**Test of Adolescent Language–3 (TOAL-3)**

The third revision of the Test of Adolescent Language: A Multidimensional Approach to Assessment (Hammill, Brown, Larsen, & Wiederholt, 1994) is a norm-referenced device designed for adolescents between the ages of 12 and 25. It is intended to identify areas of relative strength and weakness, document academic progress, and identify those who might profit from programs of language intervention. Six of the subtests may be administered to groups; two (Speaking/Vocabulary and Speaking/ Grammar) must be administered individually. The TOAL-3 was designed to assess receptive and expressive spoken and written vocabulary (semantics) and grammar (morphology and syntax).

**Subtests**

The TOAL-3 has eight subtests.

**Listening/Vocabulary** In this 35-item picture-vocabulary subtest, the adolescent must select two pictures that relate to the stimulus word read by the examiner. Credit for an item is awarded only if both correct items are selected.

**Listening/Grammar** Each of the 35 items in this subtest contains three sentences that are read aloud to the adolescent, who must select the two sentences that have the same meaning.

**Speaking/Vocabulary** In this subtest, the examiner reads a stimulus word, and the adolescent must say a meaningful sentence that includes appropriate use of the target word. The subtest contains 25 stimulus words.
**Speaking/Grammar** In this 30-item subtest, the examiner reads a sentence to the adolescent, who must then repeat it.

**Reading/Vocabulary** The adolescent is presented with up to 30 items in this subtest. Each item has three stimulus words and a multiple-choice array containing four additional words. The adolescent must select from the array two words that go with the stimulus words.

**Reading/Grammar** The adolescent is presented with up to 25 items, each of which contains five sentences. He or she must read all five sentences and then find the two that mean "almost the same thing."

**Writing/Vocabulary** The adolescent is required to read a stimulus word and to write a meaningful sentence using that word. The stimulus word must be used correctly in the exact form in which it is given. The subtest contains 30 items.

**Writing/Grammar** Each of the 30 items in this subtest contains two to six sentences of varying complexity. The adolescent is instructed to combine the sentences into one. The simple sentences prompt grammatically more complex constructions. This subtest is analogous to the Sentence Combining subtest of the Test of Language Development, Intermediate 3 (TOLD-I3), but requires written rather than verbal responses.

**Scores**

Several types of scores are available. The eight subtest scores can be transformed into standard scores (mean = 10; standard deviation = 3). Standard scores (mean = 100; standard deviation = 15) are also available for each of the following 11 composite scores:
1. Listening (Listening/Vocabulary and Listening/Grammar)
2. Speaking (Speaking/Vocabulary and Speaking/Grammar)
3. Reading (Reading/Vocabulary and Reading/Grammar)
4. Writing (Writing/Vocabulary and Writing/Grammar)
5. Spoken Language (Listening/Vocabulary, Listening/Grammar, Speaking/Vocabulary, and Speaking/Grammar)
6. Written Language (Reading/Vocabulary, Reading/Grammar, Writing/Vocabulary, and Writing/Grammar)
7. Vocabulary (Listening/Vocabulary, Speaking/Vocabulary, Reading/Vocabulary, and Writing/Vocabulary)
8. Grammar (Listening/Grammar, Speaking/Grammar, Reading/Grammar, and Writing/Grammar)
9. Receptive Language (Listening/Vocabulary, Listening/Grammar, Reading/Vocabulary, and Reading/Grammar)
10. Expressive Language (Speaking/Vocabulary, Speaking/Grammar, Writing/Vocabulary, and Writing/Grammar)
11. General Language

Norms

The TOAL-3 was normed on a total of 3,056 adolescents between 12 and 25 years of age, selected from 26 states. Of this total, 1,512 were from the original TOAL sample, 957 were added for the TOAL-2, and 587 were added for the TOAL-3. All the subjects in the TOAL-3 group were from 18–0 to 24–11 years old and were added to extend the TOAL-
3 norms upward from the 18-year level on the TOAL-2. Because each version of the TOAL has included some modifications (for example, the TOAL-3 includes extensive modifications of the Writing/Grammar subtest), combining norms from three different versions may not be appropriate.

The manual includes sample characteristics of the normative group, but these data are reported for two-year intervals, and so the information is not consistent with the scoring norms, and it is impossible to determine whether adequate sample sizes were included in the scoring norms. For example, a total of 319 thirteen-year-olds were included in the normative sample, but the authors do not indicate how many were in the 13–0 to 13–6 and 13–7 to 13–11 age ranges that are included in the scoring norms. Tables in the TOAL-3 manual show that the normative sample corresponds closely to the overall population of adolescents at the time of the 1990 census. There is no more than a 3 percent difference between the TOAL-3 sample and the population with respect to gender, residence (urban/rural), race, ethnicity, and geographic region (four U.S. regions).

Reliability

Three types of reliability data are presented. First, coefficient alpha was computed for each subtest, each composite, and the total score for each age group. Again, these data are not presented for the age groups for the reported standard scores; rather, data are given for ages 12, 13, 14, 15, 16, 17/18, 19/20, 21/22, and 23/24 for each subtest. Of the 72 age-by-subtest coefficients (9 age groups ¥ 8 subtests), 22 are less than .90, but none is
lower than .80. All 99 of the composite-by-age coefficients (9 age groups x 11 composites) exceed .90.

Second, two stability coefficients were computed. The first was on 52 adolescents attending different grades in a parochial school in Kansas City, Missouri. These adolescents took TOAL, not the revised version (TOAL-3), and 19 of the subjects were below the normative age range. In this sample, coefficients for three subtests (Listening/Vocabulary, Listening/Grammar, and Speaking/ Grammar) were greater than .70 but less than .80. On four subtests (Speaking/Vocabulary, Reading/Vocabulary, Reading/Grammar, and Writing/ Grammar), coefficients were greater than .80 but less than .90. On one (Writing/Vocabulary), the coefficient was .90. The listening and speaking composites were greater than .80 but less than .90, whereas the remaining composites exceeded .90.

The second measure of stability was completed on 59 college students from Austin, Texas, who were retested following a two-week interval. These college students ranged in age from 19 to 24 years. The results indicated that stability coefficients for two subtests (Reading/Grammar and Writing/ Grammar) were greater than .70 but less than .80, and those for the remaining subtests fell between .80 and .90. The Listening, Speaking, Reading, and Writing composites all fell between .80 and .90, and the remaining composites were greater than .90.

Finally, reliability data are presented on interscorer agreement of six raters on the three subtests that use subjective scoring. For Writing/Vocabulary, the correlations between raters ranged from .70 to .95; for Speaking/Vocabulary, correlations ranged from .86 to .99; and for Writing/Grammar, they ranged from .91 to .99. Calculations of the
percentage of interscorer agreement, based on the same data, yielded different results. Only the Speaking/Vocabulary subtest attained a minimum of 90 percent agreement among all raters. Because several items presented difficulties in scoring, the authors revised the criteria for scoring those items. Unfortunately, the revised scoring criteria were not empirically tested, and the revised items were not identified. Also, these data appear to have been gathered on an earlier version of the TOAL.

**Validity**

The authors provide a discussion of the selection of formats for subtests. Content validity is discussed in terms of the procedures and theoretical rationale for test construction. No empirical studies of content validity are presented. It would be useful to see the results of survey data from professionals, as were presented for the related tests in the series (TOLD-P:3 and TOLD-I:3).

Evidence of criterion-related validity for the first edition of the test is presented. Moderate correlations are reported between TOAL and the PPVT, a subtest of the Detroit Tests of Learning Aptitude, the reading and language totals from the Comprehensive Tests of Basic Skills (CTBS), the total score from the Test of Written Language, and the Test of Language Competence. Because the TOAL, TOAL-2, and TOAL-3 are very highly correlated, the authors assume that these correlations are for TOAL-3 as well as for TOAL and TOAL-2. Additionally, the authors note that the TOAL correlated with intelligence and that students previously identified as mentally retarded or learning disabled attained lower scores. No data are provided to indicate that the TOAL-3 is
sufficiently sensitive to monitor a student's progress, and no data are provided to
demonstrate that the TOAL-3 identifies students who might profit from programs of
language intervention (predictive validity.)

Summary

TOAL-3 is a norm-referenced device that assesses three aspects of language (semantics,
morphology, and syntax) via both the receptive and the expressive channels. Expressive
and receptive skills are sampled using both oral and written modes of communication.
The composite scores have good internal consistency, and the reported stability and
interscorer reliability are adequate, although it is difficult to interpret these data because
different versions of the test are pooled in the normative sample, and the standard score
groups are not consistent with normative sample information. Evidence of criterion-
related validity is presented, as well. Despite the limitations noted, the TOAL-3 appears
to be a useful instrument and is widely used.