The Comprehensive Receptive and Expressive Vocabulary Test–Second Edition (CREVT-2) is a norm-referenced instrument designed to assess receptive vocabulary for individuals aged 4–0 through 89–11 and expressive vocabulary for those aged 5–0 through 89–11. CREVT-2 includes a combination of items from the original CREVT, which was designed for children, and the CREVT-A, which was developed for adults. It combines receptive and expressive vocabulary assessment into one instrument; these two communication channels have traditionally been assessed using two instruments, such as the Peabody Picture Vocabulary Test–Revised (PPVT-R) and the Expressive One-Word Picture Vocabulary Test–Revised (EOWPVT-R). High-quality photographs are included in the stimulus materials. Two forms of the test (A and B) are available, and the test typically requires 20 to 30 minutes to administer. The authors state four purposes of the test: (1) identifying deficiencies in oral vocabulary, (2) identifying discrepancies in receptive and expressive vocabulary, (3) documenting progress in instructional programs, and (4) serving as a research tool.

Receptive Vocabulary

In this 76-item subtest, the individual points to one of six pictures in response to the verbal presentation of words by the examiner. Unlike the PPVT-R or the Test of Auditory Comprehension of Language–Revised (TACL-R), which require one response per page
CREVT-2 requires multiple responses on each plate. The 76 items are represented on ten plates (pages).

**Expressive Vocabulary**

This 29-item subtest uses a format similar to that of the Test of Language Development 2, Primary (TOLD-P2). The child is required to tell the examiner what a word means in response to a verbal presentation of the word.

**Scores**

The CREVT-2 yields standard scores with a mean of 100 and a standard deviation of 15 for each subtest. A composite score (general vocabulary) is generated by combining the two subtests and yields a transformed standard score with identical parameters (mean = 100; standard deviation = 15). Tables for converting these scores to percentile ranks and age equivalents are also given in the manual.

**Norms**

Normative data are provided for receptive vocabulary at six-month intervals from ages 4–0 to 12–11, at 1-year intervals from 13–0 through 17–11, and at 10–20-year intervals for the remaining ages. Normative data for expressive vocabulary begin at 5–0, are provided at six-month intervals for 5–0 to 10–11, at 1-year intervals for 11–0 to 12–11, at 2-year
intervals for 13–0 to 16–11, at 1-year intervals for 17–0 to 17–11, and at 10–20-year intervals for the remaining years. The adequacy of the normative sample is difficult to determine because of the peculiar set of procedures used in standardizing the test. The authors combined norms for the CREVT (gathered in 1993) with norms for an adult version of the CREVT (CREVT-A) that were gathered in 1996, and added 365 students tested on the CREVT-2 in 2000. It is important to note that the majority of the norm group was not tested with the actual CREVT-2. The 1993 and 1996 samples were never administered the CREVT-2. Instead, assumptions were made about how they might score on sections they were never tested on. Overall, this sample included 2,545 persons from 38 states. The sample appears representative of the U.S. population as reported by the U.S. Census Bureau (1999) for gender, geographical region, race, ethnicity, family income, and educational attainment of parents. The sample is stratified according to five age intervals (elementary, secondary, young adult, middle-aged adult, and elderly adult) for each of these variables. Although each age group appears to contain similar proportions of examinees with the given characteristics as the population, it is important to note that the age groups cover large intervals of ages, and so it is not clear whether the characteristics of individuals at more specific ages were representative of the population. Additional cross-tabulations across other participant characteristics (such as race by geographic region and gender by ethnicity) were not included in the manual. Thus, the characteristics of the comparison group are not entirely clear. Finally, a slightly higher percentage of students with disabilities were represented in the sample than are included in the U.S. population.
Reliability

Coefficient alphas were computed for Forms A and B for a variety of age groups on each subtest, as well as for the composite score. In contrast with the standard scores, coefficients are reported at 1-year intervals for ages 4–17, at 12-year intervals for ages 18–29, and at 20-year intervals thereafter. This yields a total of 104 alpha coefficients (17 age groups with two forms across two subtests and one composite, with two additional coefficients for age 4). With one exception (Expressive Vocabulary Form A for age 5), all exceed .80, with 73 at or above .90. Composite coefficients are above .90 for all age levels, except for the age interval of 70–89 years. Also, subtest and composite coefficient alphas are presented for individuals according to gender, ethnicity, and various disabilities (learning disabilities, speech–language disorders, and mental retardation). Although these were pooled across age levels, so that it is difficult to determine what proportion of subjects was at each age level, these coefficients are routinely high (ranging from .91 to .99). Alternate-forms reliability was assessed by computing correlations at each age level for test Forms A and B in the entire normative sample. Individuals were administered one form immediately after the other form. For Receptive Vocabulary, alternate-forms correlations ranged from .78 to .97, with a median correlation of .94. For Expressive Vocabulary, correlations ranged from .76 to .96, with a median of .88. Alternate-forms correlations for General Vocabulary ranged from .86 to .98, with a median of .94.

Test–retest reliability was assessed by readministering the instrument to 27 kindergarten children after two months, and to 50 high school students and 48 adults
(ages 20–59) after two weeks. All groups were selected from the Austin, Texas, area.

Coefficients are reported for each group for both forms (A and B) of the expressive and receptive subtests and for the composite (general vocabulary). These correlations ranged from .91 to .95 for the kindergarten sample, from .93 to .99 for the high school sample, and from .93 to .99 for the adult sample.

Validity

The manual includes sections on content, concurrent, and construct validity. Content validity was evaluated qualitatively by simply reporting the rationale for selecting the test format and the procedures for selecting the pool of items that ultimately were included in the test. Item analyses were also conducted to eliminate items with poor discrimination. Differential item functioning analyses were conducted to determine whether items were biased with respect to gender or ethnicity. Fifty-one significantly biased items were identified (one which had moderate to large bias). However, the authors argue that this is a small proportion of the 630 comparisons made, and therefore conclude that the test is within acceptable levels of gender, racial, and ethnic bias.

Criterion-related validity was evaluated by correlating the scores on the CREVT-2 (including each subtest and form) with a variety of other standardized tests that measure spoken language and vocabulary. These correlations were conducted primarily among students with disabilities (learning disabilities, language–speech impairments, and mental retardation). A number of correlation coefficients would suggest a high degree of concurrent validity: .70 (both forms of CREVT-2) with the PPVT-R, .75 (CREVT-2
Form A) and .72 (CREVT-2 Form B) with the EOWPVT-R, .69 (CREVT-2 Form A) and .63 (CREVT-2 Form B) with the vocabulary subtest of the Wechsler Intelligence Scale for Children–III, .89 (CREVT-2 Form A) and .90 (CREVT-2 Form B) with the Clinical Evaluation of Language Function–Revised, and .90 (CREVT-2 Form A) and .92 (CREVT-2 Form B) with the TOLD-P2. However, some correlations for the two forms of CREVT-2 were inconsistent. For example, Expressive Vocabulary correlations with the PPVT-R were .61 for Form A and .42 for Form B, and with the EOWPVT-R were .57 for Form A and .39 for Form B.

Evidence of construct validity is also included in the manual. The authors argue that construct validity is demonstrated by the fact that CREVT-2 scores follow the pattern anticipated by age; scores increase rapidly in the early years, and increase more slowly in later years. The authors also present information on how the CREVT-2 helps to differentiate groups of students with learning disabilities, speech–language disorders, and mental retardation from groups of students without disabilities. Although these groups performed somewhat lower than students without disabilities on the CREVT-2, they were still found to perform in the average range, with the exception of students with mental retardation, who received lower scores. It is not clear, therefore, whether the CREVT-2 would be useful in distinguishing students in need of language support.

**Summary**

The CREVT-2 is an individually administered receptive and expressive vocabulary test. It represents a merging of the original CREVT and the CREVT-A. The CREVT-2
normative data were taken from these original tests, with the addition of 365 students tested on the CREVT-2; therefore, it is questionable whether this normative sample allows for appropriate comparisons. A relative strength of this test is that it can be administered in a fairly short amount of time across a wide range of ages. Although adequate reliability data are presented, the validity information provided is not convincing. Despite these weaknesses, the CREVT-2 is potentially useful as a quickly administered vocabulary screening instrument, to be followed up with more detailed vocabulary assessment as needed.