Test of Written Language-Third Edition
(TOWL-3)

Publisher/Date:
• PRO-ED, Inc., 8700 Shoal Creek Blvd., Austin, TX 78757-6897. Published, 1996.

Purpose:
• Individually- or group-administered comprehensive test of writing skills for students ages 7-00 to 17-11. The test includes two parallel-forms (Forms A & B).

Provides:
• Composite scores include Contrived Writing, Spontaneous Writing, and an Overall Writing estimate. These composites estimates are based on eight individual subtests (Vocabulary, Spelling, Style, Logical Sentences, Sentence Combining, Contextual Conventions, Contextual Language, and Story Construction). Contrived writing taps responses to pre-determined stimuli and assesses sentence construction and combining, vocabulary, spelling, and conventions. Spontaneous writing taps responses from student-generated samples and assesses conventions, language, and story or theme construction. Comparing the two assesses the students understanding of conventions of writing as well as their functional applied skills.

Standardization Issues:
• 2,217 students were tested in 1995, with sample demographics matching 1990 Census data by geographic region, gender, race, urban-rural residence, disability, SES, and ethnicity.

Reliability and Validity Issues:
• Despite the inherent issues of the subjectivity involved in evaluating writing affecting reliability, the test’s coefficients appear within acceptable. Alternate-form reliability, test-retest reliability, and inter-rater reliabilities appear around .70 (minimal standards). Overall, strong construct and content validity evidence is reported. Criterion validity (comparing the test to teacher rating scales and the Comprehensive Scales of Student Abilities) was good. Item bias analysis yielded very good results across all subtests and paired groups (male/female, white/nonwhite, Hispanic/non-Hispanic).

Additional Points:
• The norms were based on 1990 Census data and may becoming outdated.
• Use with the younger ages should be approached more cautiously. The sample sizes for each age group vary between 105 (age 7) to 350 (age 10), with the former possibly being only marginally adequate. Despite the addition of easier items for younger students with academic difficulty, the test may still be too difficult for that age group. The median discriminative validity findings for ages 7 & 8 were reportedly unacceptably-low. Finally, raw scores of “0” still result in scaled scores which may be problematic when interpreting performance—especially at young ages (i.e., a 7-year old with a raw score of “0” in Sentence Combining still earns a scaled score of “9”).

• Because of the subjectivity involved in evaluating and scoring written responses, examiners may consider scoring with a knowledgeable partner. In particular, the Story Construction and Contextual Conventions subtests may be most subjective, employing rating scales asking for such various scoring judgments as “poor,” “average,” “good,” “immature,” or “interesting.” The subjectivity involved can contribute to error.

• The test provides a strong conceptual model of writing, and its analysis of writing using both contrived and student-generated responding provides a broader assessment of functional writing skills.