<td>Expressive Vocabulary</td>
<td>WISC-III</td>
</tr>
<tr>
<td></td>
<td>vocabulary</td>
<td></td>
</tr>
<tr>
<td>Word-Specific Reading</td>
<td>Accuracy</td>
<td>WRMT-R, WIAT II</td>
</tr>
<tr>
<td></td>
<td>Rate</td>
<td>TOWRE</td>
</tr>
<tr>
<td>Pseudoword Reading</td>
<td>Accuracy</td>
<td>WRMT-R, WIAT II, PAL-RW</td>
</tr>
<tr>
<td></td>
<td>Rate</td>
<td>TOWRE</td>
</tr>
</tbody>
</table>

<sup>a</sup>C TOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999)

NEPSY (Korkman, Kirk, & Kemp, 1998)

PAL-RW = Process Assessment of the Learner Test Battery for Reading and Writing (Berninger, 2001)

TOWRE = Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999)

WIAT II = Wechsler Individual Achievement Test—Second Edition (The Psychological Corporation, in press)

WRMT-R = Woodcock Reading Mastery Test—Revised (Woodcock, 1987)

<sup>b</sup>For any child who does not score above the at-risk range on the first-grade screen, also administer the kindergarten screen.

Intervention can be established based on the population a particular school serves. If all the students tend to be below the national mean, the early intervention might be done with the whole class. In schools with average performance above the national mean, only those children scoring at or below a locally-defined criterion (e.g., one-third standard deviation or more below the national mean) might be flagged for early intervention. In schools
TABLE 7.2
Screening Battery for First Tier Early Intervention for Writing

<table>
<thead>
<tr>
<th>Grade/Skill</th>
<th>Subskill</th>
<th>Tests a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kindergarten</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handwriting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Copy Task</td>
<td>PAL-RW</td>
</tr>
<tr>
<td>Automaticity</td>
<td>Alphabet Task</td>
<td>PAL-RW, WIAT II</td>
</tr>
<tr>
<td>First Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthographic</td>
<td>Alphabet Task</td>
<td>PAL-RW, WIAT II</td>
</tr>
<tr>
<td>Phonological</td>
<td>Phoneme Segmentation</td>
<td>PAL, CTOPP</td>
</tr>
<tr>
<td>Verbal Intelligence (estimated)</td>
<td>Expressive Vocabulary</td>
<td>WISC III vocabulary</td>
</tr>
<tr>
<td>Pseudoword Reading</td>
<td>Decoding</td>
<td>WRMT-R, WIAT II, PAL-RW</td>
</tr>
<tr>
<td>Spelling</td>
<td>Dictation</td>
<td>WIAT II, WRAT-3</td>
</tr>
</tbody>
</table>

aCTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte 1999)
PAL-RW = Process Assessment of the Learner Test Battery for Reading and Writing (Berninger, 2001)
WIAT II = Wechsler Individual Achievement Test—Second Edition (The Psychological Corporation, in press)
WRMT-R = Woodcock Reading Mastery Test—Revised (Woodcock, 1987)

with average performance at the national mean, those children at or below one standard deviation might be flagged. Regardless of the exact criterion used, it is based on performance on language skills related to reading and writing acquisition and also on reading or writing skills rather than on IQ-achievement discrepancy, which is not a valid predictor of which children will benefit from early intervention for reading (Stage, Abbott, Jenkins, & Berninger, 2000).

**Phonological Awareness**

Students at or below the criterion on phonological measures may benefit from the 24-lesson phonological awareness training program on pages 196 to 219 of the PAL Intervention Guides. The “sound games” described there include listening to target sounds in words, detecting a deleted sound, deleting a target sound, and substituting a target sound for the deleted sound. First graders who played these sound games, which take about 10 minutes, improved significantly in their standard scores for age in word recognition (Berninger, Abbott, & Stage, 1999). Second graders who played these sound games improved significantly in their standard scores for age in spelling (Berninger, Vaughan, et al., 1998) or word recognition (Berninger & Traweek, 1991).
Orthographic Awareness

Students at or below the criterion on orthographic measures may benefit from the “looking games” on pages 191 to 193 of the PAL Intervention Guides. These looking games include spelling back a whole word, a target letter, or letter group in a word; or retrieving the letters that come before or after a target letter in the alphabet. First graders who played some of these games, which take about 10 minutes, improved significantly in their standard scores for age in word recognition (Berninger, Abbott, & Stage, 1999). Second graders who played some of these games improved significantly in their standard scores in spelling (Berninger et al., 1998) or word recognition (Berninger & Traweek, 1991).

Talking Letters

Students who struggle with beginning reading (Berninger, Abbott, et al., 2000) or with beginning spelling (Berninger et al., 1998; Berninger, Vaughan, et al., 2000) benefit from explicit instruction in the alphabet principle. The PAL Intervention Kit contains the PAL Intervention Guides and the Talking Letters Teacher Guide and Student Desk Cards (Berninger, 1998b) used in our research to teach the alphabet principle to beginning readers and spellers. The Talking Letters Teacher’s Guide explains the theoretical concepts underlying this approach, which draws on work in linguistics and cognitive neuroscience, to teach spelling–phoneme correspondence. It also has blackline masters for overheads for group instruction, which can be used in conjunction with student cards in the PAL Intervention Kit for children to use at their desk while reading and composing.

Rapid Automatic Naming

Few research studies have investigated how to remediate a deficit in rapid automatic naming, which is the most prevalent deficit in children and adults with reading and writing disabilities (Berninger, Abbott, et al., 2001). Training studies by Levy, Abello, and Lysynchuk (1997) suggest that pretraining in speed of naming words before they are encountered in text and repeated reading of text may improve ability to access name codes for printed words quickly and automatically.

Vocabulary Knowledge

Although it is unlikely that reading instruction will substantially alter verbal reasoning ability for age, it is likely that instruction that promotes a schema for learning word meanings (Nagy & Scott, 1990) may improve reading comprehension and word-level expression in written composition. For instruc-
Handwriting Automaticity

Students at or below the criterion on the alphabet task or the finger function tasks may benefit from the Handwriting Lessons program (Berninger, 1998c) in the PAL Intervention Kit. This program is based on research in which first graders who received training that combined numbered arrow cues and writing letters from memory with frequent letter naming improved in handwriting automaticity and in composition (Berninger, Vaughan, et al., 1997). The Handwriting Lessons manual contains detailed information on the theoretical rationale, derived from cognitive neuroscience, and the scientifically validated approach to teaching handwriting automaticity. The manual also includes blackline masters for the 24 lessons used by Berninger, Vaughan, et al. with first graders at-risk for handwriting problems. Each teacher-directed lesson provides practice for each of the 26 letters as well as a composition starter to promote transfer from handwriting automaticity training to authentic written communication.

Instructional Protocols Aimed at All Levels of Language

Design experiments (Brown, 1992) integrate components needed to achieve desired outcomes. The reading tutorial on pages 233 to 256 of the PAL Intervention Guides integrates orthographic and phonological awareness, alphabet principle, whole word and word-family strategies, oral reading and rereading for fluency, and comprehension monitoring within the same instructional session. For other similar tutorials based on the principle of design experiments, see Abbott, Reed, Abbott, and Berninger (1997). Likewise, for writing tutorials aimed at all levels of language in the same instructional session, see pages 257 to 261 of the PAL Intervention Guides and Berninger, Abbott, Whitaker, Sylvester, and Nolen (1995).

SECOND TIER:
ASSESSING CURRICULUM, MODIFYING THE REGULAR PROGRAM, PROGRESS MONITORING, AND PREREFERRAL COLLABORATIVE PROBLEM SOLVING

Most psychoeducational assessment focuses on the student—that is, learner variables rather than pedagogical methods, curriculum components, or instructional materials (see Figure 7.1). In contrast, most teachers focus on pedagogy, curriculum, and instructional materials with little consideration of how individual differences in students affect their responses to

instruction. From the lay public's perspective, inadequate student learning outcome is attributed primarily to the teacher. However, all angles of The Learning Triangle in Figure 7.1 influence student learning outcome, and all angles and intersecting sides of The Learning Triangle need to be taken into account in the assessment and intervention process. Berninger (1998a) discusses the importance of assessing the curriculum before assuming that a poor student learning outcome is due to learner characteristics. Based on existing research, she outlines what necessary components of curriculum should be in place for a student to develop a functional reading system and a functional writing system. She argues that before a student is referred for comprehensive assessment to determine eligibility for special education services, the collaborative team should assess the curriculum, add missing but necessary components, and empirically monitor student progress in response to the modified general education program. In some cases students respond favorably to modifications in the curriculum and assessment for special education is not needed. In this section we review empirical research that supports an approach in which, following an assessment of curriculum, curriculum-based materials are used to monitor student progress in response to curriculum modifications.

Over the past decade, a growing number of states (e.g., Florida, Illinois, Iowa, Kansas, Minnesota, and Washington) have engaged in initiatives that utilize collaborative problem-solving prereferral interventions with general education students who experience academic difficulties (Browning, Davidson, & Stage, 1998; Ysseldyke & Marston, 1998). Initially, teachers work directly with the student and the student's parent(s). If the teacher is unable to change the student's academic progress, the student is discussed by a collaborative problem-solving team that may include the following professionals: the school psychologist, the special educator, the speech and language pathologist, the social worker, the nurse, the principal, the Title I teacher, and the general education teacher. The team uses consultation methodology in which problems are identified, possible solutions are brainstormed, a potential solution is implemented and evaluated, another potential solution is implemented and evaluated if the first approach is not effective, and so on until an effective solution is found (e.g., Cole & Siegel, 1990; Ikeda, Tilly, Stummme, Volmer, & Allison, 1996; Reschly & Ysseldyke, 1995).

The team uses curriculum-based measures (CBM) both to identify problems and to monitor student progress (Fuchs & Fuchs, 1986; Shapiro, 1996; Shinn, 1995). In the area of reading, the number of correctly read words per minute from the student's reading curriculum is often used as the CBM tool. Oral reading fluency or the number of correctly read words per minute yields adequate concurrent validity with various standardized measures in the elementary grade levels (Marston, 1989). Reading readiness CBMs include letter-naming fluency (Kaminski & Good, 1996) and letter-sound fluency (Daly, Wright, Kelly, & Martens, 1997). In order to identify the extent of a student's
academic problem compared to the average student at that grade level, normative data are collected three or four times throughout the school year (e.g., fall, winter, and spring; see Habedank, 1995, for a description of this procedure). The distribution of scores on the various CBM tools allows the collaborative problem-solving team to determine where a given student falls on the continuum of scores by grade level and classroom. With the school normative data, aim lines are graphed for the expected progress that a student needs to make in order maintain pace with classmates. Weekly progress is monitored so that the collaborative problem-solving team can determine (a) if the student is making progress toward attaining a minimal level of competency, which is usually set at the 25th percentile and (b) if there is need for further academic intervention. In summary, the collaborative problem-solving teams use prereferral reading interventions that are monitored with letter-naming fluency, letter-sound fluency, and oral reading fluency to determine whether the student is making sufficient academic gains in the general education setting.

The Student Responsive Delivery System (Browning, Davidson, & Stage, 1999), a state-wide pilot project in Washington using collaborative problem-solving teams, is described as a Tier 2 model other states could adopt, especially for schools serving a high percentage of minority and low-income students who may score poorly on standardized tests. For use of CBM to make nondiscriminatory educational programming decisions without relying solely on special education programs, see Baker, Plasencia-Peinado, and Lezcano-Lytte (1998), Canter and Marston (1994), and Shinn, Collins, and Gallagher (1998).

The second author of this chapter is a member of the team that conducted the following three studies.

**Kindergarten to First Grade CBM Study**

In the first study, the participants were 59 kindergarten students from a school serving ethnically diverse and low-income students in rural Washington (i.e., 62% Native American, 24% Hispanic, 12% European American, and 1% Asian, with 81% of the student population qualifying for free or reduced lunch) (Stage, Sheppard, Davidson, & Browning, 1999). Two measures were individually administered at the end of their kindergarten year—letter-sound fluency and letter-naming fluency. These same students received four oral-reading fluency CBMs during their first-grade year. Research has shown that (a) students who cannot form letter-sound associations are at-risk for learning the alphabet principle, which is crucial for beginning reading (e.g., Adams, 1990; Berninger, 1998a; Bradley & Bryant, 1983; Ehri & Robbins, 1992;

---

1The Student Responsive Delivery System in Washington State is sponsored by the State Office of Public Instruction, the Washington State Association of School Psychologists, the Washington State Speech and Hearing Association.
Foorman, 1995; Share & Stanovich, 1995), (b) students who cannot name letters rapidly and automatically are at-risk for automatic access to name codes for written symbols during lexical access to the mental dictionary (e.g., Perfetti, 1985; Wagner & Torgesen, 1987), which is also crucial for beginning reading (e.g., Wolf, 1984, 1991; Wolf, Bally, & Morris, 1986), and (c) oral-reading fluency is an index of the degree to which students can automatically recognize words and coordinate the word recognition and sentence-level comprehension processes temporally in working memory (see Berninger, 1999, 2001).

Three different letter-naming forms were used. Each form had 104 lowercase and uppercase letters randomly arranged across 11 rows with 10 letters on each row, with the exception of the last row, which had only 4 letters. Likewise, three different letter-sound forms were used that were arranged similarly to the letter-naming forms. Correct letter-sounds were coded in association with the Open Court Reading Series (1995) Guide to Letter-Sounds. Alternate forms of the letter-sounds and letter-naming tasks yielded high correlations and interrater reliability (rs = .90s). During first grade, four oral-reading fluency measurements were administered in October, January, March, and May (the University of Oregon CBM Reading Probes available from the School Psychology Program). A growth-curve analysis using Hierarchical Linear Modeling (HLM) (Bryk, Raudenbush, & Congdon, 1996) was used to evaluate whether kindergarten letter-sound and letter-naming fluency measures predicted the students' initial first-grade oral-reading fluency and growth in oral-reading fluency over the course of first grade.

The average letter-sound fluency score was 9.39 (SD = 9.17) correct letter-sounds per minute with a range of 0 to 40 letter-sounds per minute. The average letter-naming fluency score was 20.44 (SD = 14.41) correct letter-names per minute with a range of 1 to 63. The two measures were correlated (r = .79). The results of the growth-curve analysis for the group showed that kindergarten letter-sound fluency performance predicted initial first-grade oral-reading fluency performance (p = .001), but kindergarten letter-naming fluency performance did not (p = .464). However, kindergarten letter naming did predict slope in first-grade oral-reading fluency (p = .035), whereas kindergarten letter-sound fluency did not (p = .08). These results suggest that a certain level of letter-sound knowledge is needed initially for oral-reading skill to develop, but that rate of lexical access to name codes will predict response to reading instruction in first grade. Further evidence of the importance of rate of lexical access was that letter-naming fluency predicted word recognition performance better than the phonological awareness and phonological coding in working memory tasks when all measures were entered into regression equations simultaneously, replicating work by Torgesen et al. (1999).

However, when these group analyses were supplemented with the calculation of HLM growth curves for individual children, interesting relationships
between growth in letter-name fluency and letter-sound fluency became evident. These relationships demonstrated that children need to master both skills for fluent oral reading (see Figure 7.2). The middle curve in Figure 7.2 shows the performance of the students (44%, 26/59) who performed in the average range on the letter-naming and the letter-sound fluency tasks (i.e., between +1 SD and −1 SD). The growth curves above this line show the growth made by students (10%, 6/59) who were 1 standard deviation above average on letter-naming fluency only, growth made by students (8%, 5/59) who were 1 standard deviation above average on letter-sound fluency only, and growth by students (12%, 7/59) who were 1 standard deviation above average on both measures. The growth curves below this line show the growth made by students (13%, 8/59) who were 1 standard deviation below average on letter-naming fluency only, growth by students (5%, 3/59) who were 1 standard deviation below average on letter-sound fluency only, and growth made by students (7%, 4/59) who were 1 standard deviation below average on both measures. Clearly, those who were most fluent in oral reading were best in both letter-sound fluency and letter-naming fluency and those who were
least fluent in oral reading were the worst in letter-sound and letter-naming fluency.

The collaborative problem-solving team used the data generated in the first year of the project to monitor the progress of individual students in the next cohort of kindergartners in the same school. Aim lines based on the first-year data were used to evaluate progress of the new cohort in the second year of the project so that those who needed a modification of the regular curriculum could be identified. In this study the modification of the curriculum consisted of additional instruction in the alphabet code with an attempt to make the code as explicit as possible. Thus, local schools can generate building-based norms for identifying children who need Tier 2 intervention and for evaluating their responses to modification of the regular curriculum.

First-to Second-Grade CBM Study

In the second study, conducted at the same school, 99 first-grade students' end-of-the-year oral-reading fluency was used to predict their initial second-grade oral-reading fluency (Stage, in press). A previous study had indicated that students' reading progress is influenced by their classmates' ability level (Share, Jorm, Maclean, & Matthews, 1984). Therefore, slopes for the classroom in which students are instructed may be a more sensitive index for aim lines than slope for grade in a local school (Fuchs, 1998).

The growth-curve analysis showed that students' first-grade oral-reading fluency significantly predicted their initial second-grade oral-reading fluency ($p < .001$), but that the classroom effect did not reliably distinguish growth in oral-reading fluency over the course of the year ($p = .35$). The average growth in the students' second-grade oral-reading fluency was divided into quartiles so that students who were performing in the lower 25th percentile could be monitored. Figure 7.3 shows the growth curves in oral-reading fluency of the second-grade students by quartile. The difference in the spacing between the three quartile slopes indicates that the students' oral-reading fluency was not normally distributed. The distribution was positively skewed, with the majority of students performing between the first and second quartiles. Therefore, the growth-curve analysis was also conducted after transforming the first-grade reading-fluency variable by taking the square root of each score (see Tabachnick & Fidell, 1989, for a discussion on the transformation of non-normal distribution scores). The results yielded similar findings. Because the statistical results were not altered, the results of the initial analysis are considered adequate for interpretation and graphic representation.

At the end of the school year, 28 students who continued to score in the lower 25th percentile were invited to attend summer school that provided individualized reading instruction using the Open Court Reading Series (1995), which research has shown is an effective program for at-risk beginning read-
Second-grade reading growth by first-grade quartiles and summer-school student growth.

ers (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998). On average, students attended 25.5 summer-school days with a standard deviation of 4.4 days. The range was from 29 to 16 days. Growth-curve analysis showed that, on average, students had significant growth ($p < .0001$) during summer school, but there was individual variability in the summer-school students' growth over the course of the year ($p < .0001$). To investigate the number of students who had significant reading growth, the student's slope was divided by twice the standard error. The results indicated that all the students had significant growth at $p < .05$. A comparison of the summer-school students' August oral-reading fluency with the general education students' May oral-reading fluency showed that 71% (20 out of 28 students) of the summer-school students read at least 50 or more words per minute, which was above the 25th percentile rank as determined by the second-grade May distribution. However, it is possible that if the general education students' oral-reading fluency had been measured in August, they might have shown the same increase, negating the appearance that the summer-school students had actually improved compared to the regular education students (see Cook & Campbell, 1979).
The eight students who did not make sufficient progress were targeted for collaborative problem-solving intervention for the following academic year in order to determine whether they required specialized educational services. That is, students who did not respond to initial early intervention during the regular school year (Tier 1) or additional curriculum modification (more instruction during summer school—Tier 2) became candidates for Tier 3.

**Evaluating Prereferral Intervention across Sites**

Pilot-site buildings included 14 elementary schools, 2 middle schools, 2 high schools, and 1 middle/high school combination. Program evaluation data collected during the 1998–99 school year indicated that 215 students participated in the collaborative problem-solving process, with 138 students (i.e., 64%) needing no further intervention because their academic and/or behavioral difficulty was resolved (Browning et al., 1999). Of the 215 students, 58 (27%) were further considered for special education services, 28 (13%) were found eligible for special education services, and 19 (9%) were determined to be in need of monitoring by the problem-solving team the following academic year.

Four findings are emphasized. First, the number of students requiring full assessment for special education was reduced by 73%. Second, only 13% of the students needing academic and behavioral assistance were found to require special education services. A service delivery model is needed that includes a range of services from early intervention to modified regular program to pull-out services. Such a model will reduce the number of assessments and pull-out placements that are necessary. Third, CBM can be used to identify and monitor the progress of children who do not respond or respond slowly to Tier 1 early intervention and thus need additional intervention in the form of greater intensity (additional instruction) or modification of the regular program. Fourth, CBM is most likely to be effective if implemented within a collaborative, problem-solving framework in which the multidisciplinary team plans and evaluates curriculum modifications.

**Commercially Available CBM Tools**

The *Academic Competence Scales* (ACES) (Di Perna & Elliott, in press) provide an empirically validated approach to using teacher ratings in collaborative problem solving for students referred for academic problems. The *Academic Intervention Monitoring System* (AIMS) (Di Perna & Elliott) provides an empirically validated approach to using teacher ratings in monitoring progress. Another tool for monitoring progress is the *Standard Reading Passages* (Children's Educational Services, 1987). The PAL-RW could also be used to monitor student response to curriculum modifications for prereferral intervention.

Orthographic, phonological, and rapid automatic naming (RAN) tasks,
like the ones on the PAL-RW, predict growth curves for prereferral interventions for reading (e.g., Berninger, Abbott, et al., 2000; Berninger, Abbott, Zook, et al., 1999); and orthographic and phonological tasks, like the ones on the PAL-RW, predict growth curves for prereferral interventions for spelling (Berninger et al., 1998). For examples of prereferral intervention protocols for reading and use of standardized tests in evaluating response to intervention, see Abbott et al. (1997), Berninger, Abbott, et al. (1997), and Berninger and Traweek (1991). For examples of prereferral intervention protocols for writing and use of standardized tests in evaluating response to intervention, see Berninger et al. (1995), Berninger, Abbott, et al. (1997), Berninger et al. (1998), and Berninger (1998a, pp. 257–261). Berninger, Abbott, and Stage (1999) found that following a moderately intensive intervention for first graders at-risk for reading problems (24 twenty-minute lessons over a four-month period that aimed instruction at all components of a functional reading system), half of the 128 children reached grade level or above and maintained their relative gains at the beginning and end of second grade. The other half, who had improved, as a group, but were still not at grade level, received continuous tutoring in second grade and made additional relative gains, demonstrating that some students will need more-intensive intervention over a longer period of time. Only a few children failed to respond to early intervention by the end of second grade. We turn now to Tier 3 in the model, for those cases in which Tier 1 and Tier 2 are not sufficient.

THIRD TIER: DIAGNOSIS AND TREATMENT PLANS

Tier 1 assessment for intervention can prevent many reading and writing disabilities and/or reduce the severity of these disabilities, but it does not eliminate all disabilities (Berninger, Abbott, & Stage, 1999; Berninger, Abbott, Zook, et al., 1999). When children do not respond to Tier 1 early intervention or Tier 2 modification of the regular curriculum, or continue to struggle to keep up with the regular curriculum, then Tier 3—in-depth assessment—is called for. The purpose of this assessment is not just to decide whether the student qualifies for special education services and, if so, under which category—that is, to make an eligibility and placement decision. The purpose of the Tier 3 assessment is also to diagnose why the student is having so much difficulty; that is, if there is a discrepancy between ability and achievement, why there is a discrepancy. Simply demonstrating that a discrepancy exists is not sufficient to diagnose a learning disability, which is only one cause of underachievement. This process also helps to identify those students who may not demonstrate an ability-achievement discrepancy, but who have significant deficits in the underlying processes essential to the acquisition of reading and writing skills such that intensive instructional intervention is warranted. Further, it is necessary to demonstrate that science-validated marker
variables for reading or writing disability are also present. The PAL-RW is designed to assess such research-validated marker variables. Although education professionals are often reluctant to "label" students, many parents seek diagnostic explanations and acknowledgment of learner differences (see lower left corner of Figure 7.1 and Berninger, 1998a). A small percentage of students continue to struggle despite appropriate instruction and sometimes considerable extra help; in such severe cases the diagnosis of dyslexia or dysgraphia may be warranted. Failure to go beyond eligibility and placement decisions in assessment to diagnoses and acknowledgment of learning differences can fuel adversarial relationships between the family and school (Berninger, 1998a).

The approach to diagnosis that we recommend utilizes clinical hypothesis testing. For an example of branching diagnosis applied to clinical hypothesis testing for writing disability, see Berninger and Whitaker (1993). To test clinical hypotheses, we use multimodal assessment (e.g., Shapiro, 1996), including interviewing parents and reviewing records for developmental, educational, and family history; observing the student in and out of the classroom; interviewing the student; analyzing parent and teacher rating scales; administering standardized, normed tests and criterion-referenced tests; and reviewing portfolios or work products. Results are used for differential diagnosis involving five axes (see Berninger, 1998a): Axis 1, Development Across Multiple Domains; Axis 2, Comorbid Medical Conditions; Axis 3, Brain Functions; Axis 4, Academic Functioning in School; and Axis 5, Community, Family, School, and Classroom Factors. Students with mental retardation or with a specific reading or writing disability will be classified differently on Axis 1 and possibly Axis 2, but may have some similar problems on Axis 4 (e.g., poor reading relative to age peers). A student with a neurologically based learning disorder due to brain damage will have an Axis 2 diagnosis, whereas a student with dyslexia and/or dysgraphia will not because, although these disorders are brain-based, they are not related to brain damage or a medical condition (Berninger, 1994).

We begin the process of differential diagnosis by ruling mental retardation in or out (see Table 7.3), because this Axis 1 diagnosis is relevant to establishing reasonable expectations for rate of learning and levels of learning outcomes. If mental retardation is ruled out, we continue the differential diagnosis process by considering whether a primary language disability exists (see Table 7.3). Although children with a primary language disability will probably have reading problems, not all students with reading problems have a primary language disability. If a student has both a language disability on Axis 1 and reading problems on Axis 4, he or she may need treatment and accommodation throughout the school years because so much instruction in school involves oral language. If a primary language disability is ruled out, we proceed to determine whether a reading disability exists, with or without comorbid writing disability (see Table 7.3). When the specific reading
TABLE 7.3
Differential Diagnosis among Mental Retardation, Primary Language Disability, and Specific Reading and/or Writing Disability

1. Does the student meet criteria for mental retardation?
   A. Significant delays over time in cognitive, language, motor, and social development not attributable to cultural difference or lack of stimulation?
   D. Significant delays over time in adaptive functioning?
   E. Comorbid developmental disabilities and/or medical conditions?

2. Does the student meet criteria for primary language disability?
   A. Expressive language disorder?
   B. Mixed receptive and expressive language disorder?
   C. Comorbid developmental disabilities and/or medical conditions?

3. Does the student meet criteria for specific reading disorder (including dyslexia, i.e., severe reading disorder)?
   A. Developmental history
      1. Did the student have preschool language difficulties?
      2. Did the student struggle in learning to name letters?
      3. Did the student struggle in learning letter-sound correspondence?
      4. Did the student struggle in recognizing single words out of context?
      5. Is the student's oral reading of text slow and dysfluent?
      6. Does the student persist in making spelling errors beyond grade 3 that resemble the invented spelling of younger children?
      7. Is there a family history of reading and/or writing problems?
   B. Educational history
      1. Was the student given explicit instruction in orthographic and phonological awareness, the alphabet principle for word recognition and spelling, reading comprehension, and the planning, translating, and reviewing/revising components of writing?
      2. Was the student given daily experience in reading specific words in reading material at his or her instructional or independent reading level and in writing for varied purposes?
      3. Was the student taught by a certificated teacher with supervised experience in teaching reading and writing to at-risk students and with appropriate training in the structure of language?
      4. Was the student a slow responder or nonresponder to appropriate reading and writing instruction over a 2-year period?
   C. Are there comorbid attention, math, and/or writing problems? Can mental retardation, primary language disorder, other developmental disorder (e.g., autism and pervasive developmental disorder), or neurological condition (e.g., cerebral palsy and neurofibromatosis) be ruled out?
   D. Is there significant, unexpected achievement relative to verbal reasoning ability in the following measures?
      1. Single word reading—accuracy
         a. Word specific mechanism (WRMT-R Word Identification, WIAT II Word Reading, WRAT-3 Reading)

(continues)
**Table 7.3—Continued**

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Automaticity of single-word recognition</td>
<td>a. Real word (TOWRE real-word efficiency)</td>
</tr>
<tr>
<td></td>
<td>b. Pseudoword (TOWRE pseudoword efficiency)</td>
</tr>
<tr>
<td>3. Fluency of oral reading of text (GORT-3 rate)</td>
<td></td>
</tr>
<tr>
<td>4. Comprehension/Memory</td>
<td>a. Text-based</td>
</tr>
<tr>
<td></td>
<td>1) Cloze (WRMT-R Passage Comprehension)</td>
</tr>
<tr>
<td></td>
<td>2) Story retell (QRI, PAL-RW Test Battery)</td>
</tr>
<tr>
<td></td>
<td>b. Integration of text and background knowledge</td>
</tr>
<tr>
<td></td>
<td>1) Factual and inferential questions—multiple choice (GORT-3 comprehension)</td>
</tr>
<tr>
<td></td>
<td>2) Factual and inferential questions—open-ended (QRI, WIAT II)</td>
</tr>
<tr>
<td>5. Marker variables</td>
<td>a. Orthographic coding and/or representation in memory</td>
</tr>
<tr>
<td></td>
<td>b. Phonological coding and/or representation in memory</td>
</tr>
<tr>
<td></td>
<td>c. Rapid automatic naming</td>
</tr>
</tbody>
</table>

**E. Is there significant, unexpected achievement in writing relative to verbal reasoning ability in the following measures?**

1. Handwriting automaticity (PAL-RW, WIAT II alphabet task)
2. Spelling (WIAT II, WRAT-3 Spelling)
3. Compositional fluency and quality (WIAT II, WJ-R)

**4. Marker variables**

| a. Alphabet task                             |
| b. Orthographic coding                       |
| c. Finger function tasks                     |
| d. Phonological nonword memory               |
| e. Phonological deletion                     |
| f. Orthographic word-specific representation |
| g. Pseudoword reading                         |
| h. Verbal working memory                      |

---

*a CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte 1999)*

GORT-3 = Gray Oral Reading Test—Third Edition (Wiederholt & Bryant, 1992)

NEPSY (Korkman, Kirk, & Kemp, 1998)

PAL-RW = Process Assessment of the Learner Test Battery for Reading and Writing (Berninger, 2001)

QRI = Qualitative Reading Inventory (Leslie & Caldwell, 1990)

WJ-R = Woodcock Johnson Psychoeducational Battery—Revised (Woodcock & Johnson, 1990)
7. Assessment for Reading and Writing Intervention

Table 7.3—Continued

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT II</td>
<td>Wechsler Individual Achievement Test, Second Edition (The Psychological Corporation, in press)</td>
</tr>
<tr>
<td>WRMT-R</td>
<td>Woodcock Reading Mastery Test—Revised (Woodcock, 1987)</td>
</tr>
</tbody>
</table>

Verbal IQ may be a more realistic gauge of expected level of achievement of functional reading or writing systems than Performance IQ (see Berninger, 1998a). Reading disability is more likely to have a genetic or constitutional basis when reading ability is discrepant from IQ than when both reading and IQ are low (Olson, Datta, Gayan, & DeFries, in press).

The Process Assessment of the Learner Test Battery for Reading and Writing has subtests for each of these marker variables, which also lead to instructionally relevant recommendations for students whose achievement is not below the level expected based on Verbal IQ.

disability involves severe difficulty in learning to read and spell words, it is often referred to as dyslexia. See Berninger (2000) for a description of how the phenotype for dyslexia may express itself in different ways at different stages of development. When the specific writing disability involves severe difficulty in handwriting and/or spelling, it is often referred to as dysgraphia.

Students may overcome their reading problems (become compensated dyslexics) only to face persistent spelling and written-expression problems. However, other students may have no difficulty learning to read but have severe difficulty with handwriting, spelling, and/or written composition. The diagnosis of dyslexia or dysgraphia should be reserved, however, for those students who do not have general developmental delays (Axis 1), primary language disability (Axis 1), or brain damage (Axis 2), and continue to struggle despite considerable and appropriate instructional intervention. See Berninger (1998d) for further discussion of subtypes of learning disabilities involving written language. The case study we discuss next illustrates the process of differential diagnosis, summarized in Table 7.3, for a student referred for reading problems.

Case Study

Benjamin, who had just completed fifth grade in another state, was referred for evaluation because of increasing frustration with reading since third grade. Mental retardation was ruled out because (a) based on developmental history his motor and language milestones were early, (b) in third grade his Full Scale IQ on the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; The Psychological Corporation, 1991) was 100 (Verbal IQ 105, Performance IQ 95), and (c) his overall behavioral adaptation appeared to be age-appropriate when evaluated in third grade. In third grade the psychologist who assessed him described the boy as having "an even, usually happy
disposition" and as a child who "resists becoming discouraged by difficulties or minor setbacks." He did not have any comorbid Axis 2 medical conditions.

Likewise, primary language disabilities were ruled out on the basis of reported early language milestones and good listening and verbal expression skills. However, despite a recent hearing exam confirming that his hearing was normal, Benjy often told his mother that he cannot hear the sounds she is talking about in words. Thus, the examiner formulated a clinical hypothesis that the boy had phonological processing problems. This hypothesis was tested during the formal testing with a task in which Benjy was asked to repeat polysyllabic pseudowords without a designated syllable, phoneme, or rime unit (the part of a syllable in which the initial phoneme in the onset position is deleted) (Berninger, Cartwright, Yates, Swanson, & Abbott, 1994).

The following information gleaned from the parent interview is relevant to the differential diagnosis for dyslexia. Benjy lived with both biological parents, a younger brother who was learning to read very easily, and a sister who was a toddler. There was a family history of problems in learning to read. His mother had spent a great deal of time helping him with reading over the years, though as a preadolescent he was beginning to be reluctant to accept her help. The problems with reading were obvious to Benjy's teachers and parents from first grade on, but he did not receive any special help with reading at school. Finally, in third grade, the school psychologist evaluated him and documented a severe discrepancy between IQ and achievement in reading and written expression. His math achievement was grade appropriate. Thus, his problems were specific to reading and writing. Although he qualified for special education, the school did not develop an Individual Education Plan (IEP) and instead chose to provide pull-out support services for the regular program. To his mother's knowledge, Benjy never had any systematic phonics instruction at school. Beginning in fourth grade his parents paid for him to get phonics instruction at a private tutoring center.

In the fourth and fifth grades Benjy received all A's and B's on his report card, even though it seemed to his parents that he was having problems in reading (with word recognition, not comprehension), handwriting, spelling, and composition, and was overtly angry and aggressive every day when his mother picked him up from school. On the state's Assessment of Academic Skills, given in fifth grade, Benjy met most of the standards, which all stressed higher-level thinking skills in reading and math. Basic skills in word recognition and spelling were not considered in the standards. Benjy described a typical day in fifth grade in this way: The teacher gave the class ten worksheets and told them they had 1 hour to complete them. The teacher sat at her desk and told the class to do the work and not bother her. According to his mother, these worksheets were in preparation for taking the state's Assessment of Academic Skills. Benjy could not recall a regular spelling program in fifth grade. He thought he took maybe 10 or 12 spelling tests, but there would be a bunch in a row and then none for a long time. Benjy's at-
tention was excellent throughout the formal testing and his mother's rating of his attentional behaviors fell in the normal range.

Formal testing included many of the recommended measures in Table 7.3. Benjy's age-corrected standard scores for word recognition had decreased about 0.5 standard deviation unit since third grade and ranged from 74 (4th percentile) on the Basic Reading subtest of the Wechsler Individual Achievement Test (WIAT; The Psychological Corporation, 1992) to 72 (3rd percentile) on the Word Identification and 74 (4th percentile) on the Word Attack subtests of the Woodcock Reading Mastery Test—Revised (WRMT-R; Woodcock, 1987). He did somewhat better when words were in context, achieving the equivalent of a standard score for age of 80 (9th percentile) on the Gray Oral Reading Test—Third Edition (GOR-3; Wiederholt & Bryant, 1992), and an instructional level in the third- to fifth-grade range on the criterion-referenced Qualitative Reading Assessment (QRI; Leslie & Caldwell, 1990). However, his reading rate was slow—a standard score of 70 (2nd percentile) on GOR-3 rate. His reading comprehension varied depending on whether responses were open-ended (WIAT standard score 82, 12th percentile) or multiple choice (GOR-3 standard score 110, 75th percentile).

To determine if Benjy also had a comorbid writing disability, measures of handwriting, spelling, and composition were given. Residual problems in handwriting automaticity were evident when writing the alphabet in lowercase manuscript from memory. Benjy wrote g for j, could not remember how to make a lowercase k, and transposed the order of m and n. His standard score for age was 80 (9th percentile) on the Wide Range Achievement Test—Third Edition (WRAT-3; Wilkinson, 1993) Spelling subtest and 69 (2nd percentile) on the WIAT Spelling subtest. Written composition was a relative strength on the Woodcock-Johnson—Revised (WI-R); Benjy's standard score for age was 86 (18th percentile) on Writing Samples, which scores content independent of spelling, and 91 (28th percentile) on Writing Fluency. His spelling errors resembled the invented spellings of younger children.

Benjy had significant difficulty with phonological processing of phonemes in polysyllabic words. His speed of naming familiar letters or digits or alternating letters and digits (Wolf et al., 1986; Wolf, 1986) was significantly delayed. Although he did reasonably well on the orthographic measures out of sentence context, when he had to integrate letter information with sentence meaning on a comprehension test that required close attention to letter information, his errors reflected a tendency to reverse letters in word context. He showed strength in verbal working memory, which was grade-appropriate.

Assessment of social-emotional status showed that Benjy had signs of depression, anxiety, aggression, and low self-esteem. These emotional problems appeared to be secondary to his chronic problems with written language, as emotional problems were not evident in third grade and had appeared and escalated as the task demands at school had increased for both reading and writing. The examiner concluded that Benjy had both dyslexia
and dysgraphia. Had he gotten appropriate Tier 1 and Tier 2 assessment for intervention during the first two grades, his problems would likely have been much less severe. Although he did not appear to have received appropriate educational programming at school for his learning disability during the third through fifth grades, he had received explicit code instruction outside of school, some support services at school, and considerable help from his mom. Although he had met most of his state’s standards on the Assessment of Academic Skills and received A’s and B’s on his report card in fifth grade, he had a significant discrepancy in both reading and writing relative to his intellectual ability and low achievement in reading and spelling single words. His reading rate was slow compared to age peers. This result serves as a reminder that some students who are meeting standards on state assessments for accountability of student learning outcome and who receive good grades may have undiagnosed reading and writing disabilities. Moreover, Benjy had shown deficits on the marker language variables that research has shown are related to specific reading and writing disability (e.g., Berninger, 1994; Berninger, Abbott, et al., 2001). Benjy returned to his home state for sixth grade, where he finally received specialized instruction for his reading and writing disabilities. We recommended instruction aimed at all the necessary instructional components for the functional reading and writing systems (see PAL Intervention Guides for a list of commercially available instructional resources, organized by these instructional components, including but not restricted to phonological awareness and the alphabet principle).

**CONCLUDING REMARKS**

Many reading and writing disabilities could be prevented or reduced in severity if a three-tier model of assessment for intervention were implemented in the schools. This model would require a restructuring of how special education is funded, to provide schools with financial resources for all three tiers. In the first tier, students are screened for marker language variables and those at-risk are given early intervention. In the second tier, the regular curriculum is modified and student progress in response to the modified curriculum is monitored. Finally, in the third tier, those students who continue to struggle despite early intervention and curriculum modification are given a thorough assessment for diagnosis and treatment planning. Although not all students with reading or writing problems early in school have learning disabilities, some do. Dyslexia and dysgraphia, which, respectively, are the most severe forms of specific reading and writing disability, do exist (see Berninger, 1998d, 2000), but the learning outcome for students with dyslexia and/or dysgraphia will be much better if schools do not wait until students fail for several years before beginning the process of assessment for inter-
vention. Although biologically based (Olson, Datta, Gayan, & DeFries, in press), dyslexia and dysgraphia are treatable disorders (Abbott & Berninger, 1999; Berninger, 2000; Richards et al., 1999).

Acknowledgments

The research supporting the three-tier model was funded by the National Institute of Child Health and Human Development Grants HD 25858-10 and P5033812-04.

This chapter was based on presentations at (a) the Fall Conference for the Washington State Association of School Psychologists and the British Columbia Association of School Psychologists, British Columbia, November 1998, (b) the annual meeting of the National Association of School Psychologists, Las Vegas, March 1999, and (c) the Summer Workshop for the Washington Association of School Administrators/Directors of Special Education, Wenatchee, Washington, August 1999.

Although the PAL-RW is based on the research of the first author of this chapter, who is also author of the test battery, the last two authors have contributed substantially to its development and standardization.

References


7. Assessment for Reading and Writing Intervention


Virginia W. Berninger et al.


7. Assessment for Reading and Writing Intervention


