Journal of Psychoeducational Assessment

Use of Brief Intelligence Tests in the Identification of Giftedness
Eric E. Pierson, Lydia M. Kilmer, Barbara A. Rothlisberg and David E. McIntosh

The online version of this article can be found at:

Published by:

$SAGE

Additional services and information for *Journal of Psychoeducational Assessment* can be found at:

Email Alerts: http://jpa.sagepub.com/cgi/alerts

Subscriptions: http://jpa.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://jpa.sagepub.com/content/30/1/10.refs.html

>> Version of Record - Feb 9, 2012

OnlineFirst Version of Record - Dec 5, 2011

What is This?

Use of Brief Intelligence Tests in the Identification of Giftedness

**Eric E. Pierson[1], Lydia M. Kilmer[1],
Barbara A. Rothlisberg[1], and David E. McIntosh[1]**

Abstract

Schools often administer brief intelligence tests as the first step in the identification of students who are cognitively gifted. However, brief measures are often used without consideration of underlying constructs or the psychometric properties of the measures and without regard to the links between screening decisions and educational programming. This article provides an overview of these issues and offers recommendations for using brief intelligence measures particularly when screening children who are cognitively gifted.

Keywords

high ability, gifted, brief IQ

The identification and classification of high-ability and gifted individuals should reflect standards and goals of the program for which the assessment is being conducted. Public and private schools in the United States are often the primary systems interested in identifying and classifying high-ability school-aged children (Cross & Coleman, 2005). The importance and utility of traditional cognitive ability or IQ measures for the classification and identification of students for high-ability programs has long been recognized by researchers as a helpful starting place (Brody & Stanley, 2005; Terman, 1922). Contemporary models of giftedness such as Sternberg's WICS (Wisdom, Intelligence, Creativity, Synthesized) model (Sternberg, 2005a, 2005b) and Renzulli's conceptualizations of giftedness (Renzulli, 2005) recognize the value of traditional intelligence measures among a wide range of abilities that may be important in differentiating gifted students from others, particularly in the schools. However, both have written that older cognitive measures are outdated. Current models of high ability and giftedness incorporate other qualities in addition to intelligence, such as task commitment (Renzulli, 2003, 2005), creative ability (Cross & Coleman, 2005), and wisdom (Sternberg, 2005a, 2005b) that are typically not measured as well by IQ tests but may be evaluated by other means (Pfeiffer & Blei, 2008; Sternberg, 2005a). One of the few tools with substantial evidence to support its use in the identification and classification of individuals with high ability historically has been traditional cognitive measures. Although

[1]Ball State University, Munice, IN

Corresponding Author:
Eric E. Pierson, Department of Educational Psychology, Teachers College 512, Ball State University, Munice, IN 47306, USA
Email: eepierson@bsu.edu

contemporary accounts of gifted evaluation discredit the use of full-score IQ (FSIQ) measures, early users of IQ measures focused on the full-score performance rather than the identification of specific skill areas for classification of gifted individuals. As a result, early attempts at classification and the development of brief intelligence measures focused on this score as well.

Brief Intelligence Measures

Researchers of high-ability individuals have long desired shorter methods for assessment than the traditional full IQ test (Terman, 1922). One method of doing this has been the use of shortened or abbreviated intelligence tests. A shortened or brief intelligence test is a cognitive measure with fewer items, scales, and/or time of administration than typically required in the administration of the full test battery. The shortened or brief intelligence test is designed to provide a comparable estimate of ability and confidence interval of measured performance to that of the full battery.

The use of a single subtest or two from an established intelligence test to estimate an individual's level of performance has a long history in the field of psychology and neuropsychology (Lezak, Howieson, & Loring, 2004). In part, this is due to the early recognition that some subtest scores are more robust than others after a brain injury. For instance, Wechsler (1958) advocated the use of the vocabulary, information, object assembly, and picture completion subtests as measures that would be resilient to aging effects on the brain. These subtests were also shown to be the best predictors of educational success, particularly the vocabulary subtest (Heaton, Ryan, Grant, & Matthews, 1996). As a consequence, valuable information on cognitive performance could be gleaned from a limited sample of behavior instead of demanding the investment that administration of a comprehensive test would require.

Silverstein (1967, 1985) was one of the first to advocate for and develop short forms of intelligence tests, including brief forms of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967). His short forms, comprised of the Vocabulary-Block Design (V-BD) dyad and Vocabulary-Arithmetic-Block Design-Picture Arrangement (V-A-BD-PA) tetrad, were carefully researched and hence brought a new level of sophistication to short intelligence test development (Kaufman & Kaufman, 2001). The use of Silverstein's tetrad was suggested for both the WISC-R (Kaufman, 1976) and the WAIS (Doppelt, 1956).

Silverstein's contribution resulted in a large increase in the number of abbreviated forms that were developed. Because of the popularity of Wechsler Scales, it is not surprising that they were the source for many adaptations. The WAIS-R (Wechsler, 1981; also Boone, 1992; McCusker, 1994; Nagle & Bell, 1995), the *Wechsler Adult Intelligence Scale-Third Edition* (WAIS-III; Wechsler, 1997; see also Axelrod, Dingell, Ryan, & Ward, 2000), the Wechsler Preschool and Primary Scale of Intelligence-Revised (WIPPSI-R; Wechsler, 1989; also Tsushima, 1994), the *Wechsler Intelligence Scales for Children-Third Edition* (WISC-III; Wechsler, 1991a, 1991b), and the *Wechsler Intelligence Scales for Children-Fourth Edition* (WISC-IV; Wechsler, 2003; also Watkins, Wilson, Kotz, Carbone, & Babula, 2006) have all had a number of brief forms developed based on subtest groupings. Brief measures also have appeared based on the *Stanford–Binet Intelligence Scale-Fourth Edition* (Thorndike, Hagen, & Sattler, 1986), the *Stanford–Binet Intelligence Scales-Fifth Edition* (Roid, 2003a; Canivez, 2008), and the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). In most cases, the brief forms included two to four subtests; however, some brief forms (e.g., Ryan & Ward, 1999) are composed of seven subtests or incorporate shortened versions of the original tests (McPherson, Buckwalter, Tingus, Betz, & Back, 2000; Satz & Mogel, 1962). In recent times, researchers have begun to consider the Global Ability Index of the WISC-IV as a shortened version of the WISC-IV. The

GAI (General Ability Index) has three verbal and three performance that are heavily loaded on *g* subtests (Rowe, Kingsley, & Thompson, 2010).

Ideally, brief measures efficiently predict the range of performance one would expect from a comprehensive test of intelligence. To maximize this possibility, the subtests selected for inclusion in brief forms are often those with the highest *g* loading in the battery (e.g., Wechsler Abbreviated Scales of Intelligence [WASI], Wechsler, 1999, or the Brief Intellectual Ability [BIA] of the *Woodcock–Johnson Tests of Cognitive Abilities-Third Edition-Normative Update*; Woodcock, McGrew, Schrank, & Mather, 2007). As noted, brief measures were constructed to limit testing time. Because all early intelligence tests and test developers (Terman, 1922) had a briefer literature upon which to draw, the emphasis in test development was often on an ability to differentiate between groups. Over time these early measures have come to be viewed as having been largely atheoretical. The subtests making up early brief forms tended to relate strongest to the overall score of the comprehensive cognitive measure based on clinical experience or early statistical studies. The idea that particular cognitive factors or underlying abilities were involved in performance was only important to the degree that it helped understand better predictors of overall ability or academic achievement. However, as theories of intelligence became more refined and useful for explaining student behavior and the ability tests themselves were constructed to align with them, the selection of subtests for brief measures became more critical to capturing the components of the underlying view of ability. For example, the Brief Intellectual Ability test uses the Cattell-Horn-Carroll (CHC) model to select subtests of the *Woodcock–Johnson Tests of Cognitive Abilities-Third Edition-Normative Update* (Woodcock et al., 2007) to sample the factors, *Gf, Gc, and Gs,* and obtain an accurate estimate of *overall ability.*

Although brief cognitive measures have traditionally been composed of subtests that were highly predictive of general (*g*) intelligence in the original comprehensive measure, more recent short forms have reflected the importance of considering construct validity in subtest selection. For example, short forms of tests constructed to assess abilities according to the CHC theory of intelligence may include the factors of comprehension-knowledge (*Gc*), visual-spatial thinking (*Gv*), and fluid reasoning (*Gf*), because of their stronger relationship to *g* than other factors and their ability to predict educational performance. This broader view of assessment, taking into account different aspects of ability, addresses the concerns of researchers who argue against the limitations of a global or single-factor view of cognitive behavior. For example, the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) and its contemporary editions have been used previously to estimate intelligence because of its strong psychometric properties, high correlations with comprehensive batteries, and the ability to discriminate between important clinical populations (Duncan & Duncan, 1997). However, instruments such as the PPVT assess only a narrow facet of cognitive skill and may underestimate the performance of minority students among others (Kaufman & Kaufman, 2001). In view of the developments in theories of intelligence and the importance of recognizing multiple factors affecting intelligent behavior, practitioners and researchers need to consider the quality of their assessment data for educational decision making, even when brief forms of tests are involved.

Many researchers on giftedness have argued against the use of IQ measures because of a perception that they only measure narrow, school-based abilities (Cross & Coleman, 2005; Renzulli, 2005; Sternberg, 2005). These abilities are often closely tied to current ideas of a facet of intelligence *Gc*. However, the added features of what many of them talk about creativity, wisdom, problem solving, and speed of learning, to name a few, are recognized as part of *Gf*. Today, most test batteries include information related to which subtests measure *Gf, Gc,* or a combination of both. Further information about how contemporary theories of giftedness may be measured by *Gf* and *Gc* can be seen in Table 1. It is not our view that all aspects of these unique theories of giftedness are adequately measured by our current measures of intelligence.

Table 1. Cattell-Horn-Carroll (CHC), Giftedness Theory, and Brief IQ Measure

	Giftedness theory	Brief IQ	Subtest
CHC ability			
Gg	All	All	
Gf	Creative (Cross & Coleman)	KBIT-2 WASI	Matrices Matrix
	Intelligence and creativity (Sternberg)		Reasoning
	Creativity (Renzulli)		
	Brody and Stanley		Block design WJ-III concept formation
Gc	School (Cross & Coleman)	KBIT-2	Riddles
	Schoolhouse giftedness (Renzulli)	WASI	Vocabulary similarities
	Academic intelligence (Sternberg)	WJ-III	Verbal comprehension

Note: CHC = Cattell-Horn-Carroll; KBIT-2 = Kaufman Brief Intelligence Test-2nd Edition; WASI = Wechsler Abbreviated Scale of Intelligence; WJ-III = The Woodcock–Johnson Test of Cognitive Abilities. This table was created on the basis of a review of materials in O'Donnell (2009); Kaufman and Kaufman (2004); McGrew, Schrank, and Woodcock (2007); Cross and Coleman (2005); Renzulli (2005); Sternberg (2005); and Brody and Stanley (2005).

Examples of Brief Measures of Intelligence

Given the state of the art of evaluation of ability, professionals considering the use of a short form for screening or identification purposes have several options available, including those that have been formally published. This section will overview the characteristics of some published forms used in school-based settings. Tests such as the Stanford–Binet and the Differential Abilities Scale have suggested abbreviated forms nested within their available protocols and manuals.

The Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990) was one of the first short forms tied to a measure designed to assess intelligence based on theory—the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). Constructed to measure Luria's theory of brain organization, the K-ABC and its derivative, the KBIT, attempted to assess both verbal and nonverbal skills. The revised *Kaufman Brief Intelligence Test-Second Edition* (KBIT-2; Kaufman & Kaufman, 2004b), resulting from the *Kaufman Assessment Battery for Children-Second Edition* (K-ABC-II; Kaufman & Kaufman, 2004a), is now grounded in CHC theory, but it can also be interpreted according to Luria's neuropsychological model in cases where the child displays a diverse cultural or educational background (Kaufman, Kaufman, Kaufman-Singer, & Kaufman, 2005). Both versions of the KBIT have three subtests— two verbal (verbal knowledge and riddles) and one nonverbal subtest (matrices). After examining the three subtests separately, the KBIT-2 was standardized on a sample of 2,120 children using random sampling methods to reflect demographic data from the 2001 U.S. Census (Kaufman & Kaufman, 2004b). The lowest split-half coefficient for different age groups was .78; the rest fell in the .80s and .90s.

Validity for the use of the KBIT to estimate cognitive ability has been determined through a variety of methods. This includes cross-battery comparisons with the WISC-III in different populations of students (Canivez, 1996; Grados & Russo-Garcia, 1999; Thompson, Browne, Schmidt, & Boer, 1997). The generally positive results indicate that the KBIT was a good predictor of WISC-III scores and measured the common underlying construct of ability. Bain and Jaspers (2010) noted in their review of the KBIT-2 that previous studies demonstrated moderate correlations between the KBIT and the WRAT3 (Wide Range Achievement Test-3; Wilkinson, 1993).

They also reported that the KBIT-2 may underestimate the scores of gifted students aged between 7 and 16 for the nonverbal portion of the measure.

After many years of various researchers and clinicians using shortened forms of the Wechsler scales on the basis of altered standardization processes, the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999) was published. Intended for a broader age range than its age-specific predecessors, the WASI is normed for those aged between 6 and 89 years. This age range may make the WASI a more appealing instrument for the identification of exceptional students than short forms of the WISC-IV because the addition of harder items for measures of knowledge and reasoning as well as visual reasoning and problem solving reduce problems with subtest ceiling effects. In addition, the extended age range (6 to 89) makes it appealing for those assessing adolescents where the overlap of age levels in different versions of tests can make battery selection difficult.

The WASI is comprised of four subtests: vocabulary, similarities, matrix reasoning, and block design (Wechsler, 1999). Vocabulary and similarities make up a verbal IQ (VIQ), whereas matrix reasoning and block design make up the performance IQ (PIQ). Full-scale scores can be derived from all four subtests (FSIQ-4) or only from vocabulary and matrix reasoning (FSIQ-2). Three of the four subtests require the individual to provide open-ended responses or develop a structure in space as opposed to selecting from a multiple-choice format. This can provide insight into individual characteristics and problem-solving behaviors.

Normed on a nationally representative sample of 2,245 individuals, the two-subtest version of the WASI received reliability coefficients of .93 for children and .96 for adults, whereas the four-subtest version received reliability coefficients of .96 for children and .98 for adults (Wechsler, 1999). When compared with WAIS-III FSIQ scores, the FSIQ-4 was found to correlate at .92 and the FSIQ-2 at .87 (Axelrod, 2002). Several studies have found support for the construct validity of the WASI (Canivez, Konold, Collins, & Wilson, 2009; Ryan et al., 2003). Ryan and Brown (2005) reported that the reliability of the WASI was satisfactory. It should be noted that both the WASI and the KBIT-2 have subtests unique to them that are not part of their parent batteries. This added dimension is especially important to researchers who may be concerned with the prediction problems on test correlations previously discussed by Kaufman and Kaufman (2004b). The second edition of the WASI has an anticipated release date in the fall of 2011. At the time of this writing, little information pertaining to its development or standardization is available. However, given the history of the Wechsler Series and the high quality of typical normative samples, it is anticipated that the WASI-II will have standardization data set consistent with the demographic trends of the most recent census estimates and adequate sample sizes at all ages.

It would appear then that greater attention is now being given to the makeup of abbreviated forms of intelligence not only to improve their predictive validity but also to offer insight into factors related directly to educational performance. This awareness of test utility is pertinent in cases where short forms screen or identify giftedness because improved testing techniques may well maximize the information obtained for the students assessed.

Strengths of Brief Forms

Both schools and clinicians may have seen the administration of a comprehensive measure of intelligence as prohibitively expensive, particularly for the identification of gifted and high-ability individuals. Many school districts may choose to administer a brief measure as the first step in identifying children who may qualify for special education or gifted services (McIntosh & Dixon, 2005). Also, it is not uncommon for schools to include a brief measure, instead of including a comprehensive measure of intelligence, as one of the key components when identifying children for gifted programs. Schools also may select brief measures of intelligence for reasons of convenience

and efficiency as opposed to consideration of the test's psychometric properties. When this is the case, schools may favor a test like the KBIT-2 (Kaufman & Kaufman, 2004b) to other options because examiners can administer it with lower (Bain & Jaspers, 2004) levels of training in cognitive assessment than is possible with instruments such as the WASI. As a consequence, decisions to include a brief measure of intelligence likely are based on the need to save time, with less thought given to the constructs being assessed, predictive validity, or selection criteria of the brief form.

Limitations of Using Short Forms

Although short forms have the benefit of time savings, Thompson, Howard, and Anderson (1986) provided evidence that altering the order of subtest administration, as is done in a shortened form, may reduce the ability of the short form to accurately estimate the individual's full-scale score. Kaufman and Kaufman (2001) also questioned the validity of short forms that use only one out of two or one out of three items such as in the Satz and Mogel format to reduce test length. They argued that this changes the test-taking procedure by increasing the slope of item difficulty. Changes in the structure of the original test to create the abbreviated form may compromise the quality of the performance observed as well as the data obtained in the evaluation. In particular, the greater distance in difficulty between adjacent items on a subtest may make it that much more difficult for clinicians to discern where problems begin to emerge in regards to the difficulty of items. It is interesting to note that examiners using the current version of the Stanford-Binet (Roid, 2003) may experience a similar sense when testing children on different testlets.1

The current literature regarding the measurement of intelligence in samples of gifted and high-ability students has several limitations. Most brief and full-battery cognitive measures do not include information regarding the specificity or sensitivity of tests in classifying individuals for gifted programs or extensive information regarding the criteria for selection, goals of the program, or program philosophy. As a result it is difficult to gauge whether mean scores obtained in classification samples reflect the psychometric properties of the test or the selection criteria and philosophy of the program.

Related to this problem is the reporting of means and standard deviations of persons in gifted samples on measures of cognitive ability. Although this practice is consistent with the broader practice in psychology and education of reporting means and standard deviations, it ignores the basic principles as to why and when means are appropriate. Other measures of central tendency such as interpolated medians may be more helpful to report alongside of means and standard deviations. This would be particularly helpful given the fact that these samples are oftentimes likely to be skewed in distribution (Gagné, 2004).

Identification of Giftedness

Keeping in mind the definitions of giftedness brought forward by legislation such as NCLB (No Child Left Behind Act; U.S. Department of Education, 2002), it follows logically that a general complaint with using intelligence tests and FSIQ measures as the sole measure of giftedness is that the definition of giftedness must contain more than general intellectual ability. Thus, even if brief intelligence tests are used only in screening for gifted children, those children who are gifted in other areas of exceptionality (leadership or art) will be missed. In fact, most researchers are moving away from using FSIQ as their exclusive measure of giftedness (Feldman, 2003; Sousa, 1995). Some have suggested that such exclusive use of intelligence measures blind evaluators to other areas of intelligence such as leadership or the arts (Jarosewich, Pfeiffer, & Morris, 2002; Sternberg, 2005a). Others point out that the use of cutoff scores deprives children

just below the cutoff of education they should receive (McIntosh & Dixon, 2005). Global IQ scores also overlook the fact that many children, even those already classified as gifted, exhibit giftedness in specific domains rather than in all areas (Matthews, 1997).

The strategy of using brief measures of ability runs counter to one of the goals of many gifted programs: the identification and selection of a broad and diverse group of high-ability or gifted students (Coleman & Cross, 2005). Programs that focus on the use of brief measures of intelligence for the identification and screening of cognitive abilities are likely to identify a narrower and more homogenous group of individuals with similar cognitive strengths and abilities. An added concern with the use of a brief IQ measure relates to the manner in which estimates of cognitive ability for the full scale are estimated. Although scores for abbreviated batteries may correlate highly with full-scale scores of the comprehensive battery, this does not guarantee that the abbreviated battery will have the same range of performance or ceiling as the comprehensive instrument. The decision to develop extended normative information for the *Wechsler Intelligence Scale for Children-Fourth Edition* helps to demonstrate this fact (Zhu, Cayton, Weiss, & Gabel, 2008). An extended normative table was developed to expand the range of the full-scale scores for gifted students. In order to do this, individual subtests needed to have normative ranges recalibrated to allow for scaled scores beyond the standard 19. Another strategy for adding greater ranges and ceilings to the estimate of full-scale abilities would be to add additional subtests that measure a wider range of cognitive abilities.

Defining giftedness as only an IQ score may deny children who are members of a cultural minority the educational services that their Caucasian classmates receive. The original use of intelligence tests as a standard requirement for admittance into gifted programs put those students who were not proficient in English at a disadvantage (Sternberg, 2000). In fact, minority and impoverished students have been chronically underrepresented in gifted and talented programs (Baldwin, 2004; U.S. Department of Education, 1993). The U.S. Department of Education report, *National Excellence: A Case for Developing America's Talent* (Ross, 1993), called not only for an expanded definition of giftedness but greater support for underrepresented populations in gifted and talented programs. In response, some researchers have begun to attempt to adapt intelligence tests to ethnic minority or international students (e.g., Malda, van de Vijver, Srinivasan, Transler, & Sukumar, 2009; Moon, McLean, & Kaufman, 2003), with some success. Ford and Webb (1994) called for more holistic, multidimensional, and multimodal forms of assessment, citing this as the best way for students of ethnic minority status to achieve entry into gifted programs. However, this challenge is far from being overcome and will continue to face educators in the years to come (Kaufman, Evans, & Kaufman, 2010).

Although the use of cognitive measures has always been a key part of the identification of giftedness, researchers as far back as Terman (1922) recognized the value that parent and teacher ratings brought to the identification process. In an effort to develop giftedness evaluation scales that are more multidimensional than traditional intelligence tests, a number of teacher-rating scales have been developed (Pfeifer, 2009). In each of these scales, the teacher rates the student on a variety of characteristics. Examples of these scales are the Gifted and Talented Evaluation Scales (GATES; Gilliam, Carpenter, & Christensen, 1996), the *Gifted Evaluation Scale-Second Edition* (McCarney & Anderson, 1989), and the Scales for Rating the Behavioral Characteristics of Superior Students (Renzulli et al., 1997), as well as the Gifted Rating Scales (Pfeiffer & Jarosewich, 2003). In evaluating the quality of these scales, Jarosewich et al. (2002) examined the standardization, reliability, and validity of each. The authors found technical flaws in all three scales, ranging from poor standardization techniques to lack of information regarding predictive validity. These measures will, therefore, require continued refining, but offer options of broader-based evaluation of competencies to gifted educators that are separate from intelligence measures.

Factors to Consider When Choosing Brief Measures

Predictive Validity

The degrees to which abbreviated or brief measures of intelligence estimate a full-scale measure have largely been discussed relative to their utility. These included the high degree of intercorrelations that will be seen between the brief and composite measure and the effects of introducing an alteration of test order in administration (Kaufman & Kaufman, 2004b; Thompson et al., 1986). Additional concerns arise when assessing individuals at the extreme tail of the bell curve, as is the case in identifying exceptionally able students (McIntosh, Dixon, & Pierson, in press; Ziegler & Ziegler, 2009). In more specific terms, individuals who are administered measures of cognitive ability and who are gifted will tend to score lower than their actual level of innate measurement as a function of measurement error. So children who are tested every year or 2 years as part of an ongoing identification strategy may have years where measurement error incorrectly classifies them as *not gifted*. This last concern is a primary argument against the use of procedures involving the averaging the repeated measurements of cognitive abilities or the need for multiple-stage evaluation procedures in the identification of gifted children where they need to pass each level. This view may seem counterintuitive to those familiar to the advice given by Gallagher (1994) and Gridley, Norman, Rizza, and Decker, (2003), who advocate the use of multiple tests in assessment. It is critical for parents and teachers to understand that performances may dip below cutoff scores on similar measures of intelligence such as repeated testing on a different cognitive measure as a function of the limitations of our measures and not students or programs.

As noted earlier, the very fact that the measures are brief forms raises concerns about their validity. By definition, when a brief measure is used the observed sample of behavior will be shorter and provides the individual fewer opportunities to demonstrate superior performance. This may not be a primary concern for those researchers interested in profiles or psychometric questions regarding giftedness, but it may be to professionals involved in psychological service delivery or educational planning. Gridley and colleagues (2003) provided an outline of characteristics of intellectually gifted and talented students who may be observed during the testing session.

Cutoff Scores and Their Recommended Uses

Local schools and educational agencies often have differing criteria and methods for the identification of giftedness. As a result, programs in areas with lower population density may have a lower cutoff or minimum obtained score that they have selected as necessary to enter the only gifted and talented program in an area. In contrast, programs in urban settings or with large catchment areas involving state math and science academies may have much higher criteria for entry. As a result, gifted and talented programs in low-population-density areas may serve and identify a more diverse group of gifted students because they use lower entry criteria for a variety of measures. Those in the high-density or large catchment groups may instead offer services to a more focused and extremely talented group.

Because of this variability in identification procedure, a wide range of possible cutpoints or criteria can be applied to connote giftedness. The widespread adoption and recognition by researchers of the utility of the CHC theory of intelligence advocates its use in the screening and identification of gifted students. Most of the major test instruments available have guides on how to interpret test data according to the CHC theory. The use of the multiple strata of intelligence found in the theory allows for the identification of a richer, more diverse group of gifted students

than would have been possible through screening procedures or identification procedures using full-scale scores only. In addition, it is suggested by Gridley et al. (2003) and McIntosh and Dixon (2005) that decisions regarding eligibility involve the use of the confidence intervals of scores to better capture the scope of the student's performance.

For practical purposes, the choice of cutoff score point defining gifted performance will often be tied to the size of the desired program in an educational setting, but several factors should be considered when determining what the point or criteria should be. For example, the Flynn effect (Flynn, 1984) will cause some students with lower levels of ability to "float" into programs over time. Measurement and the tendency for test measurement to move toward the mean will force students who are gifted to score lower on selected tests at different points in time depending on the measures used and the match to their particular pattern of strengths. Multiple methods and multiple informants may improve the quality and ecological validity of such high-stakes decision making.

Use of Brief Measures With Minorities When Identifying Giftedness

The issue of diversity in the assessment and identification of giftedness is one of primary importance for societal as well as educational need. There is a long history of research that has identified the manner in which traditional measures of cognitive ability and intelligence can and have been misapplied when working with minority students (Jacob, Decker, & Hartshorne, 2010). Terman (1922) recognized that in the identification of gifted students certain ethnic groups were being selected at greater-than-expected rates whereas others were being selected at lower-than-expected rates.

Modern measures of cognitive assessment strive to minimize differences between ethnic and gender groups on measures and oftentimes view differences on items as measurement artifacts that should be minimized or removed. Contemporary batteries of cognitive assessment frequently include sample responses for words from Spanish as recognition of the growing need to have that information available to an examiner (Wendling, Mather, & Schrank, 2009). Cross and Coleman (2005) also point to the ability of objective measures of intelligence to be useful in helping to identify students who might otherwise be overlooked. As a result, short measures may introduce perspectives to interpretation of performance that might minimize bias toward the majority culture. For instance, the KBIT-2 (Kaufman & Kaufman, 2004b) can be interpreted using Luria's model of neuropsychological development that deemphasizes the crystallized or culturally derived component of ability. Another brief measure that has a relatively large normative matched to the demographics of the United States and available for ages 4:0 to 21:11 is the Wechsler Nonverbal Scale (WNV). The WNV may be of particular benefit in the evaluation of individuals who are English language learners (extensive review of the WNV can be found in Naglieri & Brunnert, 2009). Preliminary validity studies reported in the manual use categories such as good or better ability to discriminate between gifted and nongifted sample of students as other, more comprehensive, measures of intelligence. More research with diverse samples using both the KBIT-2 and the WNV is needed.

Construct Validity of Brief Measures

The current generation of brief measures of cognitive ability has increased potential and clinical as well as educational utility for the evaluation of intelligence in individuals. It is anticipated that as time and the science of psychology continues to develop so will the ability to predict and discriminate between groups and utilize time and personnel optimally. Unlike previous generations of brief IQ measures, the current generation for the most part is well grounded in theory

(albeit not gifted theory). Test developers have taken a keen eye to the changing demographics of the U.S. population and are working to incorporate better tests to serve diverse populations. Although previous generations of tests may have been atheoretical or designed to support competing models of intelligence, the preeminent theory driving their development today is the CHC model. It is unfortunate that tests based on the CHC model have not resolved all issues associated with evaluating a gifted population. In more specific terms, no current battery has shown a mean value above 125 when test publishers validated the measure. Linking all of our tests and measures to one theory ties all of our measures to its limitations and reduces our conceptualization of intelligence beyond CHC.

List of Recommendations

The authors encourage those interested in integrating brief measures into the evaluation and selection of gifted individuals to review the work of Newton, McIntosh, Williams, and Youman (2008) and the highly regarded chapter by Thompson and Morris (2008). In brief, it is recommended that programs work in accordance with the guidelines of the American Psychological Association (2002) and the National Association of School Psychologists (2010) codes of ethics and consider all high-stakes decision making within the context of the error ranges of the instruments being employed. Similarly, schools and psychologists should choose measures that are clearly linked to gifted and intelligence theory. Tests, whether brief or full batteries, should be chosen based on adequate normative information for the sample that matches the target client group being assessed. When using brief cognitive measures, any identification plan should include supplementary measures and techniques to identify other potential areas of giftedness in line with the program's philosophy and objectives.

Careful consideration should also be given to who administers cognitive measures. Schools should choose personnel who are familiar with the personality characteristics of gifted students that may lead to lowered estimates of cognitive ability on IQ measures. Schools should administer brief IQ measures only to those students who are not suspected of having a learning disability or ADHD (attention-deficit hyperactive disorder). Finally, the decision to use one or more cognitive measures in the evaluation and identification of giftedness should take into consideration the likelihood that one or more measures may end up with a lower-than-anticipated obtained score as a function of test theory and not as an innate ability on the part of the individual.

Although some brief measures such as the KBIT-2 may be administered by individuals with less training in psychological testing, we would advise against this practice. Individuals with limited training in testing may not have sufficient knowledge or experience to recognize anxiety, low motivation, cultural variations, or other factors that may lead to an underestimation of ability. We also recommend that the results of brief measures be considered in the context of the individuals cultural experiences, personal history, and psychometric properties of the instrument.

Conclusions

Brief intelligence tests appear to be a reasonable tool allowing for the prediction of full-scale intelligence scores as well as valuable in the identification of giftedness (Newton et al., 2008). The use of these measures should be used with recognition of the limitations that they inherently possess. Brief measures should not be used in isolation, and it should be recognized that measurement error may inaccurately label a student as *gifted* or not if that student's scores fall within the confidence interval of the selected cutoff for program inclusion. Educators and their prospective students may be better served to include a more broad-based, ecological view of gifted characteristics when screening for gifted education candidates.

Table of Federal Legislation

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, 20 U.S.C. § 6301 *et seq.* (West 2003). To close the achievement gap with accountability, flexibility, and choice, so that no child is left behind. (2002).

Note

1. The testlet is a feature of the *Stanford–Binet Intelligence Scales* (5th ed.) designed to replicate the age-level test format found in early editions of the Stanford-Binet. Each subtest has 5 to 6 testlets. The individual's performance across testlets is used to estimate performance on the full subtest.

References

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1060-1073.

Axelrod, B. N. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning. *Assessment, 9*(1), 17-23.

Axelrod, B. N., Dingell, J. D., Ryan, J. J., & Ward, L. C. (2000). Estimation of Wechsler Adult Intelligence Scale—III index scores with the 7-subtest short form in a clinical sample. *Assessment, 7*(2), 157-161.

Bain, S. K., & Jaspers, K. E. (2010). Test review: Review of Kaufman Brief Intelligence Test, Second Edition. *Journal of Psychoeducational Assessment, 28*(2), 167-174.

Baldwin, A. Y. (2004). I'm black but look at me, I am also gifted. In S. M. Reis & A. Y. Baldwin (Eds.), *Culturally diverse and underserved populations of gifted students* (pp. 1-9). Thousand Oaks, CA: Corwin Press.

Boone, D. E. (1992). Reliability of the WAIS-R with psychiatric inpatients. *Journal of Clinical Psychology, 48*(1), 72-76.

Brody, L. E., & Stanley, J. C. (2005). Youths who reason exceptionally well mathematically and/or verbally: Using the MVT:D[4] model to develop their talents. In *Conceptions of giftedness* (2nd ed., pp. 20-37). New York, NY: Cambridge University Press.

Canivez, G. L. (1996). Validity and diagnostic efficiency of the Kaufman Brief Intelligence Test in reevaluating students with learning disability. *Journal of Psychoeducational Assessment, 14*(1), 4-19.

Canivez, G. L. (2008). Orthogonal higher order factor structure of the Stanford–Binet Intelligence Scales—Fifth Edition for Children and Adolescents. *School Psychology Quarterly, 23*(4), 533-541.

Canivez, G. L., Konold, T. R., Collins, J. M., & Wilson, G. (2009). Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity. *School Psychology Quarterly, 24*(4), 252-265.

Cross, T. L., & Coleman, L. J. (2005). School-based conception of giftedness. In *Conceptions of giftedness* (2nd ed., pp. 52-63). New York, NY: Cambridge University Press.

Doppelt, J. E. (1956). Estimating the full scale score on the Wechsler Adult Intelligence Scale from scores on four subjects. *Journal of Consulting Psychology, 20*(1), 62-66.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MD: American Guidance Service.

Feldman, D. H. (2003). A developmental, evolutionary perspective on giftedness. In J. H. Borland (Ed.), *Rethinking gifted education* (pp. 9-33). New York, NY: Teachers College Press.

Flynn, J. R. (1984), The mean IQ of Americans: Massive gains from 1932 to 1978. *The Psychological Bulletin, 95*, 29-51.

Ford, D. Y., & Webb, K. S. (1994). Desegregation of gifted education programs: The impact of brown on underachieving children of color. *Journal of Negro Education, 63*(3), 358-375.

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies, 15*(2), 119-147.

Gallagher, J. J. (1994). *Current and historical thinking on education for gifted and talented students*. Washington, DC: Office of Educational Research and Improvement.

Grados, J. J., & Russo-Garcia, K. A. (1999). Comparison of the Kaufman Brief Intelligence Test and the Wechsler Intelligence Scale for Children—Third Edition in economically disadvantaged African American youth. *Journal of Clinical Psychology, 55*(9), 1063-1071.

Gilliam, J. E., Carpenter, B. O., & Christensen, J. R. (1996). *Gifted and talented evaluation scales*. Waco, TX: Prufrock Press.

Gridley, B. E., Norman, K. A., Rizza, M. G., & Decker, S. L. (2003). Assessment of gifted children with the Woodcock–Johnson III. In F. A. Schrank & D. P. Flanagan (Eds.), *WJ III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 285-317). San Diego, CA: Academic Press.

Heaton, R. K., Ryan, L., Grant, I., & Matthews, C. G. (1996). Demographic influences on neuropsychological test performance. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 141-163). New York, NY: Oxford University Press.

Jacob, S., Decker, D. W., & Hartshorne, T. S. (2010). *Ethics and law for school psychologists* (6th ed.). New York, NY: John WIley.

Jarosewich, T., Pfeiffer, S. I., & Morris, J. (2002). Identifying gifted students using teacher rating scales: A review of existing instruments. *Journal of Psychoeducational Assessment, 20*(4), 322-336.

Kaufman, A. S. (1976). A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities, 9*(3), 33-41.

Kaufman, S. B., Evans, M. L., & Kaufman, J. C. (2010). Gifted education. *Encyclopedia of Cross-Cultural School Psychology*, pp. 480-481.

Kaufman, A. S., & Kaufman, N. L. (1983). *K-ABC: Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Services.

Kaufman, A. S., & Kaufman, N. L. (1990). *K-BIT: Kaufman Brief Intelligence Test Manual*. Circle Pines, MN: American Guidance Service.

Kaufman, J. C., & Kaufman, A. S. (2001). Time for the changing of the guard: A farewell to short forms of intelligence tests. *Journal of Psychoeducational Assessment, 19*, 245-267.

Kaufman, A. S., & Kaufman, N. L. (2004a). *Kaufman Assessment Battery for Children* (2nd ed.). Bloomington, MN: Pearson.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman Brief Intelligence Test* (2nd ed.). Bloomington, MN: Pearson.

Kaufman, J. C., Kaufman, A. S., Kaufman-Singer, J., & Kaufman, N. L. (2005). The Kaufman Assessment Battery for Children-Second Edition and the Kaufman Adolescent and Adult Intelligence Test. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 344-370). New York, NY: Guilford.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.

Malda, M., van de Vijver, F. J. R., Srinivasan, K., Transler, C., & Sukumar, P. (2009). Traveling with cognitive tests: Testing the validity of a KABC-II adaptation in India. *Assessment, 17*(1), 107-115.

Matthews, D. J. (1997). Diversity in domains of development: Research findings and their implications for gifted identification and programming. *Roeper Review, 19*, 172-177.

McCarney, S. B., & Anderson, P. D. (1989). *Gifted Education Scale—Second Edition, technical manual.* Columbia, MO: Hawthorne Educational Services.

McCusker, P. J. (1994). Validation of Kaufman, Ishikuma, and Kaufman-Packer's Wechsler Adult Intelligence Scale—Revised short forms on a clinical sample. *Psychological Assessment, 6*(3), 246-248.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical manual* (Woodcock- Johnson III Normative Update)*.* Rolling meadows, IL: Riverside Publishing

McIntosh, D., & Dixon, F. (2005). Use of intelligence tests in the identification of giftedness. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual theories, tests and issues* (2nd ed., pp. 504-520). New York, NY: Guilford.

McIntosh, D., Dixon, F., & Pierson, E. E. (in press). Use of intelligence tests in the identification of giftedness. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual theories, tests and issues* (3rd ed.). New York, NY: Guilford.

McPherson, S. Buckwalter, J. G., Tingus, K., Betz, B., & Back, C. (2000). The Satz-Mogel short form of the Wechsler Adult Intelligence Scale—Revised short forms on a clinical sample. *Psychological Assessment, 6*(3), 246-248.

Moon, S., McLean, J. E., & Kaufman, A. S. (2003). A cross-cultural validation of the sequential- simultaneous theory of intelligence in children. *School Psychology International, 24*(4), 449-461.

Nagle, R. J., & Bell, N. L. (1995) Clinical utility of Kaufman's "amazingly" short forms of the WAIS-R with educable mentally retarded adolescents. *Journal of Clinical Psychology, 51*(3), 396-400.

Naglieri, J. A., & Brunnert, K. (2009). Wechsler Nonverbal Scale of Ability (WNV). In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement* (pp. 315-338). New York, NY: John Wiley.

National Association of School Psychologists. (2010). *Principles for professional ethics*. Bethesda, MD: Author.

Newton, J. H., McIntosh, D. E., Williams, T., & Youman, E. (2008). Assessing giftedness in children: Comparing the accuracy of three shortened measures of intelligence to the Stanford-Binet Intelligence Scales (5th ed.). *Psychology in the Schools, 45*(6), 523-536.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, 20 U.S.C. § 6301 *et seq.* (West 2003).

O'Donnell, L. (2009). *The Wechsler Intelligence Scale for Children-Fourth Edition*. In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement* (pp. 153-190). New York, NY: John Wiley.

Renzulli, J. S. (2003). Conception of giftedness and it's relationship to the development of social capital. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp 75-87). Boston, MA: Allyn & Bacon.

Renzulli, J. S. (2005). The three-ring conception of giftedness: A developmental model of promoting creative productivity. In *Conceptions of giftedness* (2nd ed., pp. 246-279). New York, NY: Cambridge University Press

Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M. ,Hartman, R. K., & Westberg, K. L. (1997). *Scales for rating the behavioral characteristics of superior students*. Mansfield Center, CT: Creative Learning Press.

Roid, G. H. (2003). *Stanford–Binet Intelligence Scale manual* (5th ed.). Itasca, IL: Riverside.

Ross, P. O. (1993). *National excellence: A case for developing America's talent*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Rowe, E. W., Kingsley, J. M., & Thompson, D. F. (2010). Predictive ability of the General Ability Index (GAI) versus the full scale IQ among gifted referrals. *School Psychology Quarterly, 25*, 119-128. doi: 10.1037/a0020148

Ryan, J. J., & Brown, K. I. (2005). Enhancing the clinical utility of the WASI: Reliabilities of discrepancy scores and supplemental tables for profile analysis. *Journal of Psychoeducational Assessment, 23*(2), 140-145.

Ryan, J. J., Carruthers, C. A., Miller, L. J., Souheaver, G. T., Gontkovsky, S. T., & Zehr, M. D. (2003). Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Applied Neuropsychology, 10*(4), 252-256.

Ryan, J. J., & Ward, L. C. (1999). Validity, reliability, and standard errors of measurement for two seven-subtest short forms of the Wechsler Adult Intelligence Scale—III. *Psychological Assessment, 11*(2), 207-211.

Satz, P., & Mogel, S. (1962). An abbreviation of the WAIS for clinical use. *Journal of Clinical Psychology, 18*(1), 77-79.

Silverstein, A. B. (1967). A short form of Wechsler's scales for screening purposes. *Psychological Reports, 21*(3), 842.

Silverstein, A. B. (1985). Two-and four-subtest short forms of the WAIS-R: A closer look at validity and reliability. *Journal of Clinical Psychology, 41*(1), 95-97.

Sousa, D. (1995). *How the gifted brain learns*. Reston, VA: National Association of Secondary School Principals.

Sternberg, R. J. (2000). Multicultural issues in the testing of abilities and achievement. In G. B. Esquivel & J. C. Houtz (Eds.), *Creativity and giftedness* (pp. 105-124). Cresskill, NJ: Hampton.

Sternberg, R. J. (2005a). The WICS model of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 327-342). New York, NY: Cambridge University Press.

Sternberg, R. J. (2005b). WICS: A model of giftedness in leadership. *Roeper Review, 28*(1), 37-44. Retrieved from EBSCO*host*.

Terman, L. M. (1922). A new approach to the study of genius. *Psychological Review, 29*(4), 310-318. doi:10.1037/h0071072

Thompson, A., Browne, J., Schmidt, F., & Boer, M. (1997). Validity of the Kaufman Brief Intelligence Test and a four-subtest WISC-III short form with adolescent offenders. *Assessment, 4*(4), 385-394.

Thompson, A. P., Howard, D. D., & Anderson, J. J. (1986). Two- and four-subtest short forms of the WAIS-R: Validity in a psychiatric sample. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *18*(3), 287-293.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale (4th ed.) technical manual*. Chicago, IL: Riverside.

Tsushima, W. T., (1994). Short form of the WPPSI and WPPSI-R. *Journal of clinical Psychology, 50*(6), 877-880.

U.S. Department of Education. (1993). *National excellence: A case for developing America's talent.* Washington, DC: Author.

U.S. Department of Education. (2002). *Executive summary: The No Child Left Behind Act of 2001.* Washington, DC: Author.

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children-Fourth Edition among referred students. *Educational and Psychological Measurement, 66*(6), 975-983.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1989). *Wechsler Preschool and Primary Intelligence Scales for Children-Revised.* San Antonio, TX: Psychological Corporation .

Wechsler, D. (1991a). *Wechsler Intelligence Scale for Children-3rd Edition (WISC-III)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991b). *Manual for the Wechsler Intelligence Scale for Children-3rd Edition (WISC-III)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-3rd Edition (WAIS-III).* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-4th Edition (WISC-IV).* San Antonio, TX: Psychological Corporation.

Wendling, B. J., Mather, N., & Schrank, F. A. (2009). Woodcock-Johnson III Tests of Cognitive Abilities. In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement* (pp. 191-232). New York, NY: John Wiley.

Wilkinson, G. S. (1993). *WRAT3: Wide Range Achievement Test manual*. Wilmington, DE: Wide Range, Inc.

Woodcock, R. W., McGrew, K. S., Schrank, F. A., & Mather, N. (2007). *Woodcock-Johnson Tests of Cognitive Abilities-3rd Edition—normative update.* Itasca, IL: Riverside.

Zhu, J., Cayton, T., Weiss, L., & Gabel, A. (2008). *Wechsler Intelligence Scale-4th Edition* (Technical Report #7: Extended Norms). Bloomington, MN: Pearson.

Ziegler, A., & Ziegler, A. (2009). The paradoxical attenuation effect in tests based on classic test theory: mathematical background and practical implications for the measurement of high abilities. *High Ability Studies, 20*(1), 5-14.