The Intelligent Testing of Children with Specific Learning Disabilities

NANCY MATHER
University of Arizona

INTRODUCTION

The category of specific learning disabilities (SLD) encompasses a heterogeneous group of disorders that adversely impacts the development of some aspect of academic functioning and proficiency. Although few doubt the existence of SLD, a lack of consensus regarding definition, as well as a failure to resolve various identification and treatment issues, has plagued the field since its inception. One major area of controversy is the use of intelligence tests for the identification of individuals with SLD (Kaufman & Kaufman, 2001). During the last four decades, the work and writings of Dr. Alan Kaufman, in association with his wife, Dr. Nadeen Kaufman, have helped to clarify, refine, and substantiate the most efficacious ways intelligence tests can and should be used with individuals having or suspected of having SLD. The purpose of this chapter is to discuss various aspects of their contributions that have had particular relevance to and impact on the field of learning disabilities through my own perspective. The chapter includes discussion of several issues that have affected SLD identification and assessment procedures, as well as consideration of the most pragmatic and valid processes and procedures for diagnosing SLD.

DEFINITION OF SPECIFIC LEARNING DISABILITIES

One component of the controversy surrounding intelligent testing of individuals with SLD has been agreeing upon the fundamental characteristics of these disorders. To identify a disorder accurately, one must first define and delineate the characteristics that typify the problem. In essence, the basic component of nearly all SLD definitions is that learning disabilities comprise a category of specific disorders in one or more of the basic
psychological processes involved in learning. Because development is uneven with some abilities being far more advanced than others, a discrepancy exists between a set of intact cognitive processes and one or more disordered processes (Hale, Naglieri, Kaufman, & Kavale, 2004; Kaufman, 2004). These weaknesses in the basic psychological processing or intracognitive discrepancies are considered to be the hallmark of SLD (Kaufman & Kaufman, 2001; Kavale, Kaufman, Naglieri, & Hale, 2005). Thus, to fulfill the requirement of the SLD definition, it is necessary to document cognitive integrities and one or more disorders within the basic psychological processes that are linked to academic difficulties (Hale et al., 2004; Kaufman, 2004).

Within both school and clinical settings, a major purpose of cognitive assessment is then to describe the fundamental disorder or disorders that have affected some, but not all aspects, of cognitive and academic performance. Unfortunately, the recent IDEA 2004 reauthorization does not realign identification procedures with the definition, and contributes further to a disconnection between the SLD definition and the selected assessment methodology. The reauthorization maintains the original definition of learning disabilities but does not provide a methodology for diagnosing SLD or make specific recommendations regarding how to identify and document the spared and disordered processes (Kaufman, Lichtenberger, Fletcher-Janzen, & Kaufman, 2005). In essence, the definition of SLD demands the identification of a processing disorder because identification is the central characteristic and defining feature of the disability (Kaufman, 2004) but the law does not specify how this should be accomplished.

During a class seminar on learning disabilities, the late Dr. Samuel A. Kirk proffered that “... a learning disability is like pornography; it’s hard to define but you know it when you see it” (personal communication, September, 1981). The analogy holds true because most skilled psychologists, diagnosticians, and special education teachers recognize various types of learning disabilities when they encounter them. These seasoned professionals are guided, however, by clinical impressions and experience, rather than by legalistic constraints that have been imposed by rigid interpretation of the law. For example, after years of remediation to improve accuracy, a person with a reading disability may obtain a score in the average range on a measure of reading decoding (i.e., standard score of 90), but their scores on measures of processing speed and reading speed still remain well below average. In school districts that demand strict adherence to a legalistic criterion, slow processing speed and reading speed by themselves may not be considered sufficient to document the existence of a reading disability. The skilled evaluator will make an accurate
diagnosis based upon a review of background information, as well as the functional limitation that results from slow reading speed on timed tests. Thus, an accurate diagnosis adheres to the definition, but must be multidimensional in nature.

ABILITY–ACHIEVEMENT DISCREPANCIES VERSUS CLINICAL TRADITION

For the last three decades, the field of SLD has been dominated by pre-occupation with mathematical formulas for identification of SLD despite the fact that intelligence tests were not originally built or designed to be used in mathematical formulas (Kaufman & Kaufman, 2001). After the passage of PL 94-142, most states and school districts based documentation of SLD on the existence of an ability–achievement discrepancy, rather than documentation of unusual variability among an individual’s cognitive, perceptual, linguistic, and academic abilities. In fact, it appears that the biggest discrepancy that existed was between the SLD definition and then how this definition was operationalized (Hale et al., 2004; Kavale et al., 2005). Systems of educational classification were based upon special claims that intelligent quotients (IQs) actually measured a person’s intellectual potential, a belief that was neither conceptually nor psychometrically justifiable (Kaufman, 1979; Stanovich, 1999). This misplaced emphasis resulted in increased attention to test scores, and decreased attention to the meaning and inferences that could be derived from a careful analysis of the results.

The many limitations inherent with the use of an ability–achievement discrepancy procedure have been enumerated repeatedly (e.g., Aaron, 1997; Berninger, 2001; Fletcher et al., 1998; Fuchs, Mock, Morgan, & Young, 2003; Lyon, 1995; Mather & Healey, 1990). Bateman (1992) observed that the problems with the use of a formula for SLD identification are “...many, serious, and too often disregarded” (p. 32). To further complicate matters, state and school district guidelines have varied in regard to the specific method used to quantify a discrepancy, as well as to the magnitude of the discrepancy that is needed to qualify for services. Therefore, a child may be identified as having a learning disability in one school district, but then denied services in another, depending upon the state and the local criteria or the personal philosophy of an independent evaluator (Berninger, 1996).

In many school settings, unrealistic rules, which overemphasized the value and role of the obtained test scores, governed the eligibility decisions.
(Kaufman, Harrison, & Ittenbach, 1990). Some school psychologists were even required to employ a specific formula to establish the existence of a significant discrepancy because of a misguided belief that this simplistic procedure was tantamount to legal compliance. Because of these state and district mandates, the focus became more upon compliance with eligibility policy, rather than on thoughtful clinical interpretation of multiple sources of information. The real problem was not, however, the clinicians, but rather the state guidelines that mandated the use of these formulas (Kaufman & Kaufman, 2001). This is not meant to imply that the test themselves were not useful. It was just that their diagnostic and clinical utility was lost because of a narrow focus upon just a few broad-based scores. As aptly noted by Willis and Dumont (2002), the determination of a disability should be more than “an exercise in arithmetic” (p. 173) and students who truly have SLD should not be denied services “simply because of the results of a statistical exercise” (Dumont, Willis, & McBride, 2001, p. 13).

Thus, in some districts, school psychologists became viewed as “gatekeepers” who held the mechanism for providing or denying special education services to students. One puzzling aspect regarding this whole process was that qualification for services was not based upon a need for help or even the level of severity of the problem, but rather upon the existence of a discrepancy between the scores from some selected intelligence test and some achievement battery. Children with low intelligence were usually not classified as students with SLD and rarely received the help that they needed (Siegel, 1989). Common sense dictates that mathematical formula should not be used to deny services to students who truly require a program of special education or to falsely “label” a student who does (Willis & Dumont, 2002). Clearly, unfairness and bias exist in a legal system that requires categorization and placement in services before students can receive help that is obviously needed, but these are different issues; unfair guidelines do not negate the fact that SLD is a meaningful category (Kaufman & Kaufman, 2001).

In spite of these constraints, well-trained clinicians have always used test instruments wisely. The beliefs of those who adhere to the clinical tradition have forever been at odds with the rigid ways in which local, state, and federal agencies enforced requirements for diagnostic and placement decisions (Kaufman, 2000). Sole reliance on a quantitative procedure clearly abdicates common sense. In fact, the use of formulae is the complete opposite of what skilled evaluators are trained to do. One may even assert that the identification of an ability–achievement discrepancy in a
rigid, quantitative fashion has no meaningful role in SLD identification (Kavale et al., 2005).

In fact, the ability–achievement discrepancy procedure was never a good idea from the start because plugging the most global score into a formula was just “plain dumb” (Kaufman, 2004). Kaufman further observed that elimination of the discrepancy requirement from IDEA 2004 frees school psychologists from the shackles of the past. Kaufman and Kaufman (2001) reiterated that: “The time has come to release professionals from the burden of using psychometric instruments for purposes that the test authors never intended, and in ways (e.g., plugging obtained scores into uncompromising formulas) that defy a common-sense understanding of psychometrics” (p. 450). To be an intelligent evaluator, psychologists must go beyond mere score generation (Kaufman & Harrison, 1991) and translate psychoeducational data into educational action (Kaufman & Kaufman, 1983b).

Problems with Full-Scale Intelligence Test Scores

Strict adherence to a formula for the diagnosis of SLD resulted in an overreliance on global or full-scale scores that did not convey useful information for educational planning. When considering the diagnostic and clinical implications of cognitive assessment results, the least useful score is the broad-based, full-scale IQ score. The concept of general intelligence or g does little to assist with the identification or diagnosis of individuals with SLD. In fact, it is antithetical to most modern theories of intelligence (Kaufman, 2004).

Stanovich (1999) aptly defined intelligence as “… the statistical amalgamation of a panoply of different cognitive processes” (p. 352). A full-scale IQ score simply represents the individual’s relative standing compared to the norm group, based on their aggregate performance at a specific point in time on a specific set of tasks that are designed to measure the test authors’ conceptualization of intelligence (Mather & Wendling, 2005). As early as 1938, Stern reflected upon the limited value of global IQ scores, stating: “To be sure there has been and there still is exaggerated faith in the power of numbers. For example, ‘an intelligence quotient’ may be of provisional value as a first crude approximation when the mental level of an individual is sought; but whoever imagines that in determining this quantity he has summed up ‘the intelligence’ of an individual once and for all, so that he may dispense with the more intensive qualitative study, leaves off where psychology should begin” (p. 60).
Furthermore, all broad-based or composite scores mask the contribution made by reading-related cognitive abilities, so for individuals with specific reading disabilities the results must be interpreted with caution (Vellutino, 2001). Orton (1925) cautioned that for individuals with reading problems, full-scale scores tend to provide an entirely erroneous estimate of intellectual capabilities. Similarly, Kaufman (1979) advised that intelligence tests can underestimate a child’s intellectual functioning and without careful interpretation, the results are unfair, and even “hazardous” to the child’s welfare.

In addition, because the existence of SLD can impact and lower the full-scale score, definitions of SLD that require normal intelligence as part of the identification criteria are also suspect (Kaufman et al., 1990). A low intelligence score does not rule out the possible existence of a learning disability; the evaluator has to examine and consider the reasons for the low score. In addition, superior intelligence does not rule out the possibility of a learning disability. The evaluator has to consider all facets of performance. A learning disability can exist in people of any age who are at any level of intellectual functioning (Cruickshank, 1983; Shaywitz, 2003).

**IMPACT OF PROCESSING AND LINGUISTIC DEFICITS**

Processing deficits are as likely to impair performance on intelligence tests as on measures of academic performance (Kaufman & Kaufman, 2001). In addition, limited school achievement impairs performance on fact-oriented items and when communication problems occur, a child’s intelligence test scores undoubtedly suffer (Kaufman, 1994). Similarly, Fletcher et al. (1998) observed that: “To the extent that the child who reads poorly has a significant language disorder, scores on a language-based IQ test will underestimate the child’s aptitude for learning in other areas” (p. 200).

For individuals with SLD, as well as those with language impairments, certain subtests measure the weak cognitive and linguistic processes that can be the defining characteristics of the disorder, thus reducing the full scale score. These tests can be primarily linguistic in nature (e.g., low vocabulary), providing supportive information for a language impairment, or perceptual in nature (e.g., slow processing speed or poor memory span), providing supportive information for SLD. When examining an individual’s performance on a subset of tests, the evaluator must consider how certain factors or abilities are suppressing the more global, broad-based scores. In these instances, an explanation must be provided.
regarding how one or more specific abilities are impacting linguistic or academic development.

For example, on many intelligence tests, considerable emphasis is placed on measures of acquired knowledge or crystallized intelligence and as a result, the low scores can be either the cause or the effect of the disorder (Kaufman, 1979). Because depressed performance may be the result of poor school achievement and limited learning opportunities, the obtained results may not provide an accurate estimate of learning potential. In addition, research supports the conclusion that verbal abilities for SLD samples decrease over time, demonstrating the gradual impact of the learning disability upon subsequent linguistic and academic performance and development (Kaufman, 1994). Kaufman and Harrison (1991) stated: “Thus, performance on an intelligence test may misrepresent the potential ability of individuals, which is less immune to the effects of culture and past learning, because the test is measuring an achievement-like component” (p. 100). In some instances, nonverbal tests provide a more accurate measure of general ability for children with SLD (Kaufman & McLean, 1986). Even with nonverbal measures, however, the evaluator must take care to acknowledge the potential impact of low specific abilities that are typically classified as nonverbal in nature (e.g., processing speed, visual-motor, visual-spatial thinking, etc.) on the broad-based score.

DISTINCTION BETWEEN INTELLIGENCE AND ACHIEVEMENT TESTS

With regard to the use of an ability–achievement discrepancy as the main criterion for SLD identification, one flaw was related to the proposed distinction between the contents of intelligence and achievement tests. If an individual’s scores on the two different measures were to be used to document SLD, then the measures should be relatively independent and distinct from each other. A clear distinction does not always exist, however, between the contents of intelligence and achievement tests. Each can measure aspects of the same underlying ability. For example, if a vocabulary test requires oral questions and responses, it would be included as part of an intelligence test battery. If, on the other hand, the test requires reading words and providing oral or written responses, the test would be included as part of an achievement battery. If a person does not have difficulty with reading or writing, the same underlying ability, that is knowledge of word meanings, is being measured.
Kaufman (1984) contended that the large, unrotated first factor (g) of many cognitive tests may actually be better described as a measure of general achievement rather than as general intelligence. If many of the tests (cognitive and achievement) load on the same factor, then the examiner may actually be comparing two aspects of a person’s performance on the same underlying mental construct (Kaufman & O’Neal, 1988). We now know that many valid measures of school achievement have high g loadings, thus the somewhat mired distinction between the content of these test batteries makes comparisons between a single ability and achievement score artificial at best (Kaufman & O’Neal, 1988). Fortunately, the mandatory discrepancy requirement has been removed from the reauthorization of IDEA 2004, providing increased flexibility in identification procedures, but conceivably, creating a new set of issues and concerns regarding SLD identification procedures. As noted before, the reauthorization does not provide clear guidance regarding how to operationalize the definition, or how to develop uniform criteria that can be applied from state to state, or even school district to district.

**RESPONSE TO INTERVENTION AS A METHOD FOR SLD IDENTIFICATION**

As specified in the IDEA 2004 reauthorization, states may permit a process that examines whether or not a student responds to scientific, research-based intervention as *part* of the learning disability evaluation procedure. This process of data collection is often referred to as response to intervention (RTI). Although RTI may be used as part of the process, it is not clear how big that part is, or how and when that part is integrated into the diagnostic assessment process.

Many states have been implementing or are planning to implement some form of RTI with the hopes of (a) reducing the number of students referred for evaluations, (b) providing early intervention to children in a more timely fashion, (c) providing targeted assistance to all children who need help, and (d) increasing the validity of actual placement decisions. Clearly, efficient progress monitoring and early intervention can provide benefits to and help improve the quality of instruction to all children. It seems, however, that RTI can be viewed most accurately as a prereferral intervention rather than a proxy for SLD.

Limited response to a treatment does not confirm or negate the existence of SLD (Kavale et al., 2005; Kavale, Kauffman, Bachmeier, & LeFever, 2008). The primary concern of RTI is not valid SLD identification and its
use as an identification procedure perverts the SLD category, making it into a convenient home for any student who otherwise might be left behind (Kavale et al., 2008). A failure to respond to treatment only indicates that a student is having problems achieving at a certain rate or level with a selected intervention as compared to his or her peers. Numerous reasons exist for why a student would not fully respond to a certain intervention, only one of which is a learning disability. Some of the reasons for low achievement are extrinsic, whereas others would be considered intrinsic.

Examples of extrinsic factors that could affect responsiveness to an intervention would include an ineffective methodology, improper implementation of an intervention, or poor school attendance. In addition to SLD, examples of other intrinsic factors that could influence responsiveness to treatment would include: attention problems, oral language impairments, behavior disorders, and cultural and linguistic differences. All low-achieving students should not be lumped into a single package (Kaufman & Kaufman, 2001). Different treatments, as well as alterations in the rate and mode of presentation of materials, will be effective with different learners (Kaufman & Kaufman, 1983b).

In fact, it seems somewhat inexplicable that RTI has only been recommended as a procedure for SLD identification, when it may be equally, if not more, effective for helping to substantiate other types of disabilities, such as behavior disorders or mild mental retardation. If a child with behavioral problems showed a poor response to increasingly structured reinforcements and behavior management techniques, this information certainly has direct relevance to the evaluation and identification procedure. In other words, a failure to respond to treatment (or more accurately a limited or insufficient response to a selected treatment) does not provide any type of unique information that is only applicable to the category of SLD. The removal of cognitive processing measures is likely to increase classification errors, as there are multiple causes of low achievement, other than SLD (Hale et al., 2004). In addition, increasing the breadth of the category to all students with low achievement is likely to increase the number of students being identified as needing special education services.

In essence, various forms of RTI have always been practiced by good psychologists and teachers. Zach (2005) provided an interesting reflection regarding this fact: “When I was working as a school psychologist some 50 years ago and received a referral from a teacher about a child who was having trouble learning, the very first thing I did was to visit the teacher to inquire about the problem. I wanted to know what the child was having
trouble with. I wanted to know what the teacher had tried that did not work and most importantly, I wanted to know what had been done that had worked. At that time I had never heard of Response to Intervention (RTI) and I certainly would not have predicted that there were going to be initials to describe what has always been good practice, as a ‘new’ procedure” (p. 151).

Furthermore, RTI does not provide the field with a consistent means of determining responsiveness, and the application of different methods results in different prevalence rates and different subsets of unresponsive children (Fuchs, Fuchs, & Compton, 2004; Reynolds, 2005). Essentially RTI is based on a discrepancy from grade level, making it a special case of severe discrepancy analysis that assumes everyone is of equal ability or possesses similar academic aptitude (Reynolds, 2005).

Consider the case of Jason, a fourth-grade student who has been identified as having SLD (poor short-term memory), fetal alcohol syndrome (diagnosed at birth), and attention-deficit hyperactivity disorder (ADHD) (diagnosed in preschool). Upon a premature birth, he was addicted to both cocaine and heroin and his birth weight was below 2 pounds. He has received special services since preschool with small group instruction from a special education teacher daily for 45 minutes. His adoptive parents have provided him with private tutoring for one hour three times weekly from a learning disability specialist since first grade. In addition, he has attended summer school every year. Although Jason is doing quite well considering the complex nature of his attention and learning problems, he is still not achieving at grade level in any area except verbal knowledge. School personnel have suggested retention as an option, but even then, his academic performance would still lag behind his peers in reading, writing, and mathematics. Jason has challenges that other children do not have to face; he also provides his teachers with special challenges in adapting and modifying the curriculum so that he can learn and succeed.

Nearly every classroom in the country has one or more students with needs as intensive as Jason’s. From years of research on developmental differences, it is apparent that fundamental differences exist in human abilities. Over eight decades ago, Caldwell and Courtis (1924) explained that educational psychology and measurement have accumulated a mass of data that demonstrates that the most constant trait about human nature is its variability. People differ and different standards are necessary to accommodate these differences. Larson (2005) admonished promotion of a uniform criteria for all children, stating: “It is time to resolve the convoluted thinking that mandates the ‘same’ high (‘rigorous’) grade-level standards
for all. One of the things that we know for sure in special education is that one size does not fit all, and that the same standards, rigorous or not, will not result in the same outcomes” (p. 248).

Thus, RTI is really subject to the major criticism that was made against the use of an ability–achievement discrepancy as a method of SLD identification: the unreliability of the diagnostic criterion (Fuchs et al., 2004). In addition, RTI only addresses the achievement criterion of the SLD definition. To help resolve these critical issues, Hale, Kaufman, Naglieri, and Kavale (2006) suggested that a multitiered approach to serving children with learning problems be used, one that begins with RTI, but then provides for a comprehensive evaluation of cognitive processes when RTI methods are unsuccessful in resolving a student’s learning difficulties. RTI procedures should be combined with psychometric testing, including cognitive ability testing, which provides both diagnostic and instructional data (Kavale et al., 2008). If a child fails to respond to intervention, then the results of a comprehensive evaluation indicate that the child has a processing deficit that impacts academic performance; both the definitional criteria for SLD and the limited response to evidence-based instruction as part of the SLD eligibility criteria have been addressed, resulting in a balanced model that promotes diagnostic accuracy (Hale et al., 2006).

CLINICAL ASSESSMENT OF STRENGTHS AND WEAKNESSES

Kaufman (2000) credits Dr. David Wechsler for converting intellectual assessment from a “psychometric exercise to a clinical art” (p. 10). The major purpose of the use of intelligence tests on individuals with SLD is to identify a person’s specific strengths and weaknesses. In fact, an understanding of this comprehensive profile is the key to interpretation (Kaufman, Kaufman, & Shaughnessy, 2007). The findings are then critical to determining appropriate accommodations and interventions. The evaluator hypothesizes the nature of the problem through careful analysis of an individual’s profile of scores. In the past, the various problems associated with interpreting individual differences through profile analysis of the Wechsler scales were discussed (e.g., Glutting, McDermott, & Konold, 1997). One must consider, however, that these studies were performed using tests that lacked specificity or were mixed measures of cognitive abilities, rather than using more psychometrically defensible and discrete factors, such as defined by the Cattell–Horn–Carroll (CHC) theory (Carroll, 1993; Horn & Cattell, 1966).
Profile interpretation for individuals suspected of having SLD can lead to insights regarding the child’s relative strengths and weaknesses, as well as suggesting the need for differentiated educational materials (Kaufman, 1976a; Kaufman, 1979). In fact, the documentation of significant variability among a person’s abilities is precisely how intelligence tests contribute to SLD determination and educational planning (Mather & Wendling, 2005). Prudent subtest analysis can hold the key to understanding a child’s unique profile (Kaufman, 1982).

Intelligent testing requires that the evaluator use test instruments as tools for helping to understand an individual’s strong and weak areas and then finding and confirming the hypothesis or hypotheses that explain the pattern of scores through multiple sources of information (Kaufman, 1979; Kaufman & Lichtenberger, 1998). In addition to measuring the weak or problems areas, an evaluator should also administer cognitive tests that circumvent the deficiency to document the person’s spared or intact neuropsychological assets (Kaufman, 1979; Kaufman & Kaufman, 2001). Individuals with SLD typically present an uneven profile of abilities demonstrating difficulty with some types of learning, but ease with others.

For example, one consistent finding across Wechsler Intelligence Test for Children – Revised (WISC-R; Wechsler, 1974) studies was that children with SLD exhibited high scores on spatial abilities or the perceptual organization subtests, average scores on verbal comprehension subtests, but consistently lower scores on the Arithmetic, Coding, Information, and Digit Span subtests, or what has been referred to as the ACID (Arithmetic, Coding, Information, and Digit) profile (Kaufman, 1981; Kaufman et al., 1990). In fact, it appears that for students with SLD, Coding and Arithmetic are the two most difficult subtests, followed by Digit Span (Kaufman et al., 1990). Thus, the intent of a comprehensive evaluation is to reveal the profile of an individual’s unique learning abilities.

As noted, individuals with superior cognitive abilities can have SLD. Because these individuals may have average scores on achievement measures (despite advanced verbal and reasoning abilities), their difficulties can only be identified through an evaluation of strengths and weaknesses and consideration of background information and educational history. After examining the strengths and weaknesses of children of superior intelligence as well as SLD on the WISC-R, Schiff, Kaufman, and Kaufman (1981) concluded that the group of children exhibited advanced verbal comprehension and oral expression skills, but weaknesses on the Arithmetic, Coding, and Digit Span subtests, as well as in emotional and motor development.
In essence, many individuals with SLD have intact verbal or oral language abilities and obtain scores in the average to above average range in many linguistic abilities (with the exception of areas like phonological awareness and verbal short-term memory). These students can orally verify and clarify knowledge with ease, but their performance falters when they have to read or write the answers to test questions. In essence, assessment of their subject-area knowledge is compromised by their reading and/or writing disabilities, and a written examination assesses the disability, rather than their actual level of understanding. Such practice results in a very biased and highly inaccurate measure of their actual knowledge.

In analyzing performance, an evaluator should consider both ipsative (within-child) strengths and weaknesses, as well as normative, standard score indexes greater than 115 and less than 85 (Kaufman et al., 2005). The noted discrepancies within performance must constitute statistically significant deviations that are unusual in the normative population. Before interpreting the amount of scatter that exists within a person’s profile, one must first consider what normal fluctuations are and what constitutes atypical scatter (Anderson, Kaufman, & Kaufman, 1976; Kaufman, 1976a; McLean, Reynolds, & Kaufman, 1990). After reanalysis of the WISC-R standardization sample, Kaufman (1976b) cautioned that one must be wary of interpreting test scatter because normally achieving children also exhibited numerous significant differences between pairs of test scores.

Furthermore, one must consider the difference between what is “normal” or in the average range as is statistically defined (between 85 and 115) and what is average or typical as compared to peers. For example, a child whose test score is better than 17% of their peers (standard score of 87) would be defined as obtaining a score within the statistically normal range. This child’s performance, however, is below average in terms of age cohorts, and the measured ability may be considered a significant weakness in light of additional quantitative information and qualitative observations. The point being that within any given categorization scheme, considerable variability in actual performance still exists. Test authors in the past have attempted to clarify and refine this classification system by referring to the standard score range between 80 and 89 as below average (Kaufman & Kaufman, 1983a) or low average (Woodcock, McGrew, & Mather, 2001).

Intelligence tests have multiple applications for SLD evaluations, including documenting the intact abilities, uncovering the processing deficits that lead to academic problems, understanding a person’s unique profile, and developing recommendations for educational intervention.
(Kaufman & Kaufman, 2001). Thus, the ipsative assessment approach can be used to hypothesize strengths and weaknesses within a profile but then the observed pattern must be supported or negated by additional relevant information (Kaufman, Lichtenberger, & Naglieri, 1999; Lichtenberger, Mather, Kaufman, & Kaufman, 2004). These practical applications are best accomplished using tests that are well designed and theory-based. When combined with behavioral observations and background information, understanding of diverse theoretical perspectives allows for interpretation of an individual’s test score profile in a unique way (Kaufman, 2000).

**Importance of Theoretical Frameworks for Test Interpretation**

In the early years of test development, theory was subservient to practice in determining what tests were developed and used (Kaufman, 2000). Kaufman (1981) insisted that the application of theory to test interpretation fostered a more process-oriented treatment of the strengths and weaknesses of children with SLD, which then would translate more readily into appropriate educational interventions. Throughout the last several decades, the Kaufmans have focused upon enhancing test interpretation by building a rich theoretical and clinical base that placed emphasis on the role of and importance of cognitive and neuropsychological theory to test development. By studying and understanding various cognitive theories, practitioners are better equipped with an understanding of the educational implications (Kaufman & Kaufman, 1983b). If practitioners are familiar with a variety of theoretical models, they can select the particular framework that provides the most insight into an individual’s cognitive structure (Kaufman, 1982).

Although the Wechsler scales were not developed from a specific theory, the scales have advanced theory and have been studied and interpreted from a variety of different theoretical perspectives (Kaufman, 2000). In summarizing research findings related to SLD assessment, Kaufman (1981) explored the usefulness of the WISC-R and concluded that: (a) little practical value existed for discrepancies between Wechsler’s verbal-performance dichotomy as a marker for SLD; (b) children with SLD obtained consistently low scores on the Arithmetic, Digit Span, and Coding subtests; (c) Bannatyne’s four-category approach of (1) verbal conceptualization (Similarities, Vocabulary, Comprehension), (2) spatial (Picture Completion, Block Design, Object Assembly), (3) sequencing (Arithmetic, Digit Span, Coding), and (4) acquired knowledge (Information,
Arithmetic, Vocabulary) had better clinical utility and led to more process-oriented treatment than any WISC-R two- or even three-factor solution; and (d) future research should attempt to clarify what the specific factors mean in either a clinical or theoretical sense. Thus, the specific factors that are measured by tests should provide meaningful information and treatment validity. In addition, tests that provide a multiple factor approach to interpretation provide richer information than those with only a two- or three-factor solution.

Fortunately, over the past 20 years, theory has become paramount in the construction and interpretation of tests (e.g., Das, Naglieri, & Kirby, 1994; Kaufman, 2000; Kaufman et al., 2007; Naglieri, 2001; Woodcock et al., 2001). Emphasis has shifted to a qualitative approach in which cognitive theories have played an important role (Gunnison, Kaufman, & Kaufman, 1982). The original Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983a) was built on a blend of theories: the integration of psychological cerebral specialization theory (Sperry, 1968), Luria’s (1966, 1980) neuropsychological theory, and Horn’s and Cattell’s (1966) distinction between fluid and crystallized intelligence.

The K-ABC was predicated on the distinction between problem-solving abilities and factual knowledge (Kaufman et al., 1999). The K-ABC also provided measures of two types of processing: sequential and simultaneous. These two types of information processing relate to how children solve problems, rather than the types of problems they are asked to solve (Kaufman & Lichtenberger, 1998). Sequential processing refers to problems that need to be solved in a serial order or where stimuli are processed one by one or in some kind of linear arrangement, whereas simultaneous involves problems requiring gestalt-like, holistic integration of data (Gunnison et al., 1992). One common finding has been that children with reading disabilities tend to have average performance in simultaneous processing but weaknesses in sequential processing, which is related to problems in decoding, spelling, and memorization (Gunnison et al., 1982; Kaufman, 1982; Kaufman & Kaufman, 1983). Thus, the theoretical rationale provided by the K-ABC made it particularly appropriate for neuropsychological evaluations and the assessment of children suspected of having SLD (Kaufman & Harrison, 1991).

The Kaufman Adolescent & Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993) was also based on theoretical constructs, including Horn and Cattell, as well as Luria’s (1980) concept of planning ability and Piaget’s (1972) formal operational theory. Within the KAIT framework, tasks measure the retrieval and application of general knowledge, as well as
problem-solving tasks that measure the ability to learn and to solve novel problems that are less dependent on prior experience and acquired knowledge (Kaufman et al., 1999).

More recently, the KABC-II (Kaufman & Kaufman, 2004) has continued the use of two modern theoretical models: the CHC model of broad and narrow abilities, and Luria’s neuropsychological processing theory. Using a subset of subtests from the KABC-II, the evaluator can interpret the test results from two different, but complementary perspectives (Kaufman & Kaufman, 2004). CHC theory places emphasis on the interpretation of specific cognitive abilities and understanding how these abilities are related to performance, whereas Luria’s neuropsychological theory emphasizes the way children process information when solving problems.

When considering the assessment of individuals suspected of having SLD, both approaches are useful for different types of referral questions. In clinical cases that involve the role of language as a factor affecting performance, application of the Luria framework deemphasizes the role and importance of factual knowledge and allows the evaluator to consider mental processing while reducing the impact of low language performance. The authors state the following fundamental principle for understanding when to include or exclude measures of acquired knowledge (Gc) from an evaluation: “measures of Gc should be excluded from any score that purports to measure a person’s intelligence or overall cognitive ability whenever the measure of Gc is not likely to reflect that person’s level of ability” (p. 4). In most instances, when assessing individuals suspected of having SLD, the authors recommend use of the CHC model over the Luria model because Gc is an important aspect of cognitive performance (Kaufman & Kaufman, 2004) and this factor is often a relative strength for individuals with SLD. Results from the KAIT and KABC-II can help identify an individual’s strengths and weaknesses in cognitive ability and mental processing, making them valuable tools for identifying strengths in processing, as well as basic psychological processing disorders, a key aspect of SLD definitions (Kaufman & Kaufman, 2004).

**ROLE OF COMPREHENSIVE EVALUATIONS**

Although school districts may become immersed in the use of RTI as part of the evaluation procedure, elimination of the discrepancy requirement and use of RTI as part of the evaluation process do not diminish the need for comprehensive evaluations or lead to the demise of clinical assessments. Although the conflict between intelligent testing and “stupid
decision-making procedures” may not be resolved soon, individual clinical tests of intelligence will always survive because of their diagnostic utility (Kaufman, 2000). These evaluations focus upon problem solving for an individual. When conducted by a skilled clinician who integrates actual classroom performance, educational history, and behavioral observations, comprehensive evaluations provide important information to help explain the nature and severity of a learning disability, as well as the various treatment options.

A comprehensive, individualized assessment includes an assessment of cognitive processes, academic performance, social-emotional functioning, and the environmental factors affecting performance (Kavale et al., 2005). Kaufman et al. (2005) outlined the three phases involved in a comprehensive evaluation. In the first phase, emotional, cultural, social, or physical factors related to the learning difficulty are ruled out; in the second phase, the cognitive and processing strengths and weaknesses of the individual are identified; and in the third phase, evidence-based interventions directly related to the child’s strengths and weaknesses are prescribed. Based upon the diagnostic conclusions, the evaluator selects specific interventions to meet the identified needs. For students with SLD, differential instruction that addresses the source of the problem is far more effective than global, generalized approaches that do not (Aaron, 1997; Gunnison et al., 1982; Kavale et al., 2005; Mather & Jaffe, 2002). In addition, taking full advantage of identified cognitive strengths such as strong verbal comprehension can help to justify critical accommodations and support and clarify the mode and methods of instructional intervention.

As noted, one facet of comprehensive, clinical evaluations is that all conclusions from the test profiles are supported with data from multiple sources (Kaufman, 1994). The evaluator relates observations of behavior to the test profile and integrates information from a variety of instruments and observations to support or clarify the diagnostic hypotheses (Kaufman et al., 1999; Lichtenberger et al., 2004). Because the obtained scores may not have a unitary explanation and children can obtain low scores for a variety of reasons, an evaluator must corroborate information with a “judicious selection” of other tests and then integrate the referral information, clinical observations of test-taking behaviors, detailed analysis of errors, and an analysis of the profile fluctuations (Kaufman, 1994).

As a case example, results from an intelligence test indicated that Michael, a fourth-grade boy, had advanced reasoning abilities and acquired knowledge, but slow performance on all timed tasks. During testing, he worked slowly on a task where he was asked to match symbols
quickly and on tasks involving the rapid naming of objects, digits, and letters. His general education teacher reported that Michael did not have behavioral, emotional, or social problems, but that he did have a slow reading rate, as well as difficulty completing most assignments in a timely manner. Additional academic testing substantiated Michael’s slow reading rate.

In consideration of his strengths in reasoning and knowledge contrasted with his slow rate of performance and production, the evaluator made recommendations for accommodations that included shortened assignments and extended time on tests, as well as recommendations for the implementation of instructional methods designed to increase reading rate and fluency. In addition, a recommendation was made for supplemental testing in the form of oral follow-up for missed or unfinished items on examinations to ensure that his less efficient reading skills did not affect comprehension of the question(s) even though extended time was granted.

All pieces of information were used to construct a diagnostic hypothesis, confirm it, and then select appropriate accommodations and interventions. In addition, the intact cognitive and linguistic abilities, the processing weakness, and the slow reading speed, all supported the diagnosis of a specific reading disability.

CONCLUSION

Some professionals have suggested that intelligence tests are outdated and are no longer needed in the field of SLD because the results have limited value (e.g., Gresham, Restori, & Cook, 2008). Siegel (1989) described the intelligence test as a “sacred cow that the LD field is not willing to abandon” (p. 518). More recently, Reschly (2005) maintained that the focus should be solely on intervention and not on “... internal child attributes that rarely have significant implications for special education interventions” (p. 512). Sole use of RTI is focused on the identification of weaknesses, and not, as a good cognitive assessment does, on the documentation of strengths as well. In this regard, an “intrinsic” child characteristic such as above-average verbal abilities has significant educational relevance. Yet if only RTI methods were used, the need and justification for emphasizing verbal methods of instruction and providing oral testing on classroom assessments of subject-area knowledge might be deemphasized or overlooked. To think that you can diagnose SLD from RTI is just simply wrong (Kaufman et al., 2007).
The history of the SLD field and the present suggest that we should challenge practices, but not the SLD construct (Hale et al., 2006). The definition of SLD requires the identification of a processing deficit which is the key aspect of contemporary SLD assessment (Kaufman et al., 2007). For accurate SLD identification, cognitive testing is a necessary and productive part of a comprehensive evaluation. Intelligence tests can also be used to help increase understanding and inform intervention when emphasis is placed on understanding the educational implications (Kaufman & Kaufman, 1983b; Mather & Jaffe, 2002; Naglieri & Pickering, 2003). Even in light of new regulations, it is highly unlikely that the demand for evaluations will diminish. If school districts cut back on evaluations, the practice of private clinicians will flourish. Kaufman et al. (2005) noted: “... there is a demand for the comprehensive assessment to drive intervention. This is the way it has always been, and this is the way it will always be because the referral questions for children with SLD have always asked, What is wrong? And how can we help? These questions demand differential diagnosis, a large part of which is determined by the cognitive abilities present in the individual child” (p. 211). We must also ask: What is right? And how can we use the spared and intact cognitive abilities to help (Kaufman et al., 2005)? Assessments can aid us in understanding a person’s information-processing capabilities, including the factors that facilitate performance. Gardner (1999) clearly made this point when he suggested, “We shouldn’t ask how smart you are, but rather how are you smart?”

Thus, the goals of intelligent testing for individuals suspected of having SLD are to identify a person’s strengths and weaknesses, attempt to understand the relationship among these abilities, and then translate these findings into meaningful intervention plans. Within these plans, instruction is individualized; different approaches are used for different people; strengths and weaknesses are considered in selecting accommodations and methodologies; and the effectiveness of the accommodations and instruction is evaluated frequently and altered as needed. The examiner is intimately involved and an integral part of all facets of the test administration and interpretation process (Kaufman, 2000). The central goal of intelligent testing of individuals with SLD is not, however, eligibility, but rather using the results to help a person improve his or her educational and life outcomes (Kaufman, 1981; Kaufman, 2000; Kaufman et al., 2007). As Kaufman (1979) stated: “The focus is the child ...” (p. 1). After all, that is what it’s all about.
ACKNOWLEDGMENT

The author wishes to express her gratitude to Dr. Michael E. Gerner for editorial suggestions on the first draft of this chapter.

REFERENCES


Kaufman, A.S. (2004). Standardized cognitive assessment and the new IDEA guidelines- Fit or misfit? Presentation at the National Association of School Psychologists Annual Convention, April 1, Dallas, TX.


Willis, J.O., & Dumont, R.P. (2002). *Guide to Identification of Learning Disabilities* (3rd ed.). Peterborough, NH: authors. Available from authors: print copy from johnzerowillis@yahoo.com or CD from dumont@fdu.edu
