

Recent Advances in the Assessment of Intelligence and Cognition

Sara S. Sparrow

Yale University Child Study Center, New Haven, U.S.A.

Stephanie M. Davis

Michigan State University, East Lansing, U.S.A.

In this paper, we review current issues in cognitive assessment. After addressing important definitional and theoretical issues, we discuss some recently developed cognitive assessment instruments as well as some recently revised instruments. Tests that are scheduled for revision will also be mentioned. As most readers are generally familiar with the widely used and nationally standardized IQ tests, we will summarize these tests according to their general usage. The testing of intelligence has been a major focus and contribution since the early days of Psychology, when the birthplace of the intelligence testing movement began in France with the work of Alfred Binet toward the end of the 19th century. Many of the most widely known and used IQ tests have been developed in the U.S.A. and are used internationally. In addition, other IQ tests have been developed in many other countries outside the U.S.A. The use of IQ tests and selected assessment considerations will be reviewed. Finally, we make some predictions about the future role of cognitive assessment in the coming century.

Keywords: Intelligence, cognition, assessment.

Abbreviations: BAS II: British Ability Scales: 2nd ed.; CAS: Cognitive Assessment System; K-ABC: Kaufman Assessment Battery for Children; KAIT: Kaufman Adolescent and Adult Intelligence Test; Leiter-R: Leiter International Performance Scale-Revised; PASS: Planning, Attention, Simultaneous, and Successive; UNIT: Universal Nonverbal Intelligence Test; WAIS-III: Wechsler Adult Intelligence Scale: 3rd ed.; WPPSI: Wechsler Primary and Preschool Scale of Intelligence.

Definitional Issues: Intelligence and Cognition

The assessment of intelligence has a long yet controversial history. In the lay literature, one often hears about the testing of “intelligence” rather than cognition. The two terms, however, are often considered synonymous. For this paper, cognition will refer to the processes whereby individuals acquire knowledge from the environment. Thus, the term cognition refers to the highest levels of various mental processes such as perception, memory, abstract thinking and reasoning, and problem solving as well as the more integrative and control processes related to executive functions such as planning, choosing strategies, and the enactment of these strategies. We will employ the term intelligence very narrowly, to refer to those abilities that are evaluated by intelligence tests (Boring, 1929).

Although there are many different definitions of and theories about cognition and intelligence (cf. Sternberg & Kaufman, 1998), almost all of them are concerned with the existence of multiple component processes. These processes and their unique features combine to produce complex cognitive tasks (e.g. problem solving). The components within the domain of cognition represent the

relationships among subsystems of the cognitive domain. Although these subsystems may differ in their degree of independence or coverage, it is suggested that each unit operates differently and through unique underlying principles. For example, as in a family, each individual member has unique characteristics. To fully understand the functioning of each individual, however, it is essential to learn about the family system. Similarly, to fully understand comprehensive cognitive functioning, one must comprehend the performance of the individual components as well as their integrated or gestalt functioning. Individuals may outperform the sum of their component processing abilities as they develop the capacity to compensate for relative weaknesses by relying on their areas of cognitive strength.

Before reviewing recent advances in assessment instruments, a selective discussion of the neuropsychological foundations of intellectual assessment is in order.

Contributions of Neuropsychology to the Assessment of Cognition

In recent years one area that has received more intense focus in the study of cognition has been the assessment of neuropsychological functioning in children and adolescents. The field of neuropsychology has emerged as the investigation of brain-behavior relations. Further, clinical neuropsychology is an applied science concerned with

Requests for reprints to: Dr S. Sparrow, Yale Child Study Center, 230 South Frontage Road, PO Box 207900, New Haven, CT 06520-7900, U.S.A.

the behavioral expression of brain dysfunction (Lezak, 1995). Although neuropsychological assessment has been studied rather extensively in adults, empirical work with children has been slower to develop.

In his classic studies, Luria (1973, 1980) proposed that three major brain regions are most useful for understanding intelligence: (1) a unit in the brain stem and midbrain structures that relates to arousal; (2) a unit comprised of the temporal, parietal, and occipital lobes that relate to sensory input; and (3) the frontal cortex, which relates to organization and planning. Whether or not one believes that the components of these systems are linked to structural regions of the brain, these neuro-behavioral or neuropsychological models of cognition can be used to evaluate cognitive, linguistic, and sensorimotor functions (Lezak, 1995).

In contrast to more traditional psychological models of cognition that distinguish between verbal and nonverbal (e.g. perceptual organizational) skills, most neuropsychological assessment models require the independent evaluation of (1) attention; (2) auditory, visual, and tactile perceptual functions; (3) verbal and language functions; (4) spatial/constructional processing abilities; (5) memory and learning; and (6) executive functions (conceptual reasoning, problem solving, planning, flexibility in cognitive strategies, and implementing cognitive plans).

Just as there are differences in rates of development across functional domains in the comprehensive assessment model, the development of the various component processes underlying cognition does not occur at the same pace. During infancy and early childhood, the more basic elements of attention and perception undergo the most rapid development, while in later childhood and adolescence, the development of higher-order linguistic, spatial, and executive elements is primary. Because of this evolving pattern of differentially emerging abilities, deficits or delays in the development of any of these various domains can lead to a wide range of configurations in the cognitive system. Such outcomes may have adaptive or maladaptive significance for a person's functional adjustment.

It must be reiterated that we do not consider the term cognition to be a synonym of IQ or intelligence. IQ represents the summary score on standardized "intelligence" tests. The IQ score may include measures of cognition, but can also include and often assesses sensory, motor, and related abilities that are not typically included under the term cognition. For example, individuals with motor dyscontrol (e.g. cerebral palsy), may score poorly on IQ measures reflecting timed tasks with motor components, despite the fact that these individuals often possess high levels of cognitive abilities (Sparrow, Carter, & Cicchetti, in press). In addition, infants have been shown to have a wide variety of cognitive abilities but, due to their heavy reliance on sensory-motor skill assessment at early ages, standardized developmental measures are not able to assess these abilities adequately (Carter & Sparrow, 1989). Finally, significant debate exists about whether IQ tests measure only what individuals have learned in the past (which is dependent in large part on their opportunity), or if they accurately assess one's innate cognitive abilities or future potential (Herrnstein & Murray, 1994).

Recent developments in the field of test construction have attempted to create new IQ tests that include more novel problem solving in an effort to focus on future learning potential and cognitive domains not always

addressed in traditional psychometric tests (e.g. the KAIT; Kaufman & Kaufman, 1983; the CAS; Nagliari & Das, 1997a). In many ways, the concept of IQ as a global index of cognitive functioning is considered to be obsolete from a comprehensive developmental assessment perspective (Sparrow et al., in press). *This evolution does not mean that the actual tests are obsolete, but rather, that their traditional use in isolation as a summary of an individual's capacities is antiquated.* Whereas some psychologists still regard an IQ score as the most fundamental measure of intelligence (Jensen, 1980), other professionals prefer to emphasize the distinct profile of strengths and weaknesses (Neisser et al., 1996). Because such measures are derived from multiple subtests and they are multifactorial and complex in their construction, an analysis of the various contributing components is required for adequate interpretation. In addition, further independent measures of cognitive functions and/or supplemental assessments of related functions are often required to interpret and integrate an IQ score into current comprehensive assessment models.

In a written discussion about the postulations put forth in Herrnstein and Murray's controversial book on the bell curve, Carroll (1997) observes that because of the need for more knowledge about cognitive skills and their properties, the nature of intelligence and the identification and measurement of cognitive skills are important topics for future research. Daniel (1997) also notes that the most cogent criticism of intelligence testing has been that the domain of assessed abilities is too narrow. Psychologists must consider both the various domains of cognition in their comprehensive assessment and continue to examine how current tests of intelligence have attempted to measure them. In the following discussion, four specific areas of cognition (attention, perception, memory, and executive functioning) are examined.

Domains of Cognition

Attention. Attention is the process that enables an individual to focus on the relevant information in a stimulus array while also inhibiting further processing of nonrelevant information (cf. Rothbart, Posner, & Hershey, 1995). Therefore, a child or adolescent must demonstrate the capacity to sustain attention by maintaining his or her focus on the important dimensions of the task presented without being confused, or pulled off-task by the unimportant distractions. In order to accomplish this focus, the child or adolescent must simultaneously ignore a wide range of environmental happenings and details. Attention is considered to be a limiting front-end feature of the perceptual process that controls the amount and quality of the information available for higher-order cognitive processes (Rothbart et al., 1995). In other words, attention is a prerequisite for the successful performance of more complex cognitive processes.

A familiar component in examining attention in currently used intelligence tests is the Freedom from Distractibility factor from the Wechsler scales (Barkley, 1998). This factor originally consisted of scores derived from the Arithmetic, Digit Span, and Coding (Digit Symbol) subtests. More recently with the WISC-III, the factor consisted of only Digit Span and Arithmetic. Klee and Garfinkel (1983) believe scores on this factor correlate to a low but significant degree with other tests of attention. Clinicians often use IQ data to assist in

determining cognitive factors that may contribute to an individual's inattention, which is probably a more appropriate use of intelligence tests for assessing attention difficulties. In fact, in administering IQ tests to children, nonquantifiable clinical information gained from observing a child's test-taking behavior may be as useful in determining problems in attention as any "scores" obtained from particular subtests. However, since problems in attention can come from motivational, biological, or emotional causes (or combinations thereof), interpretations of the attentional behaviors is usually challenging.

Perception. Perception is a central step in the processing of sensory/attentional information. Information perceived through sensory systems is later transformed into higher-order codes for use by the various higher-order cognitive subsystems. Perceptual functions include activities such as awareness, recognition, discrimination, patterning, and orientation (Lezak, 1995).

Whereas perception remains a legitimate realm to include in an assessment of cognitive ability, one needs to be cautious in the interpretation of such information. Making assumptions about the role of perception in the causality of cognitive deficits has led to questionable intervention strategies (Paul, 1995). Years ago Vellutino (Vellutino, Steger, Moyer, Harding, & Niles, 1977) cautioned about the misapplication of "the perceptual hypothesis" and the misguided interventions that were developed for reading and other learning disabilities. In addition, although perception, both visual and auditory, may often be one of the component parts of many IQ subtests (e.g. Block Design, Matrices, Digit Span, Word Order), it is rarely, if ever, a unique ability measured by a single subtest.

Memory. Another element of cognition, memory, is deemed to be the set of processes that temporarily holds new information while it is being utilized or processed for other purposes (short-term memory), or that more permanently holds learned information for future reference and use (long-term memory). In clinical evaluations, deficits in learning and memory have often been presumed to play significant role in the development of learning disabilities and other neurological and psychological disorders.

As we have discussed the lack of a universal definition for the term cognition, the same can be said for the definition of memory, particularly as it is defined by intelligence measures. Comprehensive intelligence scales have often failed to include a framework for the assessment of memory for this reason. The WAIS-III (which will be discussed in more detail later) contains a Working Memory Index that includes the Arithmetic, Digit Span, and Letter-Number Sequencing subtests. Despite its name and the introduction of the Letter-Number Sequencing subtest, which has provided more data on the cognitive function of working memory, interpretation of the Working Memory score alone is not acceptable. According to Kaufman and Lichtenberger (1999), it is best interpreted in light of observation and subtleties during testing.

The Stanford Binet IV is one commonly used IQ test that formally includes a memory domain. The four subtests of the Short-Term Memory Domain includes both meaningful and nonmeaningful tasks and both auditory and visual memory. However, the factor analytic studies have shown that the Memory domain is not appropriate for children under 7 years and that the Memory subtests appropriate for that age group all load

on other factors (Sattler, 1988). However, the Binet Short-Term Memory domain represents a well-standardized, reliable measure of some important aspects of memory.

The Kaufman Adolescent and Adult Intelligence Test (KAIT) is another IQ test which offers a memory subtest: Memory for Block Design. This nonverbal short-term memory test can also provide an important clinical picture of an individual's problem-solving style. The Information subtests on the Wechsler Scales are an example of another group of subtests which include items that attempt to measure long-term memory.

Although the study of memory is of great importance to the field of psychology, there has not been much recent development in this area with regard to instrumentation and measurement. Although there has been much research of memory functioning of adults (Hynd & Willis, 1988), there has been less documented research on memory development and functioning in children, largely due to the lack of standardized instrumentation. However, in recent years a large number of tests to measure many aspects of children's memory have been developed, such as the Rivermead Behavioral Memory Test (Wilson, Cockburn, & Baddeley, 1985), the Children's Memory Scale (CMS), and the Wide Range Attention and Memory and Learning (Sheshlow & Adams, 1990). So for the individual who wishes to supplement what is obtained from an administration of an IQ test with instruments that focus on memory, many tools are available. However, since these memory tests are not part of existing IQ tests, they will not be discussed further.

Executive functioning. The final component of a neuropsychologically oriented cognitive assessment is the evaluation of executive functioning. The term executive functioning has encompassed a number of meanings. Definitions have included those control and regulatory processes that: (1) integrate information perceived in the external world and transform perception into higher-order symbols; (2) compare incoming information with what knowledge is stored in memory; and (3) combine the incoming perceptions with information about the person's internal physiological state and biological drives. According to this terminology, executive functioning is arguably the most complex aspect of one's cognitive capacities, due to the variety of functions required to select, plan, organize and implement a behavioral response appropriate to a constantly changing world.

The assessment of executive functioning through psychological testing has been difficult to define because, as is evident from this discussion, of the excessive description of the term itself. Many "executive function" definitions have seemed so broad as to incorporate the whole range of human cognition. Some investigators' definitions of executive function have remained stable in recent years. For example, Pennington (1991) has defined the term as "the ability to maintain an appropriate problem solving set for the attainment of a future goal". Many professionals agree with this definition as research has proven that prefrontal cortical damage produces impairment on a wide variety of task including those tests that incorporate problem solving (i.e. Wisconsin Card Sorting Task, Tower tests) (Zelazo, Carter, Reznick, & Frye, 1997). Other individuals, however, have questioned the existence of the executive function altogether. For example, Parkin (1998) argues that the idea of a central executive function should be abandoned as the concept emerges from research by default when more rigorous

theoretical constructs cannot provide appropriate explanation.

As a result of the inconclusive definitions of the elusive conceptualization of executive function, it is difficult to determine how executive function is assessed with respect to current and future intelligence testing. Examiners should look carefully at instruments that purport to measure executive functions, as test makers may define this term quite differently. Some IQ subtests that include some aspects of executive function include Logical Steps, Mystery Codes, and Double Meanings, from the KAIT; Mazes, Picture Arrangement, and Similarities from the Wechsler Scales, and Matrices (found in several IQ tests). Many psychologists, however, choose to use specialized tests that have been developed specifically to measure executive functioning, rather than relying on subtests found in IQ tests which may contain some component ability that can be interpreted as executive functioning.

Summary. To summarize, within the functional domain of cognition, the psychologist must consider the following dimensions of cognition: attention, perception, memory, and executive functions, in addition to the more traditional notions of "intelligence". The psychologist must understand the contribution of these functions to intelligence quotients (i.e. performance on IQ tests) and recognize when it is critical to pursue additional testing in these areas.

As discussion of the terminology of both cognition and intelligence continues, recent advances have occurred in the development of current testing instruments. The following section describes the trend in the development of theory-based intelligence tests over the past several years. These emerging assessment measures differ considerably from the more widely used psychometric abilities tests (Wechsler Scales). Two original measures that attempt to incorporate neuropsychological understanding will be discussed, in addition to two revisions that have recently been published.

Recent Advances

Theory-based Intelligence Tests

In recent years, unlike their early Stanford-Binet and Wechsler predecessors, some newly developed measures of intelligence and cognition have adopted a different assessment approach. The majority of new cognitive assessment devices are based on psychological theory, unlike their more empirically based counterparts.

In his book *Intelligent testing with the WISC-R*, published in 1979, Kaufman decried the fact that, after years of research on intelligence and children's learning, there were still essentially no tests of intelligence based on information from newly developed theories grounded upon sound empirical research. Finally, in the 1980s, such instruments began to emerge. The leader in this endeavor was Alan Kaufman himself, who published the *Kaufman Assessment Battery for Children (K-ABC)* in 1983 with his wife Nadeen.

The K-ABC (Kaufman & Kaufman, 1983) is a battery of tests measuring intelligence and achievement of children ages 2½ through 12½ years. The K-ABC was standardized on a sample of 2000 children stratified on the basis of the U.S. census by region of the country, community size, ethnic group, and parental education. Although the K-ABC has been controversial since its inception, it is a well-accepted alternative or addition to the Wechsler

Scales in many schools, clinics, and research settings. Much of this controversy stems from the theoretical base underlying the test as well as the exclusion of language subtests from those subtests used to determine overall intelligence scores.

In 1984 an entire issue of the *Journal of Special Education* was devoted to the pros and cons of the K-ABC by numerous educators and scientists, followed by an eloquent rebuttal by Kaufman as the last paper (Kaufman, 1984). The K-ABC may also be more resistant to personality and temperament variables due to the number of sample and teaching items as well as the lack of emotionally laden content, which is characteristic of the Wechsler Scales (cf. Reynolds & Kamphaus, 1997). In addition, the theoretical basis of the K-ABC has made it a popular instrument in European countries such as France and Germany. In fact, for many children and particular diagnostic groups (i.e. Pervasive Developmental Disorders), the K-ABC has become the IQ test of choice.

The K-ABC is designed to measure the cognitive processes underlying general intellectual functioning. The test results yield a general mental processing score (Mental Processing Composite, MPC), equivalent to an intelligence quotient (IQ). The MPC is based on two global subscales that assess a child's style of problem solving and information processing. The Sequential Processing Scale measures a child's ability to process information gradually, in temporal or serial order. The Simultaneous Processing Scale assesses the child's ability to integrate several pieces of information at once and process these pieces as a whole or gestalt.

The K-ABC differs from more traditional tests of intellectual functioning because of its reduced emphasis on verbal abilities and knowledge of specific content. The K-ABC also includes a brief screening of achievement, knowledge typically acquired from school as well as the environment, in a separate section of the battery. The K-ABC intelligence scales are based on a theoretical framework of sequential and simultaneous information processing, which relates how children solve problems rather than what type of problems they must solve (e.g. verbal or nonverbal). This Sequential and Simultaneous Model stems from a variety of theories (Kamphaus & Reynolds, 1987), primarily from the information processing approach of Luria (1966, 1973, 1980), derived from his neurophysiological observations, in addition to empirical research conducted on Luria's model (Das, Kirby, & Jarman, 1975).

Several years later, the Kaufmans developed their second theoretically based intelligence test for individuals aged from 11 to 85 years, the KAIT. The KAIT was standardized on a sample of 2000 individuals appropriately stratified on the basis of the U.S. census. It is composed of separate Crystallized and Fluid Scales. The Crystallized Scale measures acquired concepts and depends on an individual's schooling and acculturation for achieving success on this scale, while the Fluid Scale examines one's ability to solve new problems.

The Horn-Cattell theory constitutes the foundation of the KAIT and defines the constructs theorized to be measured by the separate IQs. Other theories, however, guided the test development process, specifically the construction of each of the subtests. Tasks were developed from the models of Piaget's formal operations (Inhelder & Piaget, 1958; Piaget, 1972) and Luria's (1973, 1980) planning ability in an attempt to include high-level,

decision-making, and more developmentally advanced tasks. Luria's notion of planning ability involves decision-making, evaluation of hypotheses, and flexibility, and "represents the highest levels of development of the mammalian brain" (Golden, 1981, p. 285). The K-ABC and the KAIT (see Table 1) represented the beginning of a new age of intelligence tests. Even the last revision of the venerable Stanford-Binet (Thorndike, Hagen, & Sattler, 1986) represents an attempt at developing a theory-driven test.

One of the most recent intelligence tests to enter into the theory-based arena is the Cognitive Assessment System (CAS; Naglieri & Das, 1997a, b). We will describe this instrument in some detail as it is probably the least familiar to this readership and shows promise as a useful tool to assess intelligence and guide intervention due to its theory regarding cognitive functioning. McCallum and Bracken's (1997) Universal Nonverbal Intelligence Test (UNIT), another new measure of intelligence, will also be discussed.

New Measures

CAS. The CAS (Naglieri & Das, 1997a) is based on, and developed according to, the Planning, Attention, Simultaneous, and Successive (PASS) theory of intelligence (Naglieri, 1999). The PASS theory is a multi-dimensional view of ability, the result of the merging of contemporary theoretical and applied psychology recently summarized by Das, Naglieri, and Kirby (1994) and Naglieri and Das (1997b). This theory proposes that human cognitive functioning is based on the four essential PASS processes that employ and alter an individual's base of knowledge (Naglieri, 1999).

The CAS was intended to mirror the PASS theory, with subtests organized into four scales designed to provide an effective measure of each of the PASS cognitive processes. Planning subtests require the child to devise, select, and use efficient strategies or plans of action to solve the test problems, regulate the effectiveness of the plans, and self-correct these plans when necessary. Attention subtests call for the child to attend selectively to a particular stimulus and inhibit his or her attention to distracting stimuli. Simultaneous processing subtests ask the child to integrate stimuli into groups to form an interrelated whole; and the Successive processing subtests require the child to integrate stimuli in their specific serial order or appreciate the linearity of stimuli with little opportunity for interrelating the parts.

The CAS yields scales for the PASS and Full Scale, which provide normalized standard scores with a normative mean of 100 and a *SD* of 15. The Planning Scale subtests include Matching Numbers, Planned Codes, Planned Connections, and Planned Search; the Attention Scale subtests are Number Detection, Receptive Attention, and Expressive Attention; the Simultaneous Scale subtests include Nonverbal Matrices, Verbal-Spatial Relations, and Figure Memory; and the Successive Scale subtests are Word Series, Sentence Repetition, Sentence Questions, and Successive Speech Rate. Each subtest possesses a normative mean of 10 and a *SD* of 3.

Interpretation of the CAS also follows closely from the PASS theory with emphasis on each scale rather than analysis at the subtest level. Naglieri (1999) provides ample directions for evaluation of test results, integration of information about the strategies used during planning

tests, comparison of PASS scores, and methods for comparing the PASS scores to achievement using simple and predicted difference models. Additionally, illustrative case reports and a summary of relevant intervention research and their implications for treatment are provided.

The CAS was standardized on 2200 children ranging in age from 5 through 17 years, stratified by age, gender, race, ethnicity, geographic region, educational placement, and parent education. These demographics are based on recent U.S. Census reports and closely matches the U.S. population characteristics on these variables. In addition to being administered the CAS, a representative sample of 1600 included in the standardization sample were also administered achievement tests from the Woodcock-Johnson Tests of Achievement (Woodcock & Johnson, 1989). This additional testing provided a rich source of validity evidence (e.g. the analysis of the relationships between PASS and achievement) and for the development of predictive difference values needed for interpretation of ability achievement discrepancies. Finally, 872 children from special populations, including those individuals with attention deficit difficulties, mental retardation, and learning disabilities, were tested for validity and reliability studies.

Considerable validity research on the CAS is provided in the *CAS interpretive handbook* (Naglieri & Das, 1997b) and *Essentials of CAS interpretation* (Naglieri, 1999). Construct validity was supported by evidence of the tests' developmental changes, high internal consistency, the results of confirmatory factor analyses, and the utilization of strategies for completion of planning tests. Criterion-related validity was established by the strong relationships between CAS scores and achievement test scores, correlations with achievement for special populations, and PASS profiles for children with attention-deficit/hyperactivity disorders, traumatic brain injury, and reading disability. Ethnic, racial and gender fairness was determined by a series of studies of the prediction of achievement for Whites and Blacks, Hispanic and non-Hispanic, and males and females (Naglieri, 1999). In addition, the utility of the PASS scores for treatment and educational planning is demonstrated.

In summary, recent validity studies summarized by Naglieri (1999) have suggested that the CAS offers a useful alternative to traditional IQ batteries. Naglieri summarized 5 research studies involving more than 8000 children aged 5–17 years who were administered nearly all the major intelligence tests. He reported that the median correlations between the CAS global score and achievement were higher (.70) than all the other IQ tests studied (range between .59 and .63).

Kotarsky and Mason (in press, p. 9) stated that the "CAS has the potential to be a widely used test which is easily administered and scored, and it should prove to be a useful instrument among clinical and educational professionals in determining special needs of children. It is especially valuable [because] of its assessment to intervention model." Similarly, Gindis (1996) recognized that the CAS is an alternative to traditional tools, which can be used to help understand children's successes and failures.

Carroll (1995), on the other hand, provided two main criticisms of the PASS theory and thus, the CAS. First, he suggested that Planning is better described as a Perceptual Speed factor. He also argued that there was insufficient factorial support for the PASS as measures of the

constructs. Carroll's criticism that the Planning subtests actually measure speed merits further empirical investigation, but is inconsistent with two sources of data.

It remains to be seen whether the CAS will take its place among commonly used IQ tests. Its theoretical base, as well as the adequacy of its psychometric properties, may aid in its acceptance.

UNIT. The UNIT (Bracken & McCallum, 1997) is an intelligence assessment that requires no language from either the examiner or examinee. As the number of people whose first language is one other than English continues to grow, some psychologists have argued (Lopez, 1997; McCallum & Bracken, 1997) that nonverbal and other alternative models of the assessment of intelligence may be the most practical solution to accommodate these individuals. In addition, a nonverbal intelligence test should prove useful in evaluating children with language deficits or delays.

The UNIT contains six subtests, three of which were designed to measure memory (Object memory, Spatial memory, and Symbolic memory) and three intended to assess reasoning (Cube Design, Mazes, and Analogic Reasoning). The authors incorporated this two-tier model of intelligence of Reasoning and Memory (Jensen, 1980) to bring together two organizational strategies: symbolic and nonsymbolic organization (Bracken & McCallum, 1997). An individual uses symbolic organization strategies when he or she defines the environment through the use of concrete and abstract symbols (i.e. words, numbers, etc.). Nonsymbolic strategies are practised when the individual must make decisions or judgments about particular relationships within the environment. This ability is similar in definition to fluid intelligence, concerning one's ability to discriminate and solve novel problems.

In fact, the authors describe the UNIT's theoretical foundations as consistent with the Horn-Cattell Gf-Gc model of fluid and crystallized abilities, as memory and reasoning are two important identified areas of the Gf-Gc model. Memory and reasoning, as assessed by the UNIT, appear to correspond to the Gsm (short-term memory) and the Gf (fluid reasoning) factors, respectively (McCallum & Bracken, 1997). In addition, the symbolic measures are similar to Gc (comprehensive knowledge) and the nonsymbolic measures to Gv (visual processing) in the Horn-Cattell model as well.

Perhaps, unlike other nonverbal measures of intelligence, the authors of the UNIT emphatically state that this battery is a measure of intelligence obtained nonverbally. Intelligence assessed by the UNIT is defined as the ability to solve problems using memory and reason (Bracken & McCallum, 1997). Although problem solving is considered by many psychologists to be an integral aspect of intelligence (Sternberg, 1982), many practitioners consider language to be a critical element of assessment and an integral component of intelligence.

Standardization of the UNIT was based on a carefully selected sample representative of the U.S. population. Normative data was collected from 2100 children and adolescents and an additional 1765 individuals participated in reliability and validity studies (Bracken & McCallum, 1997). The sample considered several factors such as sex, race, region of the country, classroom placement, special education services, and parental educational attainment.

Reed and McCallum (1995) completed an early evaluation of both the reliability and construct and concurrent

validity of this measure. The UNIT and subtests from the Woodcock Johnson-Revised Tests of Cognitive Ability (WJ-R) were administered to elementary, middle, and high school students ($N = 104$). Spearman-Brown reliability coefficients provided evidence of the UNIT's reliability (Symbolic memory = .89, Cube Design = .92, Spatial memory = .87). Concurrent validity of the UNIT was examined through correlation analyses between the UNIT and the WJ-R. The correlation coefficients between global scores on both measures were high ($> .50$), providing validity support. The investigators also note that the memory subtests from the UNIT failed to load highly on the factor identified by subtests from the WJ-R short-term memory cluster. They believe this may be due in part to the relatively weak factor purity of the Gsm cluster.

One important limitation to recognize, particularly for a measure like the UNIT, is the predominantly White sample used (only two participants were from minority ethnic groups). If one of the authors' hopes for this measure is its utilization in the assessment of non-English speaking individuals, it would have been practical to include a more diverse sample.

As this instrument is quite new, much more research must be conducted before the UNIT can be considered a strong measure of intelligence. We look forward to components of the UNIT used to test the hypotheses generated by more standard and familiar measures of intelligence in addition to further empirical critique and review.

Current Revisions

In addition to the number of new measures that have been recently developed, revisions of two widely used measures of cognition have been published recently.

Leiter International Performance Scale—Revised (Leiter-R). The Leiter-R (Roid & Miller, 1997), an individually administered battery of 20 subtests, assesses cognition for children and adolescents between 2 years and 20 years, 11 months of age. The original scale was constructed in 1929 by Russell Leiter, whose objective was the creation of a measure to evaluate the intellectual abilities of Hawaiian children with verbal difficulties. A revision issued in 1948 was further developed as a result of testing of American children and Army recruits during the Second World War. The Leiter International Performance Scale (Leiter, 1979) has been used often in both clinical and research settings and is arguably one of the most widely used measures of nonverbal intelligence assessment. Other tests such as the Merrill-Palmer (Stutsman, 1948) and the Snijders-Oomen Nonverbal Intelligence Scale (Snijders & Snijders-Oomen, 1976), a test initially developed for deaf children (Berger, 1994), have also been used when dealing with children who may not understand verbal instructions.

The Leiter-R addresses a wide range of cognitive functions, similar to those fields found on more traditional, verbally loaded measures. Its construction was influenced by the Gf/Gc theory of the Woodcock-Johnson Psycho-educational Battery-Revised (Woodcock & Johnson, 1989). In its current form, coverage has been expanded and now encompasses four domains: Reasoning, Visualization, Attention, and Memory. The Attention and Memory areas are original to the Leiter-R and have been included to enhance the scale's clinical value. The authors believe that these new areas will be

particularly useful for individuals with attention deficits, learning disabilities, and brain injuries (Roid & Miller, 1997). Rating scales for the examiner, parent, self, and teacher have also been added.

A distinctive characteristic of the Leiter-R is the elimination of verbal instruction throughout the test's entire administration. This feature is important as it contributes to the test's validation as an impartial measure of nonverbal cognitive assessment.

The Leiter International Performance Scale (Leiter, 1979) was developed to contribute a fair nonverbal measure of intelligence. It has been widely used with children who have hearing difficulties, communication disorders, English as a second language, and in research with special populations. The scale, however, has been largely criticized for its poor norms and standardization as well as its outdated and cumbersome testing materials (Salvia & Ysseldyke, 1991; Sattler, 1988).

This most current edition of the Leiter is based on samples of over 2000 individuals from the United States. The Leiter-R standardization sample represents Caucasian, African-American, Asian-American, Hispanic, and Native American people in the same proportion to which they are found in the 1993 United States Census data. The revised Leiter scale has been improved significantly and is more sophisticated with regard to its psychometric qualities. The availability of representative contemporary norms and the expanded content should now strengthen the clinical utility of the scale (Anastasi & Urbina, 1997). Many visual and practical improvements were made in the development of the Leiter-R as well. Although administration may still be awkward due to the large number of testing materials, these materials are now lighter in weight, more visually colorful, and are expected to aid examiners in their evaluation.

Criterion-related and construct-related evidence of the validity of the Leiter-R has also been presented (Roid, Madsen, & Miller, 1997). A core of six subtests proved to have a consistent "g" factor across age groups as well as the ability to distinguish between criterion groups of gifted children (median = 115) and children with cognitive delay (median = 55). A Pearson correlation of .86 was found between the Leiter-R IQ and WISC-III Full Scale IQ ($N = 121$, $P < .001$). Regressions of Leiter-R IQ scores on achievement test scores showed reasonable equivalence in prediction across White and African-American samples.

Due primarily to its recent publication, there are no critiques and very few published studies utilizing or analyzing the Leiter-R. In one of the few available empirical investigations, Flemmer and Roid (1997) examined the nonverbal cognitive performance of adolescents from a variety of ethnic backgrounds. The study also compared the performance on the Leiter-R of speech-impaired adolescents with typical adolescents. According to the investigators, the Leiter-R produces ethnically fair results among the study sample as well as the assurance of unbiased assessment of the speech-impaired. Though Leiter-R literature is scarce, this early study seems to indicate that the new revision is a strong measure of assessment.

Although the Leiter-R may be an improvement over its predecessor in many regards, there are implications that must still be addressed. There is concern about whether this new version will be as effective in measuring the cognitive ability of individuals of lower-functioning populations (e.g. children with mental retardation or

autism). Further investigation of possible discrepancies between the two measures in addition to comparisons with other nonverbal measures are warranted.

Practical limitations of the Leiter-R exist as well. Instructions in the examiner's manual are not as clear as they should be. The acceptable amount of oral language allowed in explaining the test's directions to children is not clear and some children, particularly those individuals who possess some oral language, often find a "silent examiner" unsettling. Psychologists who are always careful not to break standardization are also concerned because there is uncertainty about how much language was actually used during test administration throughout standardization.

Wechsler Adult Intelligence Scale—3rd ed. (WAIS-III). The WAIS-III is the most recent revision of the Wechsler scales. This edition of the WAIS elicits intelligence scores for adolescents and adults between 16 and 89 years of age. The decision to revise the WAIS-R was based on criticism of the outdated norms, a desire to extend the age range to include older populations, and modification of older items and testing materials (Wechsler, 1997).

As was its predecessor, the WAIS-R, the new WAIS-III is constructed much like the Wechsler children's scale. The WAIS-III is the first of the adult Wechsler scales to provide factor scores. Three of the index scores are the same as the WISC-III factor scores: Verbal Comprehension, Perceptual Organization, and Processing Speed. In lieu of the fourth WISC factor, Freedom from Distractibility, a new factor named Working Memory has been developed. As mentioned briefly earlier, Kaufman and Lichtenberger (1999) find this new and expanded index score to be more reliable. Arithmetic, Digit Span, and a new subtest, Letter-Number Sequencing, constitute Working Memory. According to Kaufman and Lichtenberger, this new task draws from recent cognitive research and theory on working memory.

The WAIS-III standardization was completed on 2450 adult subjects. 1995 Census data was applied to the standardization sample in order to correctly represent individuals according to age, sex, race, geographic location, and educational level. The WAIS-III has generated impressive reliability and validity statistics for Verbal, Performance, and Full Scale IQs. The average split-half reliability for the IQ scores ranged from .94 to .98 and the factor indices had reliability coefficients ranging from .88 (Processing Speed) to .96 (Verbal Comprehension). Test-retest reliability scores ranged from .91 (Performance scale) to .96 (both Verbal and Full Scale IQs).

Many improvements were made with this most recent installment of the Wechsler Adult Intelligence Scale. Administration has been made easier for examiners due to the visual improvements of both the testing manuals and protocols. Floor and ceiling effects have been addressed as well in this revision. Additional step-down questions have been incorporated on each subtest of the WAIS. This adjustment should create more accurate scores for lower-functioning individuals.

Perhaps the most significant improvement of the WAIS-III is the elimination of a reference group for determination of all adults' and adolescents' test scores. The WAIS-R used a group of adults (ages 20–34 years) to compute the scaled scores for each individual, with no regard to their chronological age. Calculations for each individual's scaled scores are now based solely on the norms determined for his or her chronological age. This

Table 1
Intelligence Tests

Name/Publisher	IQ ranges		Age range	Testing time	Definition & special features
Wechsler Intelligence Scale for Children-3rd ed. (WISC-III), 1991 Psychological Corporation	Verbal Scale: 46–155 Performance Scale: 46–155 Full Scale: 40–160		6–16:11 yrs	Core subtests: 50–70 min	Measures children's intellectual ability according to Verbal, Performance, and Full Scale IQs. Also provides factor-based index scores for verbal comprehension, perceptual organization, freedom from distractibility, and processing speed.
Wechsler Primary and Preschool Scale of Intelligence-Revised (WPPSI-R), 1989 Psychological Corporation	Verbal Scale: 46–160 Performance Scale: 45–160 Full Scale: 41–160		3–7:3 yrs	75 min	Clinical intelligence instrument developed for use with younger children. Has one group of primarily perceptual motor (performance) subtests and another of verbal subtests. Both groups yield IQs and, combined, yield the full IQ.
Wechsler Adult Intelligence Scale-3rd ed. (WAIS-III), 1997 Psychological Corporation	Verbal Scale: 48–155 Performance Scale: 47–155 Full Scale: 45–155		16–89 yrs	60–90 min	Generates IQ and index scores (with the exception of working memory for freedom from distractibility) similar to the WISC-III.
Kaufman Assessment Battery for Children (K-ABC), 1983 American Guidance Service	Ages 2:6–2:11 3:0–3:11 4:0–4:11 5:0–5:11 6:0–12:5	IQs 69–160 55–160 52–160 46–160 40–160	2:6–12:6 yrs	33–85 min	Measure that assesses the intelligence and achievement levels of children. Generates global scaled scores for sequential, simultaneous, and mental processing. The achievement scale focuses on acquired facts and applied skills.
Kaufman Adult and Adolescent Intelligence Test (KAIT), 1993 Psychological Corporation	Crystallized IQ: 40–160 Fluid IQ: 40–157 Composite IQ: 40–160		11–85+yrs	Core battery: 60 min; expanded battery: 90 min	Test composed of crystallized and fluid scales of intelligence organized into a 6-subtest core battery and a 10-subtest expanded battery. Also offers a brief mental status test (supplementary). Whether the core or expanded battery is given, the standard scores are based only on the 6 core subtests.
Woodcock-Johnson Cognitive Battery-Revised (WJ-R), 1989 Riverside Publishing	Extended standard scores: 24–200		2–90+yrs	5 min per subtest	Standard battery is comprised of 7 tests, each measuring a different aspect of intellectual ability. The supplemental battery contains 14 additional measures of cognitive ability. Generates W scores, age equivalents, grade equivalents, Relative Mastery Indices (RMIs), as well as standard and extended standard scores.
Stanford Binet-4th ed. (SB-4E), 1986 Riverside Publishing	Each standard age score (SAS) and test composite: 36–164		2–23:11 yrs	33–85 min	Subtests each assess one of the following subject areas: verbal reasoning, quantitative reasoning, abstract/visual reasoning, and memory. Examinee is given a vocabulary test that serves as a routing test to determine the starting level for all other tests.

(continued)

Cognitive Awareness System (CAS), 1997 Riverside Publishing	Planning Scale Attention Scale Simultaneous Scale Successive Scale 40–160	5–17:11 yrs	40–60 min	Measure of intelligence based on the PASS theory, this test assesses planning, attention, successive, and simultaneous processing.
British Abilities Scales (BAS II), 1996 NFER-Nelson (UK)	Preschool: General conceptual ability score (GCA): Age 2:6–3:5: 45–159 Age 3:6–5:11: 37–162 Verbal ability: 50–151 Nonverbal reasoning ability: 43–162 Spatial ability: 49–150 Special nonverbal composite (SNC): Age 2:6–3:5: 51–152 Age 3:6–5:11: 44–156 School age: General conceptual ability score (GCA): 39–160 Verbal ability: 52–148 Nonverbal reasoning ability: 49–150 Spatial ability: 47–152 Special nonverbal composite (SNC): 41–158	Early years: 2:6–5:11 School Age: 6:0–17:11	30–45 min	Yields a general conceptual ability (GCA) which denotes g, the general factor. All subtests contributing to the GCA are highly g loaded. For school-aged children, the BAS-II is formed from three clusters: Verbal, Nonverbal Reasoning, and Spatial. These clusters are interpretable as Gc (crystallized ability), GF (fluid ability) and Gv (broad visualization) in the Horn-Cattell and Carroll models. These three factors are the most highly correlated with higher-order g. Also provides a special nonverbal composite (SNC) which is a measure of g with the contribution of the verbal tests removed.
Differential Abilities Scales (DAS), 1990 Psychological Corporation	Preschool: General conceptual ability score (GCA): Age 2:6–3:5: 44–169 Age 3:6–5:11: 44–175 Verbal ability: 50–153 Nonverbal reasoning ability: 43–162 Spatial ability: 49–150 Special nonverbal composite (SNC): Age 2:6–3:5: 45–158 Age 3:6–5:11: 43–162 School Age: General conceptual ability score (GCA): 45–164 Verbal ability: 51–151 Nonverbal reasoning ability: 52–152 Spatial ability: 50–155 Special nonverbal composite (SNC): 48–162	2:6–5:11 yrs (preschool); 6–17:11 (school age)	Full cognitive battery: 45–65 min	Consists of a cognitive battery of 17 subtests divided into two overlapping age levels. Produces a general conceptual ability score composed of reasoning and conceptual abilities, a nonverbal composite, and cluster composite scores derived from the core subtests.

(continued)

Table 1
(Cont.)

Name/Publisher	IQ ranges	Age range	Testing time	Definition & special features
Leiter International Performance Scale, 1948 Stoelting Company	Adjusted IQ score: 25–160	2–28 yrs	30–45 min	Nonverbal test of intelligence used to evaluate children with sensory, motor, or language deficits. Contains 54 subtests.
Leiter International Performance Scale-Revised (Leiter-R), 1997 Stoelting Company	Brief IQ screener: 36–169 Full Scale IQ: 30–170	2–20:11 yrs	Varies	Used to evaluate cognitive functions including measure of nonverbal intelligence, fluid reasoning and visualization, visual-spatial memory, and attention. Very different structure and format to the previous version of the Leiter.
Universal Nonverbal Intelligence Test (UNIT), 1998 Riverside Publishing	Full Scale (abbreviated): 45–153 Full Scale (standard): 41–159 Full Scale (extended): 40–159	5–17:11 yrs	Abbreviated battery: 10–15 min; standard battery: 30 min; expanded battery: 40 min	Measure for children and adolescents who may be at a disadvantage in traditional verbal and language loaded tests. Although it is entirely nonverbal, it is designed to provide assessment of general intelligence, cognitive abilities, and memory.
Test of Nonverbal Intelligence-3 (TONI-3) 1997 Pro-Ed	Deviation quotient: 6 → 150	5–85+yrs	15–20 min	Language-free measure that assesses intelligence, aptitude, abstract reasoning, and problem solving. Because no reading, writing, speaking, or listening is required, this test is often used with individuals whose language ability is limited or suspect.
Ravens: (1) Colored Progressive Matrices; (2) Standard Progressive Matrices; (3) Advanced Progressive Matrices, 1986 Psychological Corporation	Ranges vary depending on age: 30–141	(1) 5–11:11 yrs, physically impaired; (2) 6–16 yrs, 18+yrs; (3) 12–16 yrs, 18+yrs	20–45 min 20–45 min 40–60 min	Nonverbal test of reasoning ability presented in three different forms. Can be administered individually or to a group. Scores are converted into percentile ranks.

important modification generates intelligence scores that are significantly more appropriate and accurate than in the WAIS-R.

Another important improvement is the removal of the Object Assembly subtest from the calculation of an individual's IQ score. The fact that this subtest did not contribute substantially to any of the test's factor scores indicates that this decision was psychometrically sound. The average split-half reliability on Object Assembly in WAIS-III standardization was .70, while its test-retest reliability was .76. The authors did choose to keep the subtest in the battery, however, for those psychologists who find it clinically useful.

Although there has been much discussion about the atheoretical nature of all the Wechsler scales, they are by far the most widely used cognitive assessment instruments today. There is currently no reason to believe that the Wechsler measures will diminish in popularity any time in the near future. In recent years, the trend for tests other than the Wechsler scales has been to tie new measures to theory and what both science and research have taught us about how individuals learn, problem solve, and process information across the life span. Despite this controversy about the utility of theory-based tests, ironically, all of these measures are reasonably good predictors of success or failure in school (Sparrow et al., in press).

The British Ability Scales: 2nd ed. (BAS II). This is a revision of the well-established instrument which is used primarily in Great Britain for measuring cognitive functioning over a wide age range. In fact, the original scale was revised and standardized in the United States as the Differential Abilities Scales (Elliott, 1990). The British revision is based on a new standardization sample of 1700 children in the United Kingdom.

In the revision the coverage of BAS II has been strengthened, particularly for younger children, by more child-friendly materials. The upper age limit is now extended up to 17 years 11 months. The authors have attempted to shorten administration time by improving the accuracy of starting and decision points. At most age levels, six scales are used to derive the composite General Conceptual Ability score (GCA). The Core Scales used in GCA calculation were chosen on the basis of factor-analytic research. Those scales showing a high and consistent *g*-loading were included in the GCA. Scales not included in the GCA were classified as Diagnostic Scales, and are used to investigate functioning in specific areas that include, for example, information processing and memory.

The scales of BAS II are clearly divided into two batteries—the *Early Years* (covering the 2:6 to 5:11 years age range), and *School Age* (covering the 6:0 to 17:11 years age range).

Although there were numerous studies conducted using the original BAS, there is little yet to be found investigating the newer version.

A summary of the tests just discussed plus other commonly used IQ tests appear in Table 1.

International Perspective

As in the United States, the Wechsler Scales are the most consistently used measures around the world. According to the Psychological Corporation the WISC-R has been translated into 13 different languages and there are currently at least 17 translations of the WISC-

III either completed or in progress (personal communication). These translations include Japanese, Chinese, French, and Greek (Priftera, Weiss, & Saklofske, 1998).

Despite Wechsler dominance, issues have often been raised regarding item and test score bias. This concern is prevalent even in other English-speaking countries. Saklofske and Janzen (1990) discuss these considerations, particularly when norm-referenced tests are employed or the instrument being measured is tied to specific or cultural experiences. For this reason, standardization studies of the WISC-III have been carried out in several countries including Australia, the U.K., and Canada, with more in progress.

The K-ABC has also achieved considerable attention abroad. Translations have been conducted in Arabic, French, Swedish, Hebrew, German, and Spanish among others. Standardizations and cultural modifications have been undertaken in several countries. The fact that the K-ABC is a theory-based test has given it special appeal in many countries and have made it a strong competitor to the Wechsler scales.

Current Issues in Intelligence Testing

In the following paragraphs, we will discuss a few testing "caveats" and issues that are currently important and will continue to be significant in the next few years.

Testing of Low-functioning Populations

There has been an increased desire in both the psychological and educational fields for the development of sound testing measures for lower-functioning individuals. This development has recently had and will most likely continue to have educational implications for future assessment innovations. For example, many new intelligence batteries have attempted this (e.g. the UNIT) and newer revisions have tried to include more step-down items. As increased emphasis continues to be placed on educational programming and planning for children and adolescents with severe and profound mental retardation, well-developed batteries are required in order to calculate IQ scores below 40. In fact, most currently available standardized IQ tests have serious limitations at the lowest levels of ability. For example, on the Wechsler Primary and Preschool Scale of Intelligence-Revised (WPPSI-R), the lowest score a child can receive on the verbal scale is 46. Floor effects on many instruments illustrate the need for measures that can attain an accurate assessment, even for those individuals who are greatly impaired and those children who are very young. The current trends for developing interventions for younger and younger children and for those individuals with severe to profound developmental delays make development of appropriate assessment instruments essential.

Intelligence Screeners

Another common issue that faces researchers today, particularly for investigations involving children and adolescents, is the impact intelligence and IQ score may have on their results. For this reason, most investigators look for assessment tools that can estimate children's and adolescents' intelligence accurately and quickly, so as not to consume too much valuable research protocol time. In many cases, the assessment is for such basic reasons as

Table 2
Intelligence Screeners

Name/Publisher	IQ ranges	Age range	Administration time	Definition & special features
Wechsler Abbreviated Scale of Intelligence (WASI), 1999 Psychological Corporation	50–160	6–89 yrs	15–35 min	Brief measure of intelligence that yields Verbal, Performance, and Full Scale IQ scores. Offers flexibility with its 2- and 4-subtest options. Nationally standardized.
Kaufman Brief Intelligence Test (K-BIT), 1990 American Guidance Service	Vocabulary, Matrices, and Full Scale composites: 40–160	4–90 yrs	15–30 min	Brief individually administered screener of verbal and nonverbal intelligence. Developed specifically to be used for screening and related purposes. Norms were established on census data.
Kaufman Short Neuropsychological Assessment Procedure (K-SNAP), 1994 American Guidance Service	K-SNAP composite: 51–150	11–85+yrs	30 min	Assesses and generates scores for the areas of Attention-orientation, Memory and perceptual skills, Intellectual functioning, and Planning ability.

ruling out mental retardation, or the assurance that control and experimental groups do not differ significantly in cognitive ability.

In part, publishers and investigators have developed “screeners” or short forms of longer more standard intelligence measures for the above explanations. Unfortunately, such abbreviated batteries are also used inappropriately for clinical evaluations because of their expedience. Clinicians have utilized “short forms” (i.e. adopting portions of the Wechsler Scales [Kaufman, Kaufman, Balgopal, & McLean, 1996]) for many years. Other screener methods have been “developed” by using two domain versions of tests (e.g. the Peabody Picture Vocabulary Test and the Ravens) to approximate the verbal performance model of the Wechsler Scales. Berger (1994) considers this type of testing “incomplete” and will thus restrict interpretation of scores due to reduced reliability.

Recently, however, two “screener” IQ tests have been developed especially for this purpose. The Wechsler Abbreviated Scale of Intelligence (WASI) and the Kaufman Brief Intelligence Test (K-BIT) represent important tools in the field of intelligence testing. The WASI correlations with Full Scale IQs on the WISC-III is .87 and with the WAIS-III is .92. Correlations with the K-BIT are .75 with WAIS-R and .80 with the WISC-R. These screeners possess increased reliability and validity over “created” short forms, because of their having been appropriately normed and standardized. Both of these screeners correlate highly with more comprehensive IQ tests but still have no place in a comprehensive assessment of development. However, screeners have an appropriate research application when determining relative intellectual strengths or weaknesses between groups of subjects.

Group-administered Tests

Another type of intelligence measure we should mention briefly is the group-administered test. We want to caution clinicians and investigators about the limitations of such measures, however. We are experiencing a time

when clinical evaluation intervals continue to become smaller, and many clinicians are looking to the administration of group tests for answers. Unlike the current trends toward theory-based measures in individually administered IQ tests, theory has played a smaller role in group-administered tests of the multiple-choice variety (Kaufman, in press). Knowledge of the utility of group tests of intelligence is important, however, because our current society makes frequent use of group tests within the educational system, military services, industry, and government service (Anastasi & Urbina, 1997).

The most popular group-administered tests are multi-level batteries, whose coverage includes the primary grades through high school, as well as multiple aptitude batteries, designed for adolescents and adults, which are often used for educational and career counseling. Such tests have consistently been revised, computerized, and improved substantially by the application of new and sophisticated psychometric procedures, but they have only rarely been impacted by advances in psychological theory (Kaufman, in press). Occasional group tests have been theory-based, most notably the SOI Learning Abilities Test (Meeker, Mestyanek, Shaddock, & Meeker, 1975), developed from J. P. Guilford’s (1967) Structure-of-Intellect model of intelligence. For our purposes, however, we will not be discussing these batteries because they are not appropriate for clinical evaluations. These tests have typically not met the psychometric rigors of standardization and validation that characterize most individual tests of intelligence, and their impact has been modest. Group-administered tests have no place in a comprehensive assessment of children or adolescents.

Future Directions

In the last few years, we have seen a resurgence of interest in the assessment of intelligence and cognition. Since Kaufman’s (1979) publication of *Intelligent testing with the WISC-R* and that of the K-ABC, it is no longer acceptable for new tests of intelligence only to provide simple subtest and global scores. To be appropriate,

intelligence tests must now be based on highly sophisticated and extensive psychometric expertise. Tests must also provide elaborate interpretive information. Yet, controversy remains over how this interpretation should be implemented. Subtest interpretation has been frowned upon by some psychologists (McDermott, Fantuzzo, & Glutting, 1990; Witt & Gresham, 1985), while it is still strongly supported and elaborately explicated by others (Kaufman, 1994; Sattler, 1988).

The time constraints encountered in schools and managed care environments may, unfortunately, restrict and constrain the in-depth clinical interpretation that can be so useful in understanding individual children and adolescents in the future. We contend that the "basic tenets of the intelligent testing approach" advocated by Kaufman (1994) will still be the cornerstone of cognitive assessment for enlightened psychologists of the future.

In the last two decades of the 20th century, there have been an abundance of new and revised tests to measure cognition. The future of the assessment of cognition will probably continue to be marked by new revisions and new tests based on new and revised theories. In fact, the revisions of two Wechsler scales, the WPPSI-R and the WISC-III, have begun, and the revision of the K-ABC began in 1998. The K-ABC revision will strongly address criticisms of the original version (Kaufman, 1999, in press).

Neuropsychological assessment approaches and instruments that measure the various subdomains of cognition will also receive considerable attention in the beginning of the next century. In the infancy of neuropsychological testing instruments, very few, if any, of the tests conformed to the psychometric rigors demanded for traditional, nationally standardized intelligence tests. Neuropsychological tests followed a more medical, clinical model instead of a psychometric one. Furthermore, children in particular, represent a relatively recent focus for neuropsychological assessment.

There is a component to intelligence that is essentially connected to the values and demands of a culture (Perlman & Kaufman, 1990). Intelligence assessment of the future must reflect aspects of that culture, particularly the manner in which information is sent and received. In Western society, for example, information transmitted through television, video, and the computer is becoming the norm. With this acculturation, new types of tasks and methods of testing incorporating these changes will need to be produced.

Summary

The assessment of intelligence and cognition is in a fluid rather than crystallized state (Daniel, 1997). The recent appearance of numerous new and revised batteries, however, has not seemed to interrupt practitioners' reliance on older, more traditional instruments (i.e. the Wechsler Scales). There is reason, as a result, to expect that the high level of test development activity will continue for some time.

The evolution of intelligence testing will be shaped by further development in theory in addition to basic and applied research. Because of the strength of empirical findings for psychometric-ability tests, theory-based tests must prove that they are worthy of being considered as more than just alternative testing measures. Practical benefits with respect to both educational planning and

psychological diagnostic assessment must be demonstrated for each new published measure.

References

- Adams, W., & Sheslow, D. (1990). *The Wide Range Assessment of Memory and Learning*. Wilmington, DE: Jastak.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Baddeley, A. D., & Hitch, G. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485–493.
- Barkley, R. A. (1998). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford Press.
- Berger, M. (1994). Psychological tests and assessment. In M. Rutter, E. Taylor, & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches* (3rd ed., pp. 94–109). Cambridge, MA: Blackwell Science.
- Bigler, E. D., Lowry, C. M., & Porter, S. S. (1997). Neuroimaging in clinical neuropsychology. In A. M. Horton Jr, D. Wedding, & J. Webster (Eds.), *The neuropsychology handbook Vol. One: Foundations and assessment* (2nd ed., pp. 171–220). New York: Springer Publishing Company.
- Boring, E. G. (1929). *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts.
- Bracken, B. A., & McCallum, R. S. (1997). *The Universal Test of Nonverbal Intelligence*. Chicago, IL: Riverside Publishing.
- Carroll, J. B. (1995). Review of the book: *Assessment of cognitive processes: The PASS theory of intelligence*. *Journal of Psychoeducational Assessment*, 13, 397–409.
- Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence*, 24, 25–52.
- Carter, A. S., & Sparrow, S. S. (1989). The acquisition of adaptive behavior in infancy: Developmentally at-risk infants. In R. K. Creasy & J. B. Warshaw (Eds.), *Seminars in perinatology* (pp. 474–481). Philadelphia, PA: Grune & Stratton.
- Cohen, M. J. (1997). *Children's Memory Scale manual*. San Antonio, TX: The Psychological Corporation.
- Daniel, M. H. (1997). Intelligence testing: Status and trends. *American Psychologist*, 52, 1038–1045.
- Das, J. P., Kirby, J., & Jarman, R. F. (1975). Simultaneous and successive synthesis: An alternative model for cognitive abilities. *Psychological Bulletin*, 82, 87–103.
- Das, J. P., Naglieri, J., & Kirby, J. (1994). *Assessment of cognitive processes*. Boston, MA: Allyn & Bacon.
- Elliott, C. D. (1990). *Differential Ability Scales: Introductory and technical handbook*. San Antonio, TX: The Psychological Corporation.
- Flemmer, D. D., & Roid, G. H. (1997). Nonverbal intellectual assessment of Hispanic and speech-impaired adolescents. *Psychological Reports*, 80, 1115–1122.
- Gindis, B. (1996). Review of the book *Assessment of cognitive processing: The PASS theory of intelligence*. *School Psychology International*, 17, 305–308.
- Golden, C. (1981). *Diagnosis and rehabilitation in clinical neuropsychology* (2nd ed.). Springfield, IL: Thomas.
- Golden, C. J. (1987). *Luria-Nebraska Neuropsychological Battery: Children's revision*. Los Angeles, CA: Western Psychological Services.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Hynd, G. W., & Willis, W. G. (1988). *Pediatric neuropsychology*. New York: Grune & Stratton.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence; an essay on the construction of formal operational structures*. New York: Basic Books.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

- Kamphaus, R. W., & Reynolds, C. R. (1987). *Clinical and research applications of the K-ABC*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: John Wiley & Sons.
- Kaufman, A. S. (1980). Issues in psychological assessment: Interpreting the WISC-R intelligently. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology, Vol. 3* (pp. 177–214). New York: Plenum Press.
- Kaufman, A. S. (1984). K-ABC and controversy. *Journal of Special Education, 18*, 409–444.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: John Wiley & Sons.
- Kaufman, A. S. (in press). Tests of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence*. New York: Cambridge University Press.
- Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., & Kaufman, N. L. (1993). *Manual for Kaufman Adolescent and Adult Intelligence Test*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., Kaufman, J. C., Balgopal, R., & McLean, J. E. (1996). Comparison of three WISC-III short forms: Weighing psychometric, clinical, and practical factors. *Journal of Clinical Child Psychology, 25*, 97–105.
- Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. New York: John Wiley & Sons.
- Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance task: A new measure of attention. *Journal of the American Academy of Child and Adolescent Psychiatry, 11*, 486–496.
- Kotarsky, D., & Mason, E. (in press). A review of the Cognitive Assessment System. *Insight*.
- Leiter, R. G. (1979). *Instruction manual for the Leiter International Performance Scale*. Wood Dale, IL: Stoelting Company.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Lopez, E. C. (1997). The cognitive assessment of limited English proficient and bilingual children. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 503–516). New York: Guilford Press.
- Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.
- Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. London: Penguin Books.
- Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.
- McCallum, R. S., & Bracken, B. A. (1997). The Universal Nonverbal Intelligence Test. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 268–280). New York: Guilford Press.
- McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment, 8*, 290–302.
- Meeker, M. N., Mestyaneck, L., Shaddock, R., & Meeker, R. (1975). *SOI Learning Abilities Test*. El Segundo, CA: SOI Institute.
- Naglieri, J. A. (1997a). Planning, attention, simultaneous, and successive theory and the cognitive assessment system: A new theory-based measure of intelligence. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 247–267). New York: Guilford Press.
- Naglieri, J. A. (1997b). Intelligence knowns and unknowns: Hits and misses. *American Psychologist, 25*, 75–76.
- Naglieri, J. A. (1999). *Essentials of CAS Assessment*. New York: Wiley & Sons.
- Naglieri, J. A., & Das, J. P. (1997a). *Cognitive Assessment System*. Itasca, IL: Riverside Publishing Company.
- Naglieri, J. A., & Das, J. P. (1997b). *Cognitive Assessment System interpretive handbook*. Itasca, IL: Riverside Publishing Company.
- Naglieri, J. A., & Johnson, D. (1998). Improving math calculation: A cognitive intervention based on the PASS theory. *Communications, 27*, 23.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 5*, 77–101.
- Parkin, A. J. (1998). The central executive does not exist. *Journal of the International Neuropsychological Society, 4*, 518–522.
- Paul, R. (1995). *Language disorders from infancy through adolescence: Assessment and intervention*. New York: Mosby.
- Pennington, B. F. (1991). *Diagnosing learning disorders*. New York: Guilford Press.
- Perlman, M. D., & Kaufman, A. S. (1990). Assessment of child intelligence. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 59–78). New York: Pergamon Press.
- Piaget, J. (1972). *The child's conception of the world* (trans. by J. Tomlinson & A. Tomlinson). Totowa, NJ: Littlefield.
- Prifitera, A., Weiss, L. G., & Saklofske, D. H. (1998). The WISC-III in context. In A. Prifitera & D. H. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 1–38). New York: Academic Press.
- Reed, M. T., & McCallum, R. S. (1995). Construct validity of the Universal Nonverbal Intelligence Test (UNIT). *Psychology in the Schools, 32*, 277–290.
- Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.
- Reynolds, C. R., & Kamphaus, R. W. (1997). The Kaufman Assessment Battery for Children: Development, structure, and applications in neuropsychology. In A. M. Horton Jr, D. Wedding, & J. Webster (Eds.), *The neuropsychology handbook Vol. One: Foundations and assessment* (2nd ed., pp. 291–330). New York: Springer Publishing Company.
- Roid, G. H., Madsen, D. H., & Miller, L. J. (1997). *Validity and fairness of prediction in the Leiter-R*. Paper presented at the meetings of the American Psychological Association, Chicago, IL, August 15, 1997.
- Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised: Examiner's manual*. Wood Dale, IL: Stoelting Company.
- Rothbart, M. K., Posner, M. I., & Hershey, K. L. (1995). Temperament, attention, and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 315–342). New York: John Wiley & Sons.
- Saklofske, D. H., & Janzen, H. L. (1990). School-based assessment research in Canada. *McGill Journal of Education, 25*, 5–23.
- Salvia, J., & Ysseldyke, J. E. (1991). *Assessment* (5th ed.). Boston, MA: Houghton Mifflin Company.
- Sattler, J. M. (1988). *Assessment of children's abilities* (3rd ed.). San Diego, CA: Author.
- Sheshlow, D., & Adams, W. (1990). *Wide Range Assessment of Memory and Learning*. Wilmington, DE: Jastak.
- Snijders, J. Th., & Snijders-Oomen, N. (1976). *Snijders-Oomen Non-verbal Intelligence Scale*. Windsor, U.K.: NFER-Nelson.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales—survey form*. Circle Pines, MN: American Guidance Service.
- Sparrow, S. S., Carter, A. S., & Cicchetti, D. V. (in press). *Comprehensive psychological and psychoeducational assessment of children and adolescents: A developmental approach*. Boston, MA: Allyn & Bacon.
- Sternberg, R. J. (1982). Reasoning, problem solving, and

- intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). New York: Cambridge University Press.
- Sternberg, R. J., & Kaufman, J. C. (1998). Human abilities. *Annual Review of Psychology*, 49, 479–502.
- Stutsman, R. (1948). *Stutsman Copy Test*. Wood Dale, IL: Stoelting Company.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale: Fourth ed.* Chicago, IL: Riverside Publishing.
- Vellutino, F. R., Steger, B. M., Moyer, S. C., Harding, C. J., & Niles, J. A. (1997). Has the perceptual deficit hypothesis led us astray? *Journal of Learning Disabilities*, 10, 375–385.
- Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence—Revised*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children (3rd ed.) manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (3rd ed.) manual*. San Antonio, TX: The Psychological Corporation.
- Wilson, B. A., Cockburn, J., & Baddeley, A. (1985). *The Rivermead Behavioral Memory test*. Reading, U.K.: Thames Valley Test Company; Gaylord, MI: National Rehabilitation Services.
- Witt, J. C., & Gresham, F. M. (1985). Review of the Wechsler Intelligence Scale for Children—Revised. In J. V. Mitchell (Ed.), *Ninth mental measurements yearbook* (pp. 1716–1719). Lincoln, NE: University of Nebraska Press.
- Woodcock, R. W. (1998). The WJ-R and Bateria-R in neuropsychological assessment. *Research Report Number 1*. Itasca, IL: Riverside Publishing.
- Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery-Revised*. Itasca, IL: Riverside Publishing.
- Zelazo, P. D., Carter, A. S., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem solving framework. *Review of General Psychology*, 1, 198–226.