

# Assessment critique

Science



Assessment Critique Sandra Whitney California State University, Northridge  
SPED 501 M/M Dr. Haney A. General Information The Kaufman Test of Educational Achievement, Second Edition (KTEA-II) is an individually administered measure of academic achievement for ages 4.5 through 25. The test is available in 2 versions. The Brief Form assesses achievement in reading, math and written expression. The Comprehensive Form covers reading, math, written language, and oral language. It also provides an analysis of students' errors.

Examiners can obtain a Comprehensive Achievement Composite in about 30 minutes for younger children and 85 minutes for the oldest students. The Comprehensive Form has 2 independent, parallel forms (A and B). The KTEA-II was written by Alan and Nadeen Kaufman and is published by AGS Publishing. B. Brief Description of Test Scoring & Types of Scores Derived  
The KTEA-II was designed to measure student progress. Some of its applications include assessing achievement, identifying processes, analyzing errors, program planning, measuring academic progress, evaluating interventions/programs, and making placement decisions.

After reviewing the Manual, I believe the KTEA-II would be a good measure of academic achievement and student progress. The KTEA-II's authors examined literature reviews and recommendations from experts in different subject areas in order to define which skills should be measured in each achievement domain. Three national tryouts of the KTEA-II Comprehensive Form Materials were conducted between 2000 and 2001. These trials illustrated whether each subtest had enough items to be reliable and provided adequate coverage of skills at each grade level.

They also allowed for statistical analysis to identify and modify/remove items that had poor discrimination or were differentially difficult according to sex or ethnicity. Finally, the tryouts provided valuable information regarding item difficulties that was necessary for constructing standardization forms that would be parallel in content and level of examinee performance. I believe the KTEA-II is well designed. I especially like the fact that it provides a Clinical Analysis of Errors and that the authors utilized input from experts when designing/selecting test items.

The analysis of errors can help a teacher identify specific areas in which the student demonstrates weak, average, or strong skill development. I feel the KTEA-II's design and norms make it suitable for most populations between the ages of 4.5 and 25. As a special educator, a real positive feature is the inclusion of examinees with special classification or diagnosis. However, I do not feel the KTEA-II is suitable for English Language Learners. The manual specifically states that the test was normed to represent the US population of children and young adults who speak English.

C: Validity, Normative Population Data, & Types of Scores Derived The norm sample consisted of 3,000 examinees aged 4 through 25. The grade norms are based on 2,400 of the examinees in Grades K-12. The standardization took place from September 2001 through May 2003. All age levels had between 100 and 200 participants, except age 19, which had 80. The KTEA-II sample was based on the 2001 Current population Survey and designed to match the US population with regards to sex, parent education, ethnicity, and educational status of examinees aged 18 to 25.

The sample was representative in terms of geographic region, with a few exceptions at a couple of age levels. Examinees with special disability classification or diagnosis were also included in the standardization sample. These participants had a specific learning disability, speech/language impairment, attention deficit/hyperactivity disorder, mental retardation, emotional/behavioral disturbance or were gifted and talented. One shortcoming in the norms is the failure to provide a breakdown of rural/urban participants.

For internal consistency, the overall Comprehensive Achievement composite coefficient was very reliable at (.97). The core composites for Reading (.96), Mathematics (.96), and Written Language (.93) are also highly reliable. However, the Oral Language composite (.87) and Oral Fluency (.85) fall below the desired (.90) standard for reliability. The Sound-Symbol and Decoding composites are adequately reliable at all age levels. Because of the format for the subtests for the Reading Fluency composite, it is not possible to evaluate the internal consistency.

The internal consistency coefficients are lower for subtests than composites. Most of the Reading and Mathematics subtests, and the Spelling subtest coefficients are sufficiently reliable. The majority of coefficients for the Oral Language subtests and the Written Expression subtest are less than (.90). The coefficients for Nonsense Word Decoding are acceptable; but the majority of coefficients for the Phonological Awareness, Associational Fluency, and Naming Facility are below (.90).

To assess the stability of the KTEA-II scores over a period of weeks, the test was administered twice to 221 children from three grade ranges (Pre-K to

<https://assignbuster.com/assessment-critique/>

Grade 1, Grades 2 through 6, and Grades 7 through 12). The retest interval ranged from 11 to 60 days and averaged 3? to 4 weeks. Alternate-form reliability was also examined in this analysis because about half the students took Form A first and Form B second; the other half took the test in the opposite order. The reliability correlations for the three grade ranges for the Comprehensive Achievement composite were (. 92), (. 94), and (. 5), respectively. For Pre-K to Grade 1, only the overall Reading and Decoding composites are sufficiently reliable. Coefficients for the Mathematics (. 87), Written Language (. 85), Oral Language (. 64), Sound-Symbol (. 84) and Oral Fluency (. 59) composites are all below (. 90). Letter & Word Recognition is the only subtest for Pre-K to Grade 1 with adequate reliability (. 97). Coefficients for the rest of the subtests range from (. 47) to (. 88). For Grades 2 through 6 the Mathematics, Written Language, Reading Fluency, and Decoding Composites all had coefficients of at least (. 0). The Reading (. 87), Oral Language (. 68), Sound-Symbol (. 80), and Oral Fluency (. 67) composites are less than . 90. All subtest correlations are less than (. 90), except Spelling, Nonsense Word Decoding, and Decoding Fluency. Coefficients for the Reading (. 89), Oral Language (. 81), and Oral Fluency (. 76) composites are below . 90 for Grades 7 through 12. Correlations for the Mathematics, Written Language, Reading Fluency, and Decoding composites are all adequate. All subtest correlations, except Math Computation, are less than (. 90).

Outside reviewers note that because stability and alternate-form reliability were not separated in this analysis, it is impossible to know whether results for some components are unstable, whether the forms differ, or both. The

Oral Language composite is problematic because of its internal consistency and stability correlations are consistently below (.90). Interrater reliability was evaluated for Written Expression, Oral Expression, Reading Comprehension, Listening Comprehension, and Associational Fluency because they require judgment in scoring and are most susceptible to difference in scoring among examiners.

The cases used 50 students at each of two grade levels. Students from Grade 2 or 3 completed Form A and students from Grade 8 completed Form B. Three or four examiners scored each level of each subtest. Correlations were all above (.90), except Oral Expression at both grade levels (.82 and .88) and Associational Fluency at Grade 2 (.82). The authors took many steps to ensure the validity of items on the KTEA-II. These efforts included literature reviews, consultation with experts in the field, and field testing.

Intercorrelation of subtests and composites are provided at each age and grade level and address construct validity. Moderate to high correlations were found between the majority of subtests and composites, except for the Oral Language domain. The average correlation between Oral Expression and Listening Comprehension was (.46). Low correlations for these subtests and composite suggest they are measuring skills not closely related to other sections of the test. Factor analysis was used for the eight primary subtests of the KTEA-II Comprehensive Form, using the entire age-norm sample for Grade 1 through age 25.

Confirmatory factor analysis provided evidence for a four-factor model (math, reading, written language, and oral language), as this model had good fit statistics and high loadings on the factors for all subtests. To

evaluate concurrent validity, the KTEA-II Comprehensive Form was administered along with one or more achievement or cognitive abilities tests. Administration of the two tests occurred in counterbalanced order, with approximately half of the cases taking the KTEA-II first and the other half taking it second.

Administration of the two tests could occur on the same day or separated by as much as 60 days. When compared to the original Kaufman Test of Educational Achievement (K-TEA), the Wechsler Individual Achievement Test-Second Edition (WIAT-II), the Woodcock-Johnson Tests of Achievement-Third Edition (WJIII ACH), and the Peabody Individual Achievement Test-Revised, Normative Update (PIAT-R/NU), high overall composite correlations were found (range .84 to .94). At the composite and subtest level, moderate to high correlations were generally found for the domains of reading, mathematics, and written language.

However, the Oral Language composite correlations were mixed, with one as low as (.08). When the KTEA-II was compared to the Oral and Written Language Scales (OWLS), a (.75) correlation was found between the written expression subtests. Correlations in the (.40's) were found between the oral expression and listening comprehension subtests for the two measures. The relationship between the KTEA-II and several intelligence tests was also examined. Composites from the KTEA-II correlate in the low to moderately high range (from .13 to .4) with the Kaufman Assessment Battery for Children-Second Edition (KABC-II;), the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler), the Woodcock-Johnson Tests of Cognitive Abilities-Third Edition (WJ III COG). Students diagnosed with

reading, mathematics, and writing learning disabilities; cognitive impairment; attention deficit/hyperactivity disorder; and emotional or behavioral disturbance and those who were deaf or hard of hearing had mean scores below average on all subtests and composites. Children identified as high performing or talented received mean scores above average for all subtests and composites.

In terms of validity, comprehensive evidence is provided in support of the test's content validity. The addition of an Oral Language section to the KTEA-II seems beneficial, but is an area of particular concern in terms of both reliability and validity. I believe the Oral Language results should be interpreted cautiously. Outside reviewers caution that there is a danger of overestimating or underestimating a student's performance due to steep item gradients on the KTEA-II. When tests have steep item gradients, a 1-point change in raw scores can result in a large change in standard scores when using the KTEA-II.

In some cases, a 1-point change in a raw score results in a change of as much as 13 standard score points. Reviewers have also found that although the norms for the KTEA-II begin at 4 to 6, most subtests do not have adequate floors at this age. Instruments without adequate floors do not have enough easy items to discriminate between students with and without skill deficits. Several concerns exist in regard to the adequacy of some KTEA-II subtest floors; thus, it is suggested that examiners check floor adequacy when assessing younger children. Using a subtest with an inadequate floor may overestimate performance at certain ages.