Original TOWL published in 1978 and modified in 1983.Recommended Administrator: The TOWL–4 is a level B test and the examiner's manual recommended that examiners have "some formal training in assessment", such as a college or university course. Examiners also need to have a strong knowledge of English language usage, conventions, grammar and the proper construction of sentences and compositions to be able to score the subtests (recommended college graduate level of proficiency). Amount/Type of training needed: The TOWL–4 manual states that individuals with the qualifications stated above should be able to competently administer the subtests after familiarizing themselves with the specific administration instructions provided and practicing scoring for the spontaneous writing subtests. Reliability: The examiner's manual presents evidence to support the reliability of the instrument. Coefficient alpha values are provided for each subtest at different ages and grades. These showed Internal Consistency for most subtests (.72–.96), and for composite scores (.82–.96). Test–Retest reliability was mostly within acceptable range (93% rounded to .80 or above). Alternate Forms of composite scores (Immediate: .94 & .95) showed internal consistency but Alternate Forms Delayed (.73–.94) was not proven to be reliable in the Spontaneous Writing composite score. Interscorer Differences fell within acceptable ranges of (.80–.99). The information and evidence presented shows that the TOWL–4 has reasonable internal consistency and test–retest reliability, with the exception of a few subtest scores in the .71–.75 range, and the Spontaneous Writing composite score. Improvements in reliability over the years are clearly delineated in the manual. Validity: The TOWL–4 Examiner's Manual Evidence provided support for the content validity by providing a detailed description of the test format and content of the test. A rationale for the specific content of each subtest was provided and a rationale for changing the content form previous versions was included in the Item Analysis. The manual also provides a discussion on the Criterion–Prediction Validity and showed a comparison with 3 other tests of literacy, the WLOS, the ROS and the TORC–4. The results showed Level 1–A criterion of acceptability in two out of three comparison studies, supporting the validity. Standardization: The TOWL–4 was normed on a sample of 2,205 students in 17 states that was collected between 2006 and 2007. The authors chose four major standardization sites as representative of the regions of the United States. The demographic characteristics of the sample included "geographic region, gender, race, ethnicity, household income, education level of parents, and disabling condition" (Hammill & Larsen, p. 56), and were matched with the 2007 U. S. Census.Subtests one through five are called "contrived" and include Vocabulary (using a word in a sentence), Spelling (writing dictated sentences with an emphasis on correct spelling), Punctuation (writing dictated sentences with an emphasis on correct punctuation), Logical Sentences (editing illogical sentences) and Sentence Combining (integrating two or more sentences into one grammatically correct sentence). Subtests six and seven are called "spontaneous", and are based on a story written by the subject in 15 minutes; subtest six, Contextual Conventions, measures orthographic and grammatical conventions.These individual scores are then combined into three composite scores, Contrived Writing, Spontaneous Writing and Overall Writing.