

Early Childhood Measurement and Evaluation Tool Review

Early Childhood Measurement and Evaluation (ECME), a portfolio within CUP, produces *Early Childhood Measurement Tool Reviews* as a resource for those who conduct screening, assessment, and evaluation. To learn more about ECME and CUP, provide feedback, or to access additional reviews, visit our website at www.cup.ualberta.ca or email us at cup@ualberta.ca

Date of review: May 2012

Stanford-Binet Intelligence Scales – Fifth Edition (SB5)

Measurement Areas:

The Stanford-Binet Intelligence Scales – Fifth Edition (SB5) is designed to test intelligence and cognitive abilities. This test can be used for ages 2 years 0 months to 89 years 11 months. The scale provides a Full Scale IQ (FSIQ), Non Verbal (NV) and Verbal (V) domain scores as well as 5 factor scores:

1. Fluid Reasoning (FR)
2. Knowledge (KN)
3. Quantitative Reasoning (QR)
4. Visual-Spatial Processing (VS)
5. Working Memory (WM)

Purpose:

The SB5 is a norm-referenced assessment tool that can be used to:

- measure general cognitive functioning,
- identify developmental disabilities and exceptionalities in children, adolescents, and adults (within cautionary guidelines listed by authors),
- provide clinical and neuropsychological assessment,
- research abilities,
- assess early childhood,
- conduct a psycho-educational evaluation for special education placements,
- provide information for adult social security evaluations,
- support worker's compensation evaluations,
- inform career assessments, and
- guide treatment program development.

Length and Structure:

The SB5 is designed to be administered on an individual basis with children, adolescents, and adults. The full test takes approximately 45-75 minutes to administer. The abbreviated test takes approximately 15-20 minutes.

Administration of the assessment begins with two routing subtests that identify an appropriate developmental starting point for the examinee. Separate routing subtests are conducted to determine appropriate starting points for both verbal and non-verbal domains. After the two routing subtests, all other subtests are grouped into testlets. The testlets are arranged into levels of difficulty, with 5 levels in the verbal domain and 6 levels in the non-verbal domain. Both routing subtests and most of the testlets have example items to assist the examinee to understand each task.

In addition to converting raw scores to standard scores for each composite score, subtest scaled scores, percentile ranks, confidence intervals, age equivalents, and change-sensitive scores (CSS) can all be computed from the data.

Materials:

The SB5 is classified as a “Level C” qualification that is targeted to institutions with personnel possessing masters and doctorates of psychology or education, and/or have licensure in a relevant area of assessment with provincial or national organizations.

The SB5 complete kit is sold by the publisher for USD \$1,087.00 and includes: 3 item books, examiners manual, technical manual, child card, layout card, manipulatives kit and storage box, 25 record forms, and a carrying case. Scoring software (SB5 Scoring Pro, Version 1.2) is available for USD \$255.00 and an interpretive manual is also available for USD \$121.00.

Accessibility:

The SB5 is available in the English language using USA norms. The administration manual contains cautions for testing individuals who are English language learners or English as a second language. They recommend that a measure of English language proficiency is administered prior to the SB5 to ensure adequate English language knowledge. This caution is particularly crucial to the verbal items.

Considerations and cautions are also noted for those with special needs, or who are deaf or hard of hearing, have communication disorders, orthopedic impairments, or motor skills deficits.

Administration, Scoring, and Interpretation:

The manual suggests that only those individuals with formal graduate-level or professional training in psychological assessment should interpret test results using the 7 step method described in the manual. The manual contains a section on basic score interpretation, including an example of an interpretive report. An interpretation manual is available from the publisher.

Subscales:

The SB5 contains 5-factor indexes: Fluid Reasoning (FR), Knowledge (KN), Quantitative Reasoning (QR), Visual-Spatial Processing (VS), and Working Memory (WM). Subtests are combined into

either one of the two domains or one of the 5-factor indexes. Either the two domains or the 5-factor indexes combine to form a composite Full Scale IQ (FSIQ) score.

Documentation:

The SB5 Administration and Scoring Manual provides specific procedures for administration and scoring. The Technical and Interpretive Manual has comprehensive chapters on interpretation, test standardization, norm development, validity, and reliability.

Standardized Sample:

The SB5 was normed and standardized using an American sample of 4800 individuals (51% Female, 49% Male) ages 2 to 85+ years. The sample was stratified according to demographic variables such as age, sex, race/ethnicity, socioeconomic level, and geographic region. Demographic information is based on the 2001 United States Census bureau data. According to the manual, persons classified as being American Indian and Alaskan Native were included in the standardization sample but for the purposes of tabulation they were included in the category of “other”. Included in the category of “other” were those who indicated their race/ethnicity as Native Hawaiian, Pacific Islander, or those that indicated more than one race. The entire “other” sample makes up 2.7% of the standardization sample.

Reliability:

The SB5 Technical Manual discusses four types of reliability measures:

Internal Consistency: According to the technical manual, reliability coefficients (Cronbach’s alpha) for the Full Scale IQ scores were in the extremely high range (.97 to .98). Verbal and Nonverbal IQ scores were also found to have coefficients in the high range (average of .95 to .96). All factor indexes had reliability coefficients above .90. The Abbreviated IQ score was found to have average reliability coefficients of .91.

Test-retest Reliability: Four separate samples in four age groups (2 to 5, 6 to 20, 21 to 59, and 60 and over) were administered the SB5 on two different occasions. Subtest reliabilities ranged from .63 to .93, while test-retest reliability coefficients for the factor indexes were .79 to .95. Correlations for the Verbal and Nonverbal domains were strong, with a range of .89 to .95. Stability coefficients for the FSIQ were very strong ranging from .93 to .95 among the various age groups.

Standard Error of Measurement and Confidence Intervals: The standard error of measurement (SEM) is calculated from test reliability and the standard deviation of scores. SEM is “the standard deviation of the theoretical distribution of differences between the true score of the individual and his or her observed score across repeated measures of the same individual” (Roid, 2003). The most important use of SEM is to construct confidence intervals where the degree to which measurement error should contribute to an individual’s observed score is taken into consideration. For example, a FSIQ score of 100 has a confidence interval of 96-104 in order to capture potential measurement error.

Interscorer Agreement: Interscorer Agreement is measured by the degree to which separate examiners agree on the scoring of examinee responses. During standardization, items with poor

interscorer agreement were deleted before publication. Among published items, agreement correlations ranged from .74 to .97.

Validity:

An extensive discussion of validity is contained in the SB5 Technical Manual. Validity will be discussed in terms of content-related validity evidence, criterion-related validity evidence, and construct-related validity.

Content Validity: Content-related validity evidence for the SB5 is determined by three methods:

1. *Professional Judgment of Content:* Researchers, examiners, and experts in assessment were consulted in order to provide feedback on items and subtests. Various data collection procedures were conducted wherein experts provided comments and concerns with items. Also, both female and male content reviewers of various ethnic and religious backgrounds reviewed items and brought forth items that may be problematic.
2. *Coverage of Important Constructs:* Experts were asked to examine the items within the Cattell-Horn-Carroll (CHC) theory of intellectual abilities. The experts were asked to sort the items into CHC factor or factors. Factor analysis provided evidence for factor structure of the SB5.
3. *Item analysis:* According to the SB5 Technical Manual “Extensive item analyses, including classical and item response theory, were conducted on SB5 items” (Roid, 2003, p.79). Analyses included item discrimination, percentage correct by age level, model-data-fit statistics, and differential item functioning.

Criterion-Related Evidence of Validity: Criterion related validity is determined by demonstrating the scale is related to other scales that measure the same construct, societal outcome measures, and special group membership (i.e. gifted). Two types of evidence are used to demonstrate criterion-related validity:

1. *Concurrent validity:* In a study, 104 individuals were presented with the SB5 and the SB4 in a counter-balanced design. The study found a high correlation between the two measures for the FSIQ (.90). Other studies between the SB5 and other measures of cognitive abilities found similarly high results: SB L-M (.85), the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R; .83); the Wechsler Intelligence Scales for Children-Third Edition (WISC-III, .84); the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III, .82), and the Woodcock-Johnson III Tests of Cognitive Abilities (.78).
2. *Predictive Validity:* The importance of predictive validity is that it demonstrates how well an instrument can predict academic achievement as traditionally Intelligence is said to be related to achievement. Correlations between the SB5 IQ and factor scores and the Woodcock-Johnson III (WJ III) achievement scores are in the range of .50 to .84. In a study including 80 children presented in the technical manual, correlations between the SB5 and the Wechsler Individual Achievement Test-Second Edition (WIAT-II) are in the range of .61 to .83. These results demonstrate that the SB5 yields results that are able to adequately predict achievement.

Special Populations: Another important facet of cognitive ability tests is the presence of distinctive score profiles for those in special populations (i.e. giftedness, mental retardation, developmental delay, autism, limited English competency, speech/language disorders, learning disabilities, attention-deficit disorder, severe emotional disturbance, and motor difficulties). The technical manual describes some possible patterns in mean scores that the various groups may demonstrate on the SB5.

Construct-Related Validity Evidence: According to a review by D’Amato, Johnson, and Kush (2003), “all SB5 subtests, across all ages, demonstrated average principal component loading of greater than .70 on the g, or general factor, indicating that each subtest was a good measure of g. The proportion of SB5 variance accounted for by the g factor ranged from 56% to 61% depending on the factoring method.”

The Technical Manual also provides confirmatory factor analysis (CFA) data to provide evidence for construct related validity. One to five factor models were tested and indicated that the five factor model provided the best fit. However, D’Amato, Johnson, and Kush (2003) note that CFA was performed on split half subtests, otherwise there would have been no model to be tested and indicate that the model fit data presented in the manual may be up to desired standards.

Additional Reliability and Validity Studies

- Pomplun & Custer (2005) investigated validity evidence for the two SB5 measures of working memory. They specifically investigated construct validity and concurrent validity between the SB5 working memory and memory tests, non-verbal and verbal tests, and achievement measures. In terms of construct validity, item mapping demonstrated that as items increased in difficulty so did the demands on verbal and non-verbal working memory. It was also found that the SB5 verbal working memory scales correlated highly with other measures of verbal memory measures and had lower correlations with nonverbal measures. Conversely, the SB5 non-verbal scale correlated strongly with other measures non-verbal memory and less strongly with measures of verbal memory.

Additionally, the researchers found that verbal working memory correlated more highly with reading skills, while non-verbal working memory with mathematics skills. The authors note that this is consistent with previous findings, and thus the results support the differences between ways in which individuals process verbal and non-verbal information.

- Another study by Marusiak & Janzen (2006) investigated working memory abilities in ADHD children as compared to a control group of typically developing children. They found that for ADHD children working memory was the lowest factor score and was significantly lower than other factor scores. In addition, ADHD children were found to have lower non-verbal working memory. As children with ADHD are usually hypothesized to have working memory difficulties, the results provide some additional support for the validity of the working memory factor on the SB5.
- DiStefano and Dombrowski (2006) investigated the theoretical structure of the SB5, using the data provided from SB5 standardization. They used both exploratory and confirmatory factor analysis. Contrary to published information, using exploratory factor analyses

suggested that there were two factors (verbal and non-verbal) apparent for those younger than 10 years old. However, for the older age groups, only a one factor general intelligence model fit the data appropriately. Similar findings were found with confirmatory factor analyses. In addition, factors on the SB5 were found to have higher than expected correlations (.89 to .98) which may indicate that the factors are not necessarily measuring different factors of intelligence but rather more general intelligence. The authors of this study conclude that given the results of this study, the SB5 can be regarded as a good measure of general intelligence and is an option for early childhood assessment.

Publication Information:

This review is based on the 5th edition of the Stanford-Binet Intelligence Scales published in 2003 by Riverside Publishing.

References:

- DiStefano, C., & Dombrowski, S. C. (2006). Investigating the theoretical structure of the Stanford-Binet-Fifth-Edition. *Journal of Psychoeducational Assessment, 24*, 123-136.
- Johnson, J.A., & D'Amato, R.C. (2004). Review of the Stanford-Binet Intelligence Scales- Fifth Edition. *Mental Measurements Yearbook, 16*.
- Kush, J. C. (2004). Review of the Stanford-Binet Intelligence Scales- Fifth Edition. *Mental Measurements Yearbook, 16*.
- Marusiak, C.W., & Janzen, H.L. (2005). Assessing the working memory abilities of ADHD children using the Stanford-Binet Intelligence Scales, Fifth Edition. *Canadian Journal of School Psychology, 20*, 84-97.
- Pomplun M., & Custer. M. (2005). The construct validity of the Stanford-Binet 5 measures of working memory. *Assessment, 12*, 338-346.
- Roid, G. H. (2003). *Stanford- Binet Intelligence Scales, Fifth Edition, Examiner's Manual*, Itasca, IL: Riverside Publishing.
- Roid, G. H. (2003). *Stanford -Binet Intelligence Scales, Fifth Edition, Technical Manual*, Itasca, IL: Riverside Publishing.
- Roid, G. H. (2003). SB5 Scoring Pro [Computer software]. *Stanford-Binet Intelligence Scales, Fifth Edition*, Itasca, IL: Riverside Publishing.
- Roid, G. H. (2005). *Stanford Binet Intelligence Scales, Fifth Edition, Interpretive Manual*, Itasca, IL: Riverside Publishing.

How to cite this document: This document was created for CUP. However to cite this document use the following:

Community-University Partnership for the Study of Children, Youth, and Families (2011). *Review of the Stanford-Binet Intelligence Scales – Fifth Edition (SB5)*. Edmonton, Alberta, Canada.

