

CHAPTER 14

Please use the following reference when citing this work:

McGill, R. J., Styck, K. M., & Dombrowski, S. C. (2020). Learning disabilities. In E. A. Youngstrom, M. J. Prinstein, E. J. Mash, & R. A. Barkley (Eds.), *Assessment of disorders in childhood and adolescence* (5th ed., pp. 374-409). Guilford Press.

Chapter 14

Learning Disabilities

Ryan J. McGill¹, Kara M. Styck², and Stefan C. Dombrowski³

¹ Department of School Psychology and Counselor Education, William & Mary

² Department of Psychology, Northern Illinois University

³ Department of Graduate Education, Leadership, and Counseling, Rider University

Author Note

Please address all correspondence concerning this article to Ryan J. McGill, William & Mary School of Education, P. O. Box 8795, Williamsburg, VA 23187. E-Mail: rmcgill@wm.edu

CHAPTER 14

Learning disability (LD [a.k.a. specific learning disorder, or SLD]) is an umbrella term that describes a condition in which individuals present with low achievement that cannot be explained by other psychological disorders or environmental factors. Despite much persistence and effort, LD remains difficult to demarcate, and considerable controversy has been associated with each attempt to operationally define the construct over the last 50 years. As a result, clinicians are forced to navigate an array of complex, and in some cases, conflicting statutes, regulations, and classification criteria when attempting to identify LD in children and adolescents. As an example, a client may meet the all of the diagnostic criteria for SLD in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013) yet be denied special education and related services in that same category under local educational regulations. Further complicating the matter are the numerous classification approaches that pervade professional practice. While scholars continue to debate the merits of LD assessment methods and what new approaches may portend for the field, practitioners must adhere to state and federal guidelines, and, in some cases, implement new approaches that have yet to be empirically validated (Kavale, Kauffman, Bachmeier, & LeFever, 2008; McGill, Styck, Palomares, & Hass, 2016). Apropos of this dilemma, we begin this chapter by outlining the salient diagnostic features of LD and describe the history of approaches to LD assessment and identification before proceeding to elaborate on contemporary assessment methods and classification approaches.

Diagnostic Features

Within DSM-5, LDs fall under the broader category of neurodevelopmental disorders and is characterized by focal impairments in academic learning that limit the acquisition and performance of academic skills. Epidemiological surveys indicate that approximately 5–15% of school-age children present with an LD (Moll, Kunze, Neuhoff, Bruder, & Schulte-Korne, 2014).

CHAPTER 14

Reading disorder is the most researched and most common variant of LD, and the vast majority of assessment and intervention research has been focused in this area. Evidence suggests that approximately 70–80% of individuals with LDs have primary deficits in reading (Ferrer, Shaywitz, Holahan, Murchione, & Shaywitz, 2010).

Although conceptualizations of LD may differ across clinical settings, all operational LD definitions share the same fundamental assumption that LDs reflect *unexpected* underachievement (Beaujean, Benson, McGill, & Dombrowski, 2018). Unlike other psychological disorders, LD represents both a clinical condition and an educational policy category. As a result, the makeup of the population of individuals diagnosed with LD may vary considerably when applicable laws and regulations change (Lewandowski & Lovett, 2014). As an example, if a school district uses one assessment method for classification and a neighboring school district uses another, an individual may be classified as having LD in the former district but upon moving to the later district, their eligibility for services may be removed for failure to meet the different criteria that is employed in that particular jurisdiction (Miciak, Fletcher, Stuebing, Vaughn, & Tolar, 2014). As we discuss later, even when different jurisdictions employ the same assessment method or model, identification practices may still fluctuate because of nontrivial differences in how various features of that model are operationalized. It is important to note that these issues are not limited to educational settings, as clinicians in private practice may employ different assessment and identification methods due to the opaque diagnostic criteria presently contained in DSM-5.

Changes in DSM-5

Numerous terms have been used to describe LD in previous DSM editions, including “learning disturbance” and “academic skills disorders.” The current version of DSM-5 uses the term “specific learning disorder,” which it operationally defines as “difficulties learning and using academic skills . . . that have persisted for at least 6 months, despite the provision of

CHAPTER 14

interventions that target those difficulties” (American Psychiatric Association, 2013, p. 66). Subtypes demarcate impairment in three academic domains: reading (marked by deficits in word-reading accuracy, reading fluency, and reading comprehension), written expression (marked by deficits in spelling, grammar, and clarity and organization of writing), and mathematics (marked by deficits in number sense, memorization of math facts, calculation, and math reasoning). Whereas DSM-IV listed a nonspecified subtype of SLD (LD not otherwise specified [LD NOS]), this subtype has been omitted from the current DSM edition. As with other psychological disorders, clinicians are also required to grade the severity of symptom presentation (i.e., mild, moderate, or severe).

Despite the fact that assessment of intellectual functioning has historically played a prominent role in many of the LD identification assessment models developed since the 1960s, the potential value of cognitive assessment for diagnosis is downplayed in DSM-5. It is noted that “individuals with specific learning disorder typically (but not invariably) exhibit poor performance on psychological tests of cognitive processing. However, it remains unclear whether these cognitive abnormalities are the cause, correlate, or consequence of the learning difficulties” (American Psychiatric Association, 2013, p. 70). Consequently, evidence for the presence of a cognitive processing disorder or a significant discrepancy between measured IQ and achievement are no longer required for diagnosis. Instead, the core diagnostic feature in DSM-5 is the presence of academic dysfunction that is resistant to remediation. This change has been a source of significant criticism within the professional literature, as it seemingly moves the diagnostic criteria from what was previously an ability–achievement discrepancy model to one that is largely based on low achievement and/or response to intervention (RTI),¹ a move that is

¹DSM-5 and ICD-10 (World Health Organization, 1992) do not presently endorse any particular assessment method for LD identification.

CHAPTER 14

consistent with current professional practice trends (Cavendish, 2013).

In spite of these developments, cognitive discrepancy methods are still widely utilized in school-based settings, and nothing explicitly prohibits their use in DSM or in federal regulations (Maki, Floyd, & Roberson, 2015). Given the current absence of a consistent diagnostic approach, Schroeder, Drefs, and Cormier (2017) note that it is not uncommon for clinicians to rely on clinical judgment and elect to disregard or be lenient in their application of a particular assessment approach. As should be evident, it is imperative for practitioners, regardless of the setting in which they are working, to be cognizant of the regulations (i.e., state and federal educational codes) and assessment practices that are prevalent in their particular jurisdiction, in order for their assessment results and diagnostic recommendations to have ecological validity across various clinical settings.

Assessment and Identification: Historical Context

Many authorities credit Samuel Kirk for originating the term “learning disability.” In the original edition of *Educating Exceptional Children*, Kirk (1962) provided one of the first operational definitions of LD, on which he explicated a year later in a speech delivered at the annual meeting of the organization that eventually became the Learning Disabilities Association of America. According to Hallahan and Mercer (2002), this was the first time that the term was used at an educational conference. The importance of these events cannot be overstated, as they gave rise to a decade-long series of events that led to federal involvement in LD and the eventual passage of the Education for All Handicapped Act (Public Law 94-142) in 1975. This act contained a federal operational definition for LD that continues to be used to this day. The pace of adoption in clinical practice was much faster, as the original DSM published in 1968 contained a preliminary category (learning disturbance) that eventually morphed into SLD.

Although federal involvement in LD has only been evident since the 1960s, scholars have

CHAPTER 14

traced the roots of LD all the way back to the early 1800s, and much of the assessment logic that clinicians employ today can be sourced to this “prefoundational” period; that is, many assessment and identification approaches were conceived of prior to the development and eventual acceptance of formal operational definitions of the construct. As a result, typical LD assessment practices have not kept pace with advances in evidence-based practices (Fletcher, Stuebing, Morris, & Lyon, 2013).

Prefoundational Period

The origins of the LD field can be traced to a series of case study publications by prominent physicians and researchers in early 19th-century Europe. Most of these investigations were conducted with adult patients who had sustained traumatic brain injury resulting in peculiar language deficits. Although limited by the technology of the time, seminal discoveries during this era continue to serve as points of reference for the field. For example, Joseph Gall was one of the first individuals to explore relationships between brain injury and impairment. Based on these observations, Head (1926) concluded that each of the “intellectual qualities of the mind” (p. 4) was localized in a different portion of the brain. These findings gave rise to the practice of phrenology—an approach in which the shape and size of an individual’s cranium were evaluated in order to make inferences about his or her character and mental abilities. Although phrenology was soon disavowed by the medical and psychological communities, the localization of function idea (i.e., that specific cognitive tests can be used to make inferences about potential deficits in focal areas of the brain) continues to undergird many LD assessment methods (McGill & Busse, 2017).

It is interesting to note that early assessment approaches focused mostly on using assessment to inform intervention by administering rudimentary reading and language tasks to individuals with reading difficulties (i.e., direct measures of academic skills). Clinical

CHAPTER 14

evaluations focused mostly on academic strengths and weaknesses and identifying the component skills that were in need of remediation. For example, Monroe (1932) developed a “reading index” that calculated the discrepancy between actual and expected reading achievement for a student, and it was thought that this index could be used to identify students in need of special assistance. Additionally, she was an early proponent of using error analysis to help guide treatment selection, and she encouraged practitioners to give equal consideration to a child’s quantitative and qualitative performance on academic achievement tasks. With the advent of IQ tests, the administration of these measures soon became routine in case studies that emerged in the early 20th century. Initially, the relative value of IQ tests for the diagnosis or remediation of reading disability was called into question. Orton (1925) noted that “the test [IQ test] is inadequate to gauge the equipment in a case of such a special disability” (p. 584).

Developments in psychological assessment and individual-differences research led to a paradigm shift in LD assessment in the mid-20th century. Principle among these were Samuel Kirk’s research program, which stressed the importance of intraindividual differences and assessment-based instruction, and the rise of commercial IQ tests measuring multiple factors that were thought to reflect different cognitive attributes. Whereas IQ tests were given modest consideration by early LD practitioners, they soon became prominent features of emerging LD classification models.

After obtaining his doctorate from the University of Michigan, Kirk obtained a faculty position at the University of Illinois and established an experimental preschool for children with intellectual disabilities. In order to better educate these children, Kirk believed that he needed assessments that could amplify relevant psychoeducational strengths and weaknesses. Put simply, the goal was to develop a series of diagnostic tests that would be useful for instruction. This experimental research culminated in the development of the Illinois Test of Psycholinguistic

CHAPTER 14

Abilities (ITPA; Kirk, McCarthy, & Kirk, 1961). The ITPA contained 12 subtests measuring perceptual, linguistic, and memory abilities, and it was believed that particular ITPA profiles could be matched to training activities that were developed in conjunction with the instrument. A critical goal of the ITPA research program was to identify relevant aptitude-by-treatment interactions (ATIs), the notion that some instructional strategies are more or less effective for particular individuals depending on their abilities.

Although use of the ITPA was widespread in the 1960s and 1970s, numerous critiques of the instrument's psychometric properties and efficacy of its training procedures emerged (e.g., Hammill & Larsen, 1974; Mann, 1979) and it soon fell out of favor. On the basis of 18 years of largely unsuccessful research, Cronbach (1975) concluded that the search for ATIs was plagued by the fact that "once we attend to interactions, we enter into a hall of mirrors that extends to infinity" (p. 119) and that scientific psychology would be better served by eschewing the correlational approach in favor of short-run empiricism. Nevertheless, Hallahan, Pullen, and Ward (2013) argue that the development of the ITPA remains historically important because it reinforced the notion that children with LDs have important individual differences and that psychoeducational assessment may be used to guide instruction. Despite their intuitive appeal, few ATIs have been empirically validated for children with LD since Cronbach's seminal critique, and the ATI validation remains a topic of controversy within the field (see Burns et al., 2016).

Origins of the Ability–Achievement Discrepancy Model

Although the term "learning disability" was introduced in the early 1960s, existing operational definitions are relatively opaque and shed little insight on how LD should be assessed or identified. This began to change when Barbara Bateman (1965) expanded on Kirk's definition and offered the following definition for LD:

CHAPTER 14

Children who have learning disorders are those who manifest an educationally significant discrepancy between their estimated potential and actual level of performance related to basic disorders in the learning process, which may or may not be accompanied by demonstrable central nervous system dysfunction. (p. 220)

Bateman's definition served as a critical inflection point in the field, as it reinforced Monroe's (1932) earlier notion that a discrepancy between achievement and potential may be used to formally diagnose students with LD and, more importantly, that cognitive tests are vital for assessment and identification. As a result of this development, discrepancy soon became linked with the identification of LDs and a focal point of subsequent modifications to the operational definitions and regulations produced by federal task forces.

As previously mentioned, with the passing of Public Law 94-142 in 1975, the federal government produced an operational definition for SLD that has subsequently undergone only slight modifications and has additional implications for LD classification in school-based settings. That definition, as quoted in the final regulations adopted by the U.S. Office of Education (USOE) in 1977 is as follows:

The term "specific learning disability" means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken, or written, which may manifest itself in an imperfect ability to listen, speak, read, write, spell, or do mathematical calculations (p. 65083).

Given the integral role of "psychological processes" in the federal definition, it is frequently asserted that LD assessment methods should focus on identifying relevant processing strengths and weaknesses (PSWs; Hale et al., 2010). However, in an interesting paradox, more recent federal regulations seem to deemphasize processing assessment entirely (Lichtenstein, 2014). Nevertheless, as this definition implies that a disorder in cognitive processing is the putative

CHAPTER 14

cause of LD, many assessment methods (e.g., PSW approaches) focus exclusively on documenting the presence of a processing deficit, and practitioners have long been encouraged to evaluate ability profiles for unique patterns and signs that portend to have implications for the presence of LD (Fletcher et al., 2013).

Prior to Public Law 94-142 implementation in 1977, the USOE (1976) furnished additional regulations pertaining to the identification of individuals with LD and proposed a formula that defined a severe discrepancy as “when achievement in one or more of the areas falls at or below 50% of the child’s expected achievement level” (p. 52405). Public response to the formula was overwhelmingly negative (Hallahan & Mercer, 2002). Although the USOE continued to endorse the idea of an ability–achievement discrepancy in the regulations, no formula was included, giving each state the power to determine how the discrepancy model would be operationalized within its borders. The discrepancy model soon became the dominant method for conferring an educational classification of SLD in a most states, as well as a prominent element in the diagnostic criteria outlined in the DSM.

What is particularly striking about the process that led to the reification of the discrepancy model is the limited empirical evidence available at that time to support use of the method. The only compelling evidence validating the IQ discrepancy hypothesis that could be located came from the Isle of Wight epidemiological studies conducted by Rutter and Yule (1975), who administered the Performance IQ scale from the Wechsler Intelligence Test for Children (WISC) and additional reading measures to a large sample of children. Using a regression-based definition, they found that they were able to distinguish between “specific reading retardation [*sic*]” and “general reading backwardness,” with the bulk of the children in the former group producing reading scores that were two standard deviations or more below their IQ scores. This resulting “hump” was interpreted as a useful cutoff point for distinguishing

CHAPTER 14

between children with LDs and “garden variety” poor learners. However, subsequent attempts to replicate these findings have been unsuccessful (e.g., Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992) and the discrepancy model has been maligned since its inception (Aaron, 1997).

Given the widespread dissatisfaction with the discrepancy model and additional concerns by some researchers with the field’s perceived devotion to processing assessment and cognitive process training (see Mann, 1979), the USOE funded five LD research institutes to identify and advance research-based assessment and intervention practices. The institutes were housed at the University of Illinois at Chicago, the University of Kansas, the University of Minnesota, the University of Virginia, and Columbia University. In spite of making substantive advances in the development of curriculum-based assessment procedures and the identification of numerous empirically supported treatments, researchers were unable to come to a consensus on an alternative assessment method to replace the discrepancy model. In summarizing the results of 5 years of research findings produced by the institute at the University of Minnesota, Ysseldyke and colleagues (1983) noted, “After five years of trying, we cannot describe, except with considerable lack of precision, students called LD. We think that LD can best be defined as ‘whatever society wants it to be, needs it to be, or will let it be’ at any point in time” (p. 89). Whereas this may be viewed by some as an overly draconian account of the state of affairs at that time, we suspect that there are likely many clinicians who believe this quote is as accurate today as it was 25 years ago.

The Search for Diagnostic Signs and Unique PSWs

Although dissemination of the discrepancy method was instrumental in popularizing the use of IQ tests for LD assessment and identification, this was not the first attempt to utilize cognitive tests in this capacity. It has long been speculated that cognitive test scatter and variability could serve as a potential pathognomonic sign for a host of psychological disorders,

CHAPTER 14

including LD. Early researchers hypothesized that subtest scatter would predict scholastic potential and membership in exceptional groups (Harris & Shakow, 1937), and formal methods for these types of analyses have been proposed in the literature for well over 70 years. Rapaport, Gil, and Schafer (1945) developed a formal process for evaluating cognitive scatter in a two-volume series devoted to diagnostic testing. Their system involved graphically plotting subtest profiles, then visually inspecting the peaks and valleys in an examinee's scores and generating pathognomonic inferences from these observations. Given the intuitive nature of these procedures, they soon became a staple of clinical tradition. As tests expanded, clinicians were provided with more scores and score comparisons to interpret, and questionable interpretive practices emerged and remained popular through the 1970s (Kaufman, Raiford, & Coalson, 2016).

As a remedy, Kaufman (1979) proposed a step-by-step profile analysis interpretive approach for the WISC-R² that he termed “intelligent testing.” According to Kaufman, Raiford, and Coalson (2016), Kaufman was motivated by a need to “impose some empirical order on profile interpretation; to make sensible inferences from the data with full awareness of errors of measurement and to steer the field away from the psychiatric coach” (p. 7). In the intelligent testing approach, practitioners are encouraged to interpret test scores in a systematic fashion, beginning with the full scale intelligence quotient (FSIQ) and culminating at the subtest level. However, users are encouraged to focus most of their interpretive weight on the scatter and elevation (i.e., strengths and weaknesses) that are observed in lower-order scores (e.g., subtest and broad ability composite and index scores), and interpretation of the FSIQ is deemphasized. Inferential hypotheses are then generated from these observations, as well as the qualitative

²The intelligent testing approach can be used with other measures, and the levels-of-analysis approach to test interpretation is featured in virtually every test technical manual and clinical guidebook.

CHAPTER 14

behaviors observed during the test administration.

Consonant with the publication of the WISC-III, Kaufman (1994) produced a revision of the so-called “Kaufman method,” outlining several schemes based on different configurations of Wechsler subtests that were thought to be useful for the diagnosis of LD. For instance, Kaufman noted that individuals with disabilities tended to score lower on the subtests comprising the SCAD profile (Symbol Search [S], Coding [C], Arithmetic [A], and Digit Span [D]). Additional profiles included the ACID profile, Bannatyne (1968) pattern, and the Learning Disability Index (LDI). Although it is common to find abnormal scatter and patterns of PSWs within clinical groups, the uniqueness of these differences tends to evaporate whenever normal controls are included in the samples due to the large amount variability that is endemic in the population (Zimmerman & Woo-Sam, 1985). Empirical research studies have consistently indicated that the diagnostic accuracy of these profiles rarely exceeds chance levels, rendering them ineffectual for LD classification (e.g., Smith & Watkins, 2004; Watkins, Kush, & Glutting, 1997a, 1997b; Watkins, Kush, & Schaefer, 2002).

Another popular heuristic emerging out of the Kaufman tradition is the hypothesis that children with LD are characterized by significant Verbal IQ–Performance IQ (VIQ–PIQ) discrepancies, and these types of composite score pairwise comparisons have long been a core feature of the Kaufman interpretive approach. This notion stems from belief that such discrepancies reflect underdevelopment in focal areas of the brain or neural circuits that may have implications for matching instruction to a client’s learning style (Elliott & Resing, 2015). Beyond the fact that the learning styles concept has been the subject of significant research criticism, the evidence base for matching students with learning difficulties to effective instruction based on the information furnished by cognitive tests is less than compelling (Burns et al., 2016; Cronbach & Snow, 1977; Fletcher & Miciak, 2017; Pashler, McDaniel, Rohrer, &

CHAPTER 14

Bjork, 2008). Furthermore, in a comprehensive meta-analysis of 94 studies, Kavale and Forness (1984) found that the mean VIQ–PIQ difference in individuals that were diagnosed with LDs was only 3.46 points. Based on this finding they concluded that 79% of the population was likely to exhibit the same or greater discrepancy between those scores. As a result, they concluded, “V-P differences appear to be of little value in LD diagnosis” (p. 139).

Additional studies have illuminated psychometric limitations of profile analysis that may have implications for the application of these procedures for the diagnosis of LD. McDermott, Fantuzzo, and Glutting (1990) surveyed the extant literature on intraindividual and interindividual subtest analysis and concluded that there was little empirical support for interpretation of these metrics. These findings that have since been replicated (McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992; Watkins, 2005). As an example, Watkins tested the diagnostic validity of four different configurations of WISC-III subtest scatter, and receiver operating characteristic curve (ROC) analyses revealed that the scatter indices correctly diagnosed preidentified LD only 50–55% of the time. The results produced by Macmann and Barnett (1997) may be instructive for understanding why psychometric researchers investigating the diagnostic utility of profile analysis methods have consistently obtained negative research results. Macmann and Barnett used computer simulations to measure the impact of measurement error on the reliability of cognitive profile analysis interpretations (e.g., pairwise comparisons, PSWs, and scatter) and found that 62.4% of the sample presented with at least one significant composite score strength or weakness, and the base rate was even higher at the subtest level. As a result, they concluded that there was strong potential for clinical error and confirmation bias when attempting to ascribe *meaning* to these observations. Relatedly, Watkins and Canivez (2004) examined the temporal stability of PSWs and found low longitudinal agreement. As a consequence, they suggested that the inferences generated from these patterns of scores are likely

CHAPTER 14

to be unreliable.

More recently, McGill (2018) utilized the evidence-based assessment framework outlined by Youngstrom and Van Meter (2016) to evaluate whether cognitive scatter accurately discriminated between individuals predetermined to have LD and non-LD controls in a nationally representative normative sample ($N = 2,025$). Diagnostic efficiency statistics revealed that increasing levels of scatter did not function as a useful diagnostic sign. Area under the curve (AUC) values ranged from .46 to .51, indicating that cognitive scatter, at best, functioned as the diagnostic equivalent of flipping a coin, and positive predictive values (.04 to .05) were even more dismal. As a result, probability nomograms for each level of scatter indicated that scatter assessment did not improve the posterior odds of correct diagnosis from prior base rates. Based on these results, it appears that practitioners who engage in scatter and other, related PSW analyses may be spending a significant amount of time and resources to obtain information that may not be clinically useful.

Nevertheless, Kaufman (1994) has long argued that these limitations are managed by clinical acumen. For example, Kaufman and Lichtenberger (2006) argue that validity studies using group data (e.g., Macmann & Barnett, 1997) may obscure important individual differences; thus, clinicians should use their professional judgment to discern when these results are applicable when interpreting the assessment data obtained from individuals; that is, it is possible for a child to have a LD even when specific pattern or sign hypothesized to reflect that disorder is not present and clinicians may be able to discern when this phenomenon may occur at the level of the individual through skilled detective work. In spite of a long-standing body of literature recommending that these practices be eschewed (e.g., Bray, Kehle, & Hintze, 1998; Canivez, 2013; Glutting, Watkins, & Youngstrom, 2003; Kranzler et al., 2016a; Watkins, 2000, 2003), surveys reveal that these procedures remain a core staple of LD assessment training and practice

CHAPTER 14

(Benson, Floyd, Kranzler, Eckert, & Fefer, 2018; Sotelo-Dynega & Dixon, 2014).

Decline of the Discrepancy Model and the Rise of RTI

By the 1990s, the majority of states had adopted some variant of the discrepancy method as part of their identification procedures (Hallahan & Mercer, 2002). However, empirical investigations began to identify considerable conceptual and psychometric problems with the method. For example, in a meta-analysis, Stuebing and colleagues (2002) evaluated the cognitive correlates of poor reading groups containing individuals with significant discrepancies and those that were not discrepant. Results indicated that there was significant overlap between the two groups on the core cognitive skills that were most associated with reading, indicating that those dimensions did not discriminate between children with and without discrepancies.

The discrepancy model has also been criticized as a “wait to fail” approach due to the fact that younger children, who are referred for an evaluation, often do not present with a discrepancy that is large enough to meet diagnostic criteria for LD (Restori, Gresham, & Cook, 2008). Accordingly, these students must continue to fail until their achievement is sufficiently low compared to their IQ. To buttress this claim, critics point to surveys indicating that identification rates tend to peak between third and fourth grade, even though the manifest symptoms of academic dysfunction are often apparent much earlier (Lyon, Fletcher, Fuchs, & Chhabra, 2006). Of concern, some “gold standard” treatments may begin to lose their effectiveness by the time these children are identified (Wanzek et al., 2013). This limitation also appears to have impacted the integrity of identification practices in some jurisdictions. MacMillan, Gresham, and Bocian (1998) conducted an audit of students who were identified as LD in a California school district employing the discrepancy model and found that less than half of students identified as LD met the regulatory criteria for eligibility. These findings are not intrinsic to California, and similar results have been obtained in numerous jurisdictions since the imposition of the discrepancy

CHAPTER 14

model in the 1970s. As noted by Peterson and Shinn (2002), school personnel tend to base eligibility decision on an “absolute low achievement” criterion, even at the expense of the law.

By the turn of the century, resistance to the discrepancy model reached a crescendo, culminating in a Learning Disabilities Summit of leading researchers in 2001. A majority of the white papers that were presented argued for an alternative (RTI) to the discrepancy model that focused on providing targeted interventions to children with academic difficulties and using low-inference progress monitoring assessments to evaluate treatment outcomes (Otaiba, Wagner, & Miller, 2014). It is interesting to note that much of this effort grew out of the assessment and intervention research that begun at the federal LD institutes in the 1970s. By the turn of the century, there was abundant evidence to indicate that the discrepancy model did not predict treatment response between difficult-to-remediate and readily remediated children (e.g., Vellutino et al., 1996) and that children with reading difficulties were able to benefit from targeted reading interventions, in particular, remedial phonics-based instruction.

In 2004, the Individuals with Disabilities Education Act (IDEA) was reauthorized, and the regulations pertaining to the identification of SLDs were modified, based largely on the findings and recommendations produced from the 2001 Summit. The final regulations authored by the Office of Special Education and Rehabilitation Services (OSERS) specify that use of the discrepancy method is no longer required (but is not prohibited), that states must allow for the use of RTI procedures for identification, and that states are free to adopt a *research-based* alternative to discrepancy and/or RTI. Over the course of the last decade, numerous states and jurisdictions (some states permit local educational agencies to adopt their own local LD identification criteria) have implemented RTI models and procedures, and the dissemination of these methods has been a prominent topic within the professional literature.

The debate that produced the RTI “consensus” was not without controversy. Several rival

CHAPTER 14

models that emphasized the importance of a processing-based approach for diagnosis were presented and considered at the Summit. Critics argue that the RTI model is not consistent with the federal operational definition of LD, as its assessment procedures do not yield information that is useful for identifying a deficit in cognitive processing (e.g., Hale, Kaufman, Naglieri, & Kavale, 2006). Emerging out of the radical behavior tradition, RTI represents a profound paradigm shift in which assessment is focused exclusively on treatment utility, and cognitive testing is given little, if any, consideration in the diagnostic process outside of ruling out the presence of intellectual disability. Scruggs and Mastropieri (2002) suggest that such radical alterations to identification criteria are akin to “throwing the baby out with the bath water” (p. 155). Since 2000, a number of rival processing-based approaches for assessment and identification (e.g., PSW) have been proposed in the literature.

In general, the PSW approach is essentially a cognitive discrepancy method. According to McGrady (2002), “The discrepancy comparisons are made *among* apparently discrete skills and abilities, not as a comparison of such skills with performance on a general estimate of intellectual potential—the traditional discrepancy formula approach” (p. 628). Although these models were debated and ultimately found wanting at the Learning Disabilities Summit, enthusiasm for their implementation by practitioners and scholars has increased over the last 5 years. As a result, LD identification practices vary considerably across states. In a survey, Maki and colleagues (2015) found that 67% of states continue use the discrepancy method, 16% require the use of RTI procedures as a primary method of LD identification, and 28% now permit use of a PSW approach.

Summary

Although there is considerable agreement about the definition of LD as a construct, a consensus method for identifying LD has been elusive. As a result, it is important for

CHAPTER 14

practitioners to understand the history and context that has led to the development of various classification models and the perceived weaknesses for each of these approaches. Ironically, the very method that has long dominated the field (discrepancy) was adopted by default when researchers and practitioners were unable to agree on a suitable alternative (Dombrowski & Gischlar, 2014). The field has essentially been on a path of course correction ever since.

Presently, these debates largely focus on the role of cognitive testing in assessment and diagnosis and whether any approach can *validly* and *reliably* distinguish between individuals with LD and those with broader academic dysfunction (i.e., “garden variety” poor learners). As should be evident in this review, these issues are not new and have plagued the field since its inception. In the following sections of this chapter, we outline and describe commonly used assessment tools and contemporary classification schemes in more detail.

Assessments Commonly Used in LD Classification

Contemporary surveys (e.g., Sotelo-Dynega & Dixon, 2014) reveal that norm-referenced tests of intelligence and achievement, and curriculum-based measurement are the most commonly used measures across settings for the assessment of LD. In this section we provide examples and describe each of these assessment types in more detail. It should be noted that even in classification models that employ a cognitive discrepancy approach, a comprehensive evaluation should include additional elements that are common to other areas of psychopathology (e.g., review of records, diagnostic interview, direct behavioral observations) to rule out relevant exclusionary factors or to examine other areas of psychological functioning that may contribute to a child’s learning problems. For example, behavioral rating scales may be used to screen for social–emotional dysfunction or conduct differential diagnosis for relevant comorbid disorders such as ADHD. As these applications are described in more detail in other areas of this text, we defer elaborating on them further here.

CHAPTER 14

Wechsler Intelligence Scale for Children—Fifth Edition

The Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V; Wechsler, 2014) is the latest version of one of the most frequently used intelligence tests for children. It includes 16 subtests; five factor index scores (Verbal Comprehension, Visual–Spatial, Fluid Reasoning, Working Memory, and Processing Speed); and a hierarchically ordered FSIQ score. Users can also use various combinations of secondary and complimentary subtests to produce a series of ancillary index scores (General Ability, Cognitive Proficiency, Quantitative Reasoning, Nonverbal, and Auditory Working Memory). However, these scores are not derived from factor analysis; instead, they are logically or theoretically constructed and should be interpreted with caution.

The WISC-V is a substantial revision of the previous version of the instrument. The Word Reasoning and Picture Completion subtests were eliminated, and several new subtests were added. The Picture Span subtest (adapted from the Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition [WPPSI-IV]) was added to measure visual working memory, and the Visual Puzzles and Figure Weights subtests (adapted from the Wechsler Adult Intelligence Scale, Fourth Edition [WAIS-IV]) were added to measure visual–spatial and fluid reasoning, respectively. A major goal of the WISC-V revision was to split the former Perceptual Reasoning Index (PRI) into separate Visual–Spatial and Fluid Reasoning indices.

The technical manual suggests that primary index scores should be the main focus of clinical interpretation and that the patterns observed in these scores may be used to make logical inferences about PSWs that may have implications for LD assessment. In the descriptions of the subtests in the manual, there are descriptions of relations to cognitive and neuropsychological theories to facilitate inferences onto which narrow abilities these measures may map. As a result, the FSIQ (or a substitute ancillary index) may be used to identify children with a severe

CHAPTER 14

discrepancy, and the lower-order scores may be used to generate unique PSW profiles.

Kaufman Assessment Battery for Children—Second Edition

The Kaufman Assessment Battery for Children—Second Edition (KABC-II; Kaufman & Kaufman, 2004) measures the processing and cognitive abilities of children and adolescents between ages 3 and 18 years. According to the test authors, KABC-II underwent a major structural and conceptual revision. Eight subtests were eliminated from the original K-ABC, 10 measures were created and added to the current battery, and the KABC-II theoretical foundation was updated. The KABC-II utilizes a dual-theoretical foundation: Cattell–Horn–Carroll theory (CHC; Schneider & McGrew, 2018) and Luria’s neuropsychological theory of cognitive processing (Luria, 1966). One of the features of the KABC-II is the flexibility that it affords the examiner in determining the theoretical model to administer to the examinee. Although examiners may select either the Luria or CHC interpretive models, users are advised to interpret the KABC-II primarily from the CHC perspective.

The CHC model for school ages features 16 subtests (10 core and six supplemental), which combine to yield five first-order factor scale scores (Short-Term Memory [Gsm], Long-Term Storage and Retrieval [Glr], Visual Processing [Gv], Fluid Reasoning [Gf], and Crystallized Ability [Gc]), and a second-order full scale Fluid Crystallized Index (FCI) that is thought to represent psychometric *g* (general intelligence). Each CHC factor scale comprises two subtest measures, and the FCI is derived from a linear combination of the 10 core subtests that comprise the constituent factor scores. Although the KABC-II manual encourages a stepwise progression of interpretation from the FCI to the factor scores (consistent with the Kaufman method), clinicians are encouraged to use the CHC factor scores as the primary point of interpretation for the instrument. The Luria model omits measures of Crystallized Ability and thus features eight core subtests that combine to form an alternative full scale Mental Processing

CHAPTER 14

Index (MPI). If interpreting from the Luria perspective, the factor scores also employ a different nomenclature (Sequential [Gsm], Learning [Glr], Simultaneous [Gv], and Planning [Gf]). Given its versatility, it is suggested that it may be a particularly useful measure for examinee's who are culturally and linguistically diverse. However, an independent review by Braden and Ouzts (2005) suggests caution in employing the verbiage associated with the Luria model to interpret KABC-II scores, as it is psychometrically implausible for subtests to measure two distinct and theoretically divergent constructs simultaneously.

Kaufman Test of Educational Achievement—Third Edition

The Kaufman Test of Educational Achievement—Third Edition (KTEA-3; Kaufman & Kaufman, 2014) is an individually administered, norm-referenced test of achievement for individuals ages 4–25. It features 19 subtests (six core, 13 supplemental) that combine to form three subscale scores (Reading, Math, Writing), as well as an omnibus full scale score of overall academic achievement (Academic Skills Battery [ASB]). Various combinations of the core and supplemental tests can also be combined to form four reading-related composite scores (Sound–Symbol, Decoding, Reading Fluency, and Reading Understanding), two oral language composites (Oral Language and Oral Fluency), and four cross-domain composites (Comprehension, Expression, Orthographic Processing, and Academic Fluency). The measure provides both age- and grade-based normative scores which is a useful feature when assessing clients who may have been retained.

Another useful feature that is unique to the KTEA-3 is the incorporation of error analysis procedures for 10 of the 19 subtests. For example, when administering the Math Computation subtest, examiners can inspect the completed worksheet and complete a supplemental standardized protocol, highlighting specific reasons for a student's incorrect responses to individual items. In contrast to previous editions, base rates for these observations are now

CHAPTER 14

provided by the test publisher. Error analysis on the KTEA-3 was the subject of a special issue of the *Journal of Psychoeducational Assessment* in 2017 (Breaux, Bray, Root, & Kaufman, 2017).

In that issue, containing 13 articles and commentaries, preliminary evidence was presented, using the normative sample data for the KTEA-3, that error analysis patterns may be useful for discriminating among various LD subtypes.

Woodcock–Johnson IV

The Woodcock–Johnson IV (WJ-IV; Schrank, McGrew, & Mather, 2014) is a comprehensive battery of psychoeducational abilities for individuals ages 2–90 years. The measure contains 47 subtests allocated to three different batteries that allow for a comprehensive assessment of cognitive ability, achievement, and oral language. The battery was principally designed, and is the only commercial ability test to measure, all of the consensus factors posited by CHC theory (see Table 14.1). The WJ-IV presently serves as the preeminent reference instrument for making refinements to CHC and the understanding of cognitive–achievement relations in the psychological sciences. Each of the batteries is designed to be administered in isolation, although an examiner may elect to administer selected subtests from each battery utilizing a cross-battery assessment (XBA) framework. Interested readers are encouraged to consult Schrank and Wendling (2018) for an in-depth description of the panoply of scores afforded by the instrument. In addition to CHC-based factor scores, there are well over 20 broad and narrow cognitive–achievement clusters that can be calculated from different configurations of WJ-IV subtests. Whereas all of the tests that have been previously reviewed in this section can be hand-scored, standardized scores for the WJ-IV can only be obtained from an online scoring platform that charges a per-use fee.

Given its relation to CHC theory, the WJ-IV features prominently in several interpretive approaches based on that theory that may be used as part of LD diagnosis. For example, McGrew

CHAPTER 14

and Wendling (2010) provide a summary of relations between CHC abilities and specific areas of achievement which, when combined with the WJ-IV CHC subtest classifications reported in the most recent XBA handbook (Flanagan, Ortiz, & Alfonso, 2013), may be used to guide selective assessment. In fact, an emerging PSW method, the Core-Selective Evaluation Process (C-SEP; Schultz & Stephens, 2017) seemingly was designed to align with core sets of tests from the WJ-IV. There are two sources that describe how to utilize the WJ-IV within that model (Schrunk, Stephens-Pisecco, & Schultz, 2017; Schultz & Stephens, 2015), and the C-SEP was outlined in an advertorial by the publisher in the *Communique*, a practitioner newsletter published each month by the National Association of School Psychologists.

Curriculum-Based Measurement

Although curriculum-based measurement (CBM) is frequently used to document treatment response within an RTI intervention framework, it can also be used to conduct a diagnostic assessment of academic skills. CBM represents a series of alternative probes of core academic skills (e.g., Oral Reading Fluency, Phonemic Awareness, Math Computation, Writing Fluency, Spelling, Maze [as a proxy for reading comprehension]). The measures are timed (e.g., 1–3 minutes' duration) and yield measurements of both accuracy and response rate. Probes are designed to be brief, repeatable, and sensitive to small increments in change. As a result of these properties, CBM is a useful technology for screening, as well as progress monitoring, children with academic difficulties (Deno, 2003). CBM probes are considered to be general outcome measures (GOMs); that is, they provide a consistent scale for decision making across subskills that serve as the focal target for intensive interventions. CBM works much like GOMs in other disciplines (e.g., thermometer in medicine, stock market index in economics). For example, even though a reading intervention may target a specific subskill of reading (e.g., sound blending), the effects of that intervention will be captured by corresponding changes in a person's reading

CHAPTER 14

fluency rate over time. In a comparison of progress monitoring data from reading fluency probes and specific subskill mastery measures of reading (SSMM) that more closely matched intervention targets, Van Norman, Maki, Burns, McComas, and Helman (2018) found that while some SSMM's provided an incremental benefit beyond GOM's for early struggling readers, this distinction became less meaningful as readers were exposed to more complex phonetic patterns as they got older.

A number of studies have validated the use of CBM with children suffering from academic dysfunction, and grade norms and expected growth rates have been established for every major type of CBM (e.g., Deno, Fuchs, Marston, & Shin, 2001). Thus, if a child's CBM level (absolute difference) or slope (growth) of learning deviates significantly from expected levels of performance, this may indicate the presence of LD. More importantly, CBM data directly inform academic intervention and instructional planning and have been shown to predict year-end high-stakes test outcomes. Several resources for CBM assessment and interpretation are provided in Appendix 14.1.

Contemporary Classification Approaches

LD assessment methods have been influenced by changes and modifications to federal regulations over the course of the last 40 years. The final regulations adopted by the USOE (1977) focused on a severe discrepancy resulting in a conceptual shift away from a deficit in processing to explain underachievement. According to Fletcher and colleagues (2013), this shift was a response to concerns about the validity of processing-based approaches to identification and the perceived lack of efficacy of related process training (e.g., Mann, 1979). Ironically, the field is presently experiencing a “back to the future”-type moment as enthusiasm for new processing-based approaches (i.e., PSW) has increased. Nevertheless, existing federal regulations and diagnostic criteria permit the use of multiple assessment methods for LD identification, and

CHAPTER 14

these methods have different assumptions regarding LD as a construct. What follows is a review of the three classification methods that have been featured prominently in the professional literature. We should note that even though our discussion focuses on the salient inclusionary criteria in each of the models, all identification models stipulate that LD classification should be based on a comprehensive evaluation of other relevant factors beyond these inclusionary criteria (i.e., exclusionary factors). We point this out because proponents (e.g., Hale et al., 2010) of various methods frequently invoke the term *comprehensive assessment* to suggest that the use of particular types of assessments (i.e., neuropsychological measures) may be required by existing codes and guidelines, even though this may not actually be the case (Zirkel, 2013). When considering whether to add a new test to a preferred battery, it is important for practitioners to consider the incremental validity associated with the use of that instrument in clinical practice. More testing does not always result in improved decision making and may in some circumstances actually create a dilution effect in which aggregate validity is lowered (Kraemer, 1992).

Ability–Achievement Discrepancy

The discrepancy method is based on the logic of cognitive referencing, in which a child's IQ is used as a benchmark for achievement. When measured achievement deviates significantly from measured IQ, this underachievement is thought to be unexpected. As the federal regulations do not specify a preferred discrepancy formula, numerous variations of this model have been employed. Most jurisdictions employ a “simple-difference” standard score comparison, in which a discrepancy is considered severe only if the difference between IQ and achievement meets or exceeds an a priori cutoff point or threshold. Most existing simple-difference thresholds range from one to two standard deviations (15–30 standard score points). For example, in a state that utilizes a 1.5 standard deviation cutoff point (22.5 points), a child with an IQ score of 100 would

CHAPTER 14

have to obtain an achievement score < 77 in order to have a severe discrepancy.

Other methods that have been employed include expectancy formulas (based on chronological age) and regression-based formulas that seek to control for regression to the mean. Using the regression method, a critical value is established between two scores, and the observed discrepancy must be greater than this value in order to be considered severe. According to Reynolds (1984–1985), the latter is the most psychometrically defensible method, but it is not commonly utilized by practitioners due to its computational complexity. Whereas in the majority of states employing the discrepancy model, a severe discrepancy alone represents the de facto inclusionary criteria for LD eligibility, a handful of states also require examiners to furnish evidence of a concomitant deficit in cognitive processing.

Strengths of the Discrepancy Model

The discrepancy model has several strengths. Most notably, it is the easiest of the three methods to implement, and it aligns well with the idea that LD is marked by unexpected underachievement. Among existing models, it most closely approximates an actuarial model in which the effects of clinical judgment are minimized and decision making *should be* relatively consistent across practitioners and jurisdictions in which the same formula is employed (Canivez, 2013).

Weaknesses of the Discrepancy Model

In addition to the issues that were raised previously, additional measurement issues have been raised about the method. Before proceeding, it is important to point out that all of the cognitive and achievement attributes that are commonly assessed as part of a comprehensive evaluation appear normally distributed, representing continua with no natural demarcations. In the discrepancy model, these attributes are reduced to categorical (yes–no) classifications. The problems with artificially dichotomizing continuous data have long been known (MacCallum,

CHAPTER 14

Zang, Preacher, & Rucker, 2002). In particular, decision error (i.e., diagnostic misses) will likely be exacerbated around the cutoff point due to the fact that we are unable to measure these attributes without error. Even though discrepancy models require practitioners to interpret observed differences as “point” estimates, traditional norm-referenced tests may not be capable of measuring attributes with that level of precision (Beaujean et al., 2018).

Whenever scores are positively correlated (which is almost always the case with IQ and achievement), the discrepancy between those scores will always be *less* reliable than its reference scores. If the confidence band associated with a discrepancy score straddles a diagnostic threshold, the likelihood of *actually* meeting or not meeting discrepancy criteria is relatively equivalent. Not surprisingly, in a longitudinal study of the reliability of LD classification using the discrepancy method, Francis and colleagues (2005) found that 30% of the children in the LD and non-LD groups changed group membership between grades 3 and 5 as a function of being reassessed. Put simply, many examinees who present with a severe discrepancy at Time 1 may not maintain the discrepancy when assessed again at Time 2 due to factors (e.g., measurement error, test selection³) not having to do with actual individual differences.

Furthermore, studies indicate that a severe discrepancy only accounts for approximately 1–2% of the variance in treatment response (e.g., Stuebing, Barth, Molfese, Weiss, & Fletcher, 2009) and that it disproportionally favors individuals with higher IQs as an artifact of regression to the mean. There has also been significant concern about the lack of decision-making consistency within and between jurisdictions (Peterson & Shinn, 2002). As a consequence,

³In our experience, it does not take long for skilled practitioners to become adept at selecting instruments and/or scores as the focal point of discrepancy analyses that are more likely or not to yield a severe discrepancy depending on which outcome is preferred. For example, practitioners are frequently taught to “invalidate” the FSIQ score when an examinee presents with significant variability in his or her cognitive profile. Although this practice is popular, its validity has been questioned (see McGill, 2016).

CHAPTER 14

Wilson (1987) concludes that the discrepancy model represents an “atheoretical, psychologically uninformed solution to the problem of LD classification” (p. 28). Even so, the method continues to be widely used in clinical and educational settings, and epidemiological studies indicate that the vast majority of individuals diagnosed with LD are likely identified using some variation of this method (McDermott, Goldberg, Watkins, Stanley, & Glutting, 2006).

Processing Strengths and Weaknesses

In contrast to the IQ–achievement discrepancy model, which places greater importance on the interpretation of an individual’s FSIQ score, an emerging class of LD identification models focuses more on inspecting the variability and scatter among cognitive scores in order to make inferences about the PSWs that are thought to underlie LD, and it is thought that LD subtypes can be distinguished by these unique score patterns. This approach to LD identification is broadly referred to as patterns of PSWs.

To date, several models have been proposed that attempt to operationalize PSW, including (1) the concordance/discordance model (C/DM; Hale & Fiorello, 2004), (2) the dual-discrepancy/consistency model (DD/C; Flanagan et al., 2018), and (3) the discrepancy/consistency model (D/CM; Naglieri, 2011). Although they have different theoretical orientations and employ different criteria to identify PSWs, all three models apply the same fundamental logic for identifying a confirmatory PSW pattern in assessment data; that is, in order to confirm LD, there must be evidence of cognitive and academic weaknesses in the presence of otherwise spared abilities, and the cognitive weakness should be linked theoretically⁴

⁴The decision about whether a particular cognitive weakness meets this standard is based largely on the clinical judgment of the assessor. Recent attempts to operationalize these procedures are based largely on qualitative interpretations of the professional literature.

CHAPTER 14

to the area of academic concern.

Concordance/Discordance Model

The C/DM approach suggests that LD is demonstrated by an exclusive pattern of concordant (consistent performance with a reference variable) and discordant (inconsistent performance with a reference variable) in an ability profile. Potential concordances and discordances are determined to be statistically significant if they exceed critical values obtained using the standard error of the difference (SE_d) formula (see Hale, Wycoff, & Fiorello, 2011, for a demonstration). The C/DM is one of the few PSW models that explicitly endorses use of a statistical approach for evaluating pairwise discrepancies. To facilitate these comparisons, users are encouraged to select subtests and/or composite scores from different batteries that are most likely to yield beneficial information based on the referral concern. When a confirmatory PSW pattern is observed in these data, examiners must cross-validate that finding with additional sources of information (i.e., records reviews, observations) to ensure that the assessment has ecological validity. This notion is stressed in every approach to PSW assessment. Although these types of default statements imply that a confirmatory PSW pattern, in and of itself, is not diagnostic, it is important to note that the presence of a unique PSW represents the de facto inclusionary criteria for the method.

Discrepancy/Consistency Model

It can be argued that the D/CM model, first described by Naglieri (1999), was the first systematic PSW model to be proposed in the professional literature. The D/CM utilizes the same conceptual approach as C/DM to examine profile variability for the presence of LD; however, it employs different criteria to identify a cognitive weakness. According to Naglieri (2011), dual criteria are applied to determine whether the score reflects a legitimate cognitive weakness: The score must be both a relative weakness (via ipsative analysis) and a normative weakness (e.g.,

CHAPTER 14

standard score < 90). If a child presents with a cognitive weakness that is related to an achievement weakness in the presence of otherwise spared abilities, this may be regarded as a confirmatory PSW pattern. Although most descriptions of the D/CM have illustrated its application using various iterations of the cognitive assessment system, the method can be applied to any cognitive tests. In contrast to other methods, the D/CM is designed primarily to be applied to scores obtained from the same test battery.

Dual-Discrepancy Consistency Model

The DD/C (Flanagan et al., 2018) was originally proposed in the XBA literature but is conceptually different from that approach. The DD/C uses the CHC model as a theoretical foundation for creating an operational definition for LD. Clinicians are encouraged to administer cognitive measures that correspond with the broad abilities posited in the CHC model, then compare those scores to scores obtained from a norm-referenced test of achievement. SLD identification using the DD/C requires users to (1) identify an academic weakness, (2) determine that the academic weakness is not primarily due to exclusionary factors, (3) identify a cognitive weakness, and (4) determine whether a student displays a confirmatory PSW consistent with several criteria specific to the DD/C model (see Flanagan et al., 2018, for an in-depth description of these criteria). Although normative cutoffs (e.g., standard score < 90) have been provided in the DD/C literature, clinicians may use professional judgment in determining whether a child manifests a relevant weakness in cognition or achievement.

To aid decision making, Flanagan, Ortiz, and Alfonso (2017) have developed a cross-battery assessment software system (X-BASS) that contains a PSW score analyzer inspired by XBA/CHC theory. It should be noted that the software requires that users input scores for all of the consensus broad ability factors in CHC theory. As a result, clinicians favoring a test other than the WJ will likely have to administer measures from other tests in order to utilize the

CHAPTER 14

program. Whereas the DD/C approach is presently the most commonly referenced PSW method in the professional literature, it remains unclear to what degree users *actually* adhere to the operational procedures described in DD/C materials (Beaujean et al., 2018).

Strengths of the PSW Model

Before proceeding, we must acknowledge that any discussion regarding PSWs associated with the PSW model should be interpreted as speculative, as this is a relatively new method of LD identification, and empirical data for its potential efficacy have only recently begun to accumulate. Proponents of these methods point to studies showing the differential predictive effects of cognitive abilities for achievement across the lifespan (e.g., Cormier, Bulut, McGrew, & Singh, 2017; McGrew & Wendling, 2010) and that groups formed using PSW approaches have statistically significant means differences on scores (e.g., Feifer, Nader, Flanagan, Fitzer, & Hicks, 2014). The latter form of evidence has been interpreted as indicating that PSW may be useful in isolating and amplifying various LD subtypes. However, Miciak and colleagues (2014) note that “evidence for the existence of distinct disability subtypes is not *ipso facto* evidence for the reliability, validity, or utility of PSW methods for LD identification” (p. 23).

Additionally, as compared to the traditional discrepancy model, there is a greater focus on prevention and using assessment data for intervention. As an example, all of the PSW models that we have described assume that each child has been the recipient of preventive interventions (i.e., RTI) prior to being referred for a comprehensive evaluation. In this way, these models may be regarded as *hybrid* or “third-method” approaches, as they seek to marry elements of RTI and PSW, and it is suggested that assessment data generated from both approaches can be used as part of a comprehensive assessment to determine whether a child has LD (Flanagan, Fiorello, & Ortiz, 2010).

Potential Weaknesses of the PSW Model

CHAPTER 14

At face value, it may appear that these models provide users with the best of both worlds; however, they have been the subject of considerable research criticism. In a review, McGill and colleagues (2016) identified several concerns with the PSW method. Specifically, (1) cognitive weaknesses are ill-defined, and their identification is not consistent across the models; (2) there is presently limited evidence for PSW diagnostic utility; (3) scores that are the focal point of PSW analyses may not be suitable for individual decision making; and (4) it remains unclear whether psychologists have adequate training to implement these assessment procedures with integrity given prior integrity concerns with less complex approaches such as the discrepancy model. Available research has largely supported these concerns.

In a simulation study, Stuebing, Fletcher, Branum-Martin, and Francis (2012) evaluated the technical adequacy of three PSW methods and found that all three methods were very good at identifying “not-LD” but had low to moderate sensitivity and very low positive predictive values indicating that the methods were not very good at identifying LD as defined in the simulations. These findings were replicated in a more recent investigation by Kranzler and colleagues (2016a), who reported diagnostic efficiency statistics associated with the application of DD/C procedures to assessment data for 900 participants in the WJ-III normative sample. Consistent with DD/C parameters, a true positive was indicated when participants had an academic weakness, and a predicted cognitive weakness also occurred in the presence of average or better general ability. Prevalence rates for SLD, as defined in the study, ranged from 0 to 7% depending on the target area of cognitive weakness. Mean specificity and negative predictive values were 92 and 89% across CHC cognitive abilities and achievement domains, indicating that the absence of a cognitive weakness was very accurate in detecting what they deemed to be “true negatives” (individuals without an achievement weakness). On the other hand, sensitivity and positive predictive value (PPV) estimates (21 and 34%, respectively) were quite low,

CHAPTER 14

indicating that the presence of a cognitive weakness may not be very useful at accurately identifying what they classified as “true positives.”

Of additional concern, diagnostic decisions based on PSW assessment appear to be highly unstable, and the reliability of decisions appears to worsen as more assessment data are gathered (Taylor, Miciak, Fletcher, & Francis, 2017). These findings seem to counter one of the core axioms of psychological assessment that additional information about a client helps to reduce uncertainty. As we mentioned previously, additional information is only useful when it is not redundant. Furthermore, overlap between the methods has been found to be low, and classification agreement may be impacted by nontrivial factors such as test selection (Miciak, Taylor, Denton, & Fletcher, 2015), model choice (Miciak et al., 2014), and the arbitrary use of cutoff points (McGill et al., 2016). Even so, proponents of these methods suggest that these studies merely illustrate a mechanistic approach to decision making and fail to take into account that “other data gathered through multiple methods and multiple sources need to be considered and must corroborate any conclusions that are drawn from the PSW analyses” (p. 488). Despite the intuitive appeal of such declarative statements, it remains unclear what these *other* data sources are or how practitioners should weight various pieces of information they encounter in the data gathering process. Although clinicians often claim to base conclusions on the complex integration of most or all of available data, studies emanating out of the decision science literature have consistently indicated that individuals often overweigh nondiagnostic information and have difficulty accounting for interactions in as little as two or three pieces of information during the decision-making process (e.g., Faust, 1989; Nisbett, Zukier, & Lemley, 1981). As a consequence, it is difficult to envision the psychometric shortcomings of PSW analyses being overcome by additional data collection alone.

As we mentioned previously, it is important to realize that, at its heart, the PSW method

CHAPTER 14

may very well be a reparameterization of the discrepancy model; that is, a confirmatory pattern in many of the models requires users to identify whether there are significant differences (i.e., discrepancies) between cognitive–achievement scores. Except, instead of evaluating one pairwise comparison (IQ–achievement), users are required to appraise multiple pairwise comparison simultaneously, which may result in inflated Type I decision error. Thus, the measurement issues associated with the discrepancy model are likely to be exacerbated. Whereas PSW proponents (e.g., Hale et al., 2011) raise these concerns when discussing rival models, their potential implications for PSW decision making are rarely discussed.

To be fair, some PSW models do not employ traditional discrepancy procedures and instead adopt ad hoc cutoff points for determining what constitutes a cognitive–achievement weakness. Typically, suggested thresholds are derived on a normative basis wherein a standard score ≤ 85 or 90 indicates the presence of a cognitive or achievement weakness. According to Streiner (2002), dichotomizing a continuous variable results in lost information and increased probability of decision error regardless of where the cutoff point is imposed. Figure 14.1 illustrates the logic of cutoff point analyses in the PSW model using simulated test scores for a hypothetical cognitive and achievement score. As can be seen in Figure 14.1, many hundreds of cases cluster together right at the intersection of the cutoff point threshold (< 85) for both tests. Similar to the issues we discussed when describing the issues with cutoff points associated with the discrepancy model, the corresponding confidence intervals associated with scores that reside right around a cutoff point are likely to contain values that fall above and below the threshold. In this scenario, it is incredibly difficult for a practitioner to validly determine whether a patient is a true positive (lower left-hand quadrant) *or* a true negative case (all other quadrants) as posited by the PSW model.

Although it is often suggested that PSW assessment may be useful for treatment

CHAPTER 14

planning, these claims run counter to a long-standing body of empirical evidence that calls into question the utility of ATI prescriptions (Fletcher & Miciak, 2017). For example, in a meta-analysis, Burns and colleagues (2016) found that the effect sizes associated with academic interventions guided by cognitive data were mostly small, with only the effects associated with interventions informed by phonological awareness providing moderate treatment effects. As a result, they concluded that “the current and previous data indicate that measures of cognitive abilities have little to no utility in screening or planning interventions for reading and mathematics” (p. 37). It should be noted that even those who support the use of these methods have begun to acknowledge discrepancies between the laudatory ATI rhetoric and available research evidence. To wit, Schneider and Kaufman (2017, p. 8) asserted, “After rereading dozens of papers defending such assertions, including our own, we can say that this position is mostly backed by rhetoric in which assertions are backed by citations of other scholars making assertions backed by citations of still other scholars making assertions.”

In summation, it remains to be seen whether PSW implementation represents a legitimate step forward for identification and treatment. Although billed as new and revolutionary, these models rely on the use of profile analytic logic that has considerable psychometric limitations. According to Fletcher and colleagues (2013) “It is ironic that methods of this sort continue to be proposed when the basic psychometric issues are well understood and have been documented for many years” (p. 40). Unfortunately, this body of literature is *rarely* cited by proponents of these methods, a practice that should cause concern in an era of psychological science that stresses evidence-based practice. As a result, practitioners are encouraged to carefully consider the value of the information yielded by these assessment practices relative to the costs in time and resources associated with their adoption costs (Williams & Miciak, 2018). Although PSW is described as a “research-based” method, compelling evidence that a confirmatory PSW pattern is

CHAPTER 14

necessary for the diagnosis or treatment is presently lacking (Kranzler et al., 2016b; Miciak, Taylor, Stuebing, & Fletcher, 2018).

Response to Intervention

It is somewhat difficult to define RTI as a method for LD identification, as RTI represents a prevention and intervention framework and not an assessment model per se. However, due to revisions in the federal regulations in 2006, RTI has become synonymous with LD identification in the professional literature (Burns, Jacob, & Wagner, 2008). RTI is an umbrella term that generally refers to three tiers of increasingly intense instruction, using ongoing progress monitoring to evaluate treatment response and movement between the tiers. Tier 1 represents instruction in the general education classroom. It is assumed that if instruction is effective, then this will benefit the overwhelming majority of children. Tier 2 involves providing supplemental instruction to children who do not benefit from the general education curriculum at Tier 1. This represents the first attempt to remediate academic dysfunction and usually involves some type of evidence-based group or individualized treatment. Tier 3 can differ significantly across models and may represent providing a child with a more intensive intervention in comparison to that provided at Tier 2 (i.e., higher dose, targeted individualized treatment based on further assessment) or a referral for a comprehensive evaluation. The core tenets of RTI are providing evidence-based instruction with fidelity and monitoring treatment progress frequently using low-inference assessment tools such as CBM. As such, even though classification may be a terminal outcome of an RTI model, it is not an explicit goal of the framework.

Whether implicated once a child reaches Tier 3, or after he or she has not benefited from the intervention(s) employed at that tier, a comprehensive evaluation typically includes multiple measures to supplement available RTI intervention data (i.e., CBM). These assessments may

CHAPTER 14

include a norm-referenced test of achievement to further validate the presence of low achievement, behavioral rating scales, direct behavioral observations, and/or a diagnostic interview to rule out exclusionary factors or provide information that is relevant for differential diagnosis. Thus, classification in the RTI model is not based solely on treatment outcome data. However, this information can be important for establishing that a child's rate and level of learning falls below expected level of performance.

Fuchs and Fuchs (1998) describe what is referred to as a dual-discrepancy identification model in which LD is characterized by concomitant deficits in learning rate (i.e., slope) and level (i.e., below average performance on a norm-referenced test of achievement). "Low achievement" models (e.g., Dombrowski, Kamphaus, & Reynolds, 2004), in which LD is marked by deficient achievement across multiple measures, may also be applied within the context of an RTI classification model. In contrast to cognitive discrepancy models, unexpected underachievement is documented by the presence of spared achievement indicating that the area(s) of deficit are focal in nature. Cognitive testing is typically deemphasized in RTI, although it may be used to rule out intellectual disability or to assess areas of cognitive processing that have been linked to beneficial treatment outcomes (i.e., phonological awareness).

Strengths of the RTI Model

As we discussed previously, the RTI model is a direct response to the perceived shortcomings of the traditional discrepancy method. In particular, the explicit focus on prevention illustrates well that a child no longer has to "wait to fail" in order to receive intervention. Although RTI methods may require a profound shift in how psychologists conceptualize and assess LD, surveys indicate that clinicians prefer it to the discrepancy method (O'Donnell & Miller, 2011). Thus, there is evidence of social validity among practitioners in areas in which it has been implemented.

CHAPTER 14

Although “true LD” is a somewhat amorphous construct in RTI because eligibility is not based on previously accepted inclusionary criteria (i.e., discrepancy, processing deficit), prevalence studies suggest that the model may be useful in identifying a subset of children that are likely to be LD. Speece and Case (2001) evaluated the technical adequacy of several LD identification methods and found that use of a dual-discrepancy CBM approach identified 8% of the population before Tier 2 intervention, which is close to the ~5–10% estimated rate of LD. The results have been replicated by Fuchs and colleagues (2005) who found similar estimates in a rare RTI model for math intervention. Relatedly, Fuchs, Compton, Fuchs, and Bryant (2008) found that the CBM assessments commonly used in RTI had adequate sensitivity and specificity for determining future disability status in 252 children evaluated in first grade. Additionally, there is also some evidence to indicate that RTI may be useful at reducing disproportionality in LD classification (VanDerHeyden, Witt, & Gilbertson, 2007), a major concern that was expressed in the 2001 LD Summit.

Weaknesses of the RTI Model

In spite of the many positives associated with the model, Kavale and colleagues (2008) warn that there may be a rhetoric of self-congratulation in the RTI literature regarding its use as a classification method. One of the major factors complicating its use in this context is the fact that classification is essentially conferred by default as a result of poor treatment outcomes, and little consideration is given to the underlying etiology of the disorder. Group studies indicate that individuals with LDs have large cognitive processing deficits compared to normal controls (Johnson, Humphrey, Mellard, Woods, & Swanson, 2010). Although, the implications of these findings have been difficult to translate into practice, Fuchs, Hale, and Kearns (2011) suggest that even if one supports the preventive mission of RTI, it is difficult to argue against the notion that it may be necessary for practitioners to go beyond RTI data, including the careful use of

CHAPTER 14

cognitive measures.

Additionally, Reynolds and Shaywitz (2009) suggest that a rigid application of RTI, in which children at-risk for LD are required to progress linearly through various tiers of instruction that may be of dubious quality, seemingly shifts the process of LD identification from a “wait to fail” to a “watch them fail” approach. Emerging research on the efficacy of RTI assessment and intervention methods seems to support this notion. In an RTI efficacy study, Compton and colleagues (2012) found that, among 129 first grade children who were unresponsive to Tier 1 instruction, chronic nonresponders (i.e., children who failed to respond to instruction at Tiers 1 and 2) could be identified from Tier 1 data alone. These results suggest that a different approach to RTI assessment may be necessary to avoid prolonged periods of failure.

Burns and colleagues (2008) have identified a series of threats to valid RTI practice, including but not limited to (1) the lack of established research-based interventions for diverse groups and academic domains; (2) uncertainty about when a comprehensive assessment is warranted; (3) difficulty translating implementation science to conventional school settings; and (4) inadequate training and fidelity. As a result of these threats, it is not surprising that a recent large-scale Institute of Educational Sciences (IES) outcome study (Balu et al., 2015), purporting to evaluate the effectiveness of RTI models, found that assignment to a targeted intervention in Tiers 2 or 3 had little effect on reading performance in elementary schools nationwide and, in some cases, the intervention outcomes were contraindicated (i.e., students’ performance worsened in response to treatment).

Summary

Prominent LD identification methods (discrepancy, PSW, RTI) provide different conceptualizations of LD, employ different assessment procedures, and possess different strengths as well as limitations relative to rival methods. As should be evident, model selection is

CHAPTER 14

not arbitrary, as different models will likely identify different subsets of children who present with academic difficulties. Similar to the previous edition of this chapter, LD assessment remains a subjective exercise in clinical judgment; regardless of the method used, clinicians do not appear to consistently employ or adhere to extant identification criteria.

In a relatively recent study, Maki, Burns, and Sullivan (2017) evaluated the consistency of practitioners in rendering LD classifications using a series of vignettes representing the three major classification approaches. No differences between the methods were found, but overall consistency was low. In an interesting follow-up to that study, Maki and colleagues (2018) found that diagnostic (over)confidence may be an important moderator variable within these results. In a survey of 376 school psychologists, practitioners who reported moderate levels of confidence produced less consistent decisions with ambiguous assessment data than clinicians who reported being “not very confident” in their appraisals. We would argue that part of the problem is that the data (i.e., cognitive test scores) used to render these types of judgments contain measurement error that will likely degrade any decisions-making model in which these data are the primary focus of clinical interpretation. In confronting this reality, it is important for practitioners not to fall prey to what Lilienfeld, Wood, and Garb (2006) term the *alchemist’s fantasy*—the belief that powers of intuition enable them to transform questionable test scores into clinical gold.

As a result of these limitations, the search for a “gold standard” method for LD identification remains elusive. Even so, we are mindful that many clinicians practice in jurisdictions where they may be compelled to engage in one particular method or another. Thus, “saying no” to methods that have been shown to be problematic (e.g., cognitive discrepancy methods) simply is not an option. In these circumstances, we gently remind practitioners of their responsibility to fully disclose the potential limitations of various assessment methods as part of their informed consent mandate.

Current Issues with LD Assessment

Other difficulties for assessment are the lack of a diagnostic “gold standard” for LD, which complicates any attempt to validate proposed identification frameworks. Additional psychometric and conceptual concerns include issues that have been raised about some of the scores that are the focal point of clinical interpretation in many diagnostic schemes, as well as potential conflicts of interest that pervade the field. We next provide a brief overview of these factors.

Lack of a Diagnostic “Gold Standard”

Presently, there is no diagnostic “gold standard” for LD, which makes translating LD assessment research to clinical settings difficult. As an example, although diagnostic validity studies of various PSW permutations have consistently furnished negative results, the lack of a “gold standard” means that there is no way to truly know the rate at which the method may or may not accurately identify true LD. Thus, at best, these studies should only be regarded as a sort of “best guess” for the potential utility of the method.

In a review of LD research published from 2001 to 2013, Williams, Miciak, McFarland, and Wexler (2016) found that identification varied widely and nearly one-third of all studies investigating LD failed to describe how the participants were identified, illustrating well that LD identification remains ill-defined. According to Meehl (1978), substantive theory building from an open concept is conceptually difficult because, by definition, the boundary conditions for the construct have not been established; thus, it is difficult to articulate the conditions necessary for knowledge claims about it to be falsified. Apropos, Kranzler and colleagues (2016b) raise this very concern with respect to the DD/C model and encourage practitioners to view these methods with skepticism until additional empirical evidence is furnished to support its use.

Validity of Scores from Commercial Ability Tests

CHAPTER 14

Validation of commercial ability measures involves consideration of evidence on test content, internal structure (i.e., structural validity), and relations with other measures. Whereas each of these elements is important in its own right, structural validity is especially important because it provides the theoretical and statistical rationale for the scores that are provided to users. For example, if a factor that is thought to measure a particular cognitive ability (e.g., long-term memory) is not located by a factor analysis, the score representing that construct may be illusory. Factor validity is a foundational validity measure and supports all other attempts at establishing construct validity.

Since 2000, a body of independent factor analytic research has emerged raising concerns about the integrity of many of the structures and interpretive models promoted by the publishers of commonly used intelligence tests (Canivez, 2013). In some cases, these discrepancies are small and involve only one or two factors (e.g., Canivez, Watkins, & Dombrowski, 2017; McGill & Dombrowski, 2018); in other cases, whole aspects of the posited structure cannot be replicated (e.g., Canivez, 2008; Dombrowski, McGill, & Canivez, 2018).

Even when posited dimensions can be located, these indices often contain insufficient unique variance for confident clinical interpretation and mostly reflect systematic variance that is attributable to *g*. As a result, many broad ability scores and composites may lack the requisite incremental validity to predict meaningful achievement outcomes beyond *g* (Canivez, 2013). These results are not unique to any particular instrument, and virtually every current ability measure has been implicated by these analyses. As structural and incremental validity have important implications for the potential clinical utility of scores, these results challenge the interpretive foundations of identification models that rely predominantly on the clinical interpretation of subscale scores (i.e., PSW). To date, proponents of these methods (e.g., Decker, Hale, & Flanagan, 2013) have yet to demonstrate a method by which these potential

CHAPTER 14

psychometric limitations may be overcome when using these scores to make decisions about individuals.

Processing Assessment: Just Say “Maybe”?

The assumption that a deficit in cognitive processing underlies an LD is an idea that predates the advent of PSW and can be traced to the initial conceptualization of the current federal definition of SLD. Nevertheless, many PSW guidebooks and interpretive manuals encourage practitioners to engage in elaborate processing assessment procedures and speculate about the potential linkages between cognitive processing and achievement. For example, in the Ventura County PSW manual (Ventura County SELPA, 2017), users are provided a matrix of cognitive–achievement relationships to help facilitate the identification of confirmatory PSW patterns in assessment data. The strength of these relationships are evaluated on a 5-point scale after appraising the “research base of processing-achievement relations” (p. B8). In that document, it is suggested that long-term storage and retrieval (Glr) is one of the strongest predictors (4; highest rating) for mathematics achievement. However, the results furnished in a recent study did not support this claim; McGill, Conoyer, and Fefer (2018) employed elements of the evidence-based assessment framework to shed insight on how well a processing weakness (< 85) in Glr discriminated between individuals with and without achievement weaknesses in mathematics using data from the KABC-II normative sample.

Results indicate that individual decisions based on the Glr and mathematics linkage posited in the Ventura County PSW manual are likely specious. The true positive rate (sensitivity) was .381 and the true negative rate (specificity) was .888. Positive predictive power was .394 and negative predictive power was .882, indicating that the presence of a processing weakness in Glr functions below chance as a potential rule-in test of an achievement weakness in math. Accordingly, the diagnostic odds ratio for a positive test (+3.40) falls well below

CHAPTER 14

recommended guidelines for a quality diagnostic indicator (Streiner, 2003). An AUC value of .634 was found, indicating that this particular PSW pattern has relatively low classification accuracy overall (Youngstrom, 2014). To better understand the utility afforded by assessment procedures, Meehl (1954) encouraged clinicians to “bet the base rate” to determine whether assessment information meaningfully improves our understanding of a clinical phenomenon in comparison to conducting no assessment at all. Probability theory dictates that ~15% of the population has an academic weakness in any given area (i.e., < 85). So how much does the presence of a cognitive weakness in Glr increase the posterior probability of having an academic weakness in mathematics? The probability nomogram in Figure 14.2, produced from these data, results in a posterior probability of only 37%. This means that out of 100 cases presenting in a clinical setting with similar assessment results (evidence of processing weakness), only 37 would be expected to have an academic disorder in mathematics. Whereas this is an improvement from the base rate, clearly, additional assessment data are needed in order for us to feel confident in treating this individual as having LD.

However, establishing linkages between cognitive–achievement weaknesses represents only one step in the broader PSW process. Flanagan and Schneider (2016) note that there are many *potential* reasons why cognitive deficits may not lead to academic deficits for some individuals, as “cognitive abilities are causally related to academic abilities, but the causal relationship is of moderate size, and only probabilistic, not deterministic” (p. 141). Even so, it is difficult to regard an assessment variable as important when it predicts a relevant diagnostic outcome at lower than chance levels. Proponents of these methods must also come to terms with the fact that a specific pattern, ruling in or ruling out LD has yet to be established (Mather & Schneider, 2015). To be clear, we are not suggesting that practitioners should “just say no” to cognitive testing as a matter of course. Although we are sympathetic to position expressed by

CHAPTER 14

Fuchs and colleagues (2011), we believe that available empirical evidence indicates that the use of IQ tests should be limited to ruling out exclusionary factors (i.e., intellectual disability), with additional applications related to the assessment of cognitive processing employed cautiously, if at all.

Potential Conflicts of Interest in the Assessment Literature

Over the course of the last 20 years, proponents of various methods have tended to marginalize competing frameworks, while magnifying the strengths and minimizing the weaknesses of a favored assessment approach (Dombrowski, Ambrose, & Clinton, 2007). The potential contraindicated effects of this insularity should be considered given the potential conflicts of interest that pervade assessment training and practice. For example, a nontrivial proportion of workshops where LD assessment methods are disseminated to practitioners feature authors who receive royalties from books, chapters, and commercial tests featuring the use of these methods, and the use of disclosures of conflict of interest has historically been inconsistent in clinical science (Truscott, Baumgart, & Rogers, 2004). There have been marked changes in the standards of practice in terms of disclosing conflicts of interest in medical research, publishing, and continuing medical education, and a similar move toward increased reporting and transparency is happening in other areas of psychology. Still, given these complexities, it is imperative that practitioners develop a skills set that helps them to discern between evidence-based and non-evidence-based practices.

Conclusion

As a result of the numerous psychometric and conceptual limitations that continue to plague LD identification and assessment, it is relevant to ask, “What should practitioners do?” Admittedly, this is a difficult question to answer, as there is limited empirical evidence to indicate that any of the prevailing approaches can be used to reliably and validly identify LD,

CHAPTER 14

and endorsing any particular approach at the expense of the another will no doubt engender significant controversy. Accordingly, we believe that the most defensible position at the present time is to encourage practitioners and scholars to consider adopting more parsimonious assessment methods that align with DSM-5. The dual-deficit functional academic impairment (DDFAI) model described by Dombrowski and colleagues (2004) and the “hybrid” approach articulated by Fletcher and colleagues (2013) are model frameworks that bear further consideration. As can be seen in Table 14.2, both approaches emphasize that an academic deficit must be established using dual-deficit criteria. Whereas both approaches require that a deficit be established using conventional norm-referenced achievement tests as part of the first criterion, they differ on the process used to indicate the secondary criterion. Dombrowski and colleagues suggest that, beyond low achievement, a child must also have a *functional* impairment in academics, which can be established through measures of performance in the child’s academic setting (i.e., grades, test scores). Alternatively, the hybrid approach suggested by Fletcher and colleagues requires evidence specifically of inadequate response to instruction. Whereas RTI could be used to establish a functional impairment in the DDFAI approach, it is not required. Aside from this difference, both approaches are similar in that once the dual criteria for academic impairment have been met, the goal of the comprehensive assessment is to rule out rival hypotheses that might better explain why a child is not performing well academically. In this way, assessments are administered to answer specific clinical questions. For example, although cognitive testing is not required in both of these approaches, a clinician may administer an intelligence test to rule out the presence of intellectual disability when that condition is suspected. Appendix 14.2 contains a case study and sample assessment illustrating the application of the DDFAI model.

Whereas both hybrid models meet the IDEA statutory requirements for a comprehensive

CHAPTER 14

assessment and are consistent with the diagnostic criteria for LD in DSM-V, we realize that some clinicians may be wary of adopting these models in clinical practice because of the deemphasis on cognitive testing and the perceived value afforded by these practices. However, if the implicit goal of clinical assessment is *actually* to remedy the presenting problem, clinicians continuing to endorse classification approaches centered around cognitive testing will have to consider the present evidentiary status for these assessment practices (McGill, Dombrowski, & Canivez, 2018). As noted by Schneider and Kaufman (2017, p. 18), “It is not irrational to believe that comprehensive cognitive assessment is more beneficial than can be supported by current evidence. It is irrational to pretend that the evidence is not needed and that all is well in our field.”

References

- Aaron, P. G. (1997). The impending demise of the discrepancy model. *Review of Educational Research, 67*, 461–502.
- Alfonso, V. C., & Flanagan, D. P. (Eds.). (2018). *Essentials of specific learning disability identification* (2nd ed.). Wiley.
- Al Otaiba, S., Wagner, R. K., & Miller, B. (2014). “Waiting to fail” redux: Understanding inadequate response to intervention. *Learning Disability Quarterly, 37*, 129–133.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). *Evaluation of response to intervention practices for elementary school reading*. Institute of Education Sciences.
- Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *Journal of Learning Disabilities, 1*, 242–249.
- Bateman, B. B. (1965). An educational view of a diagnostic approach to learning disorders. In J. Hellmuth (Ed.), *Learning disorders* (Vol. 1., pp. 219–239). Special Child Publications.
- Beaujean, A. A., Benson, N. F., McGill, R. J., & Dombrowski, S. C. (2018). A misuse of IQ scores: Using the dual discrepancy/consistency model for identifying specific learning disabilities. *Journal of Intelligence, 6*(3). [Epub ahead of print]
- Benson, N., Floyd, R. G., Kranzler, J. H., Eckert, T. L., & Fefer, S. (2018, February). *Contemporary assessment practices in school psychology: National survey results* [Paper presentation]. Meeting of the National Association of School Psychologists, Chicago, IL.
- Braden, J. P., & Ouzts, S. M. (2005). Review of Kaufman assessment battery for children. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (2nd ed., pp.

CHAPTER 14

- 517–520). Buros Center for Testing.
- Bray, M. A., Kehle, T. J., & Hintze, J. M. (1998). Profile analysis with the Wechsler Scales: Why does it persist? *School Psychology International*, 19, 209–220.
- Breaux, K. C., Bray, M. A., Root, M. M., Kaufman, A. S. (2017). Introduction to the special issue and to KTEA-3 error analysis. *Journal of Psychoeducational Assessment*, 35, 4–6.
- Burns, M. K., Jacob, S., & Wagner, A. R. (2008). Ethical and legal issues associated with using response-to-intervention to assess learning disabilities. *Journal of School Psychology*, 46, 263–279.
- Burns, M. K., Peterson-Brown, S., Haegele, K., Rodriguez, M., Schmitt, B., . . . VanDerHeyden, A. M. (2016). Meta-analysis of academic interventions derived from neuropsychological data. *School Psychology Quarterly*, 31, 28–42.
- Canivez, G. L. (2008). Orthogonal higher order factor structure of the Stanford–Binet Intelligence Scales—Fifth Edition for children and adolescents. *School Psychology Quarterly*, 23, 533–541.
- Canivez, G. L. (2013). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwenn (Eds.). *The Oxford handbook of child psychological assessment* (pp. 84–112). Oxford University Press.
- Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2017). Structural validity of the Wechsler Intelligence Scale for Children—Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*, 29, 458–472.
- Cavendish, W. (2013). Identification of learning disabilities: Implications of proposed DSM-5 criteria for school-based assessment. *Journal of Learning Disabilities*, 46, 52–57.
- Compton, D. L., Gilbert, J. K., Jenkins, J. R., Fuchs, D., Fuchs, L. S., Cho, E., . . . Bouton, B.

CHAPTER 14

- (2012). Accelerating chronically unresponsive children to tier 3 instruction: What level of data is necessary to ensure selection accuracy? *Journal of Learning Disabilities*, 45, 204–216.
- Cormier, D. C., Bulut, O., McGrew, K. S., & Sing, D. (2017). Exploring the relations between Cattell–Horn–Carroll (CHC) cognitive abilities and mathematics achievement. *Applied Cognitive Psychology*, 31, 530–538.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116–127.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Irvington.
- Decker, S. L., Hale, J. B., & Flanagan, D. P. (2013). Professional practice issues in the assessment of cognitive functioning for educational applications. *Psychology in the Schools*, 50, 300–313.
- Deno, S. L. (2003). Developments on curriculum-based measurement. *Journal of Special Education*, 37, 184–192.
- Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30, 507–524.
- Dombrowski, S. C., Ambrose, D., & Clinton, A. (2007). Dogmatic insularity in learning disabilities identification classification and the critical need for a philosophical analysis. *International Journal of Special Education*, 22, 2–10.
- Dombrowski, S. C., & Gischlar, K. L. (2014). Ethical and empirical considerations in the identification of learning disabilities. *Journal of Applied School Psychology*, 30, 68–82.
- Dombrowski, S. C., Kamphaus, R. W., & Reynolds, C. R. (2004). After the demise of the

CHAPTER 14

- discrepancy: Proposed learning disabilities diagnostic criteria. *Professional Psychology: Research and Practice*, 35, 364–372.
- Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2018). An alternative conceptualization of the theoretical structure of the WJ IV Cognitive at school age: A confirmatory factor analytic investigation. *Archives of Scientific Psychology*, 6, 1–13.
- Elliott, J. G., & Resing, W. C. M. (2015). Can intelligence testing inform educational intervention for children with reading disability? *Journal of Intelligence*, 3, 137–157.
- Faust, D. (1989). Data integration in legal evaluations: Can clinicians deliver on their premises? *Behavioral Sciences and the Law*, 7, 469–483.
- Feifer, S. G., Nader, R. G., Flanagan, D. P., Fitzer, K. R., & Hicks, K. (2014). Identifying specific reading subtypes for effective educational remediation. *Learning Disabilities: A Multidisciplinary Journal*, 20, 18–30.
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, 21, 93–101.
- Flanagan, D. P., & Alfonso, V. C. (2017). *Essentials of WISC-V assessment*. Wiley.
- Flanagan, D. P., Alfonso, V. C., Sy, M. C., Mascolo, J. T., McDonough, E. M., & Ortiz, S. O. (2018). Dual discrepancy/consistency operational definition of SLD: Integrating multiple data sources and multiple data gathering methods. In V. C. Alfonso & D. P. Flanagan (Eds.), *Essentials of specific learning disability identification* (2nd ed., pp. 431–476). Wiley.
- Flanagan, D. P., Fiorello, C. A., & Ortiz, S. O. (2010). Enhancing practice through application of Cattell–Horn–Carroll theory and research: A “third-method” approach to specific learning disability identification. *Psychology in the Schools*, 47, 739–760.

CHAPTER 14

- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). Wiley.
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2017). *Cross-Battery Assessment Software System 2.0 (X-BASS)*. Wiley.
- Flanagan, D. P., & Schneider, W. J. (2016). Cross-battery assessment? XBA PSW? A case of mistaken identity: A commentary on Kranzler and colleagues' "Classification agreement analysis of cross-battery assessment in the identification of specific learning disorders in children and youth." *International Journal of School and Educational Psychology*, 4, 137–145.
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2019). *Learning disabilities: From identification to intervention* (2nd ed.). Guilford Press.
- Fletcher, J. M., & Miciak, J. (2017). Comprehensive cognitive assessments are not necessary for the identification and treatment of learning disabilities. *Archives of Clinical Neuropsychology*, 32, 2–7.
- Fletcher, J. M., Stuebing, K. K., Morris, R. D., & Lyon, G. R. (2013). Classification and definition of learning disabilities: A hybrid model. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 33–50). Guilford Press.
- Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, 38, 98–108.
- Fuchs, D., Compton, D. L., Fuchs, L. S., & Bryant, J. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study at the National Research Center on Learning Disabilities. *Reading and Writing: An Interdisciplinary Journal*, 21, 413–436.

CHAPTER 14

- Fuchs, D., Hale, J. B., & Kearns, D. M. (2011). On the importance of a cognitive processing perspective: An introduction. *Journal of Learning Disabilities, 44*, 99–104.
- Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204–219.
- Fuchs, L. S., Compton, D. L., Fuchs, D. Paulsen, K. Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*, 493–513.
- Glutting, J. J., Watkins, M. W., & Youngstrom, E. A. (2003). Multifactorial and cross-battery assessments: Are they worth the effort? In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence aptitude, and achievement* (2nd ed., pp. 343–374). Guilford Press.
- Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., . . . Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disabilities Quarterly, 33*, 223–236.
- Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. Guilford Press.
- Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools, 43*, 753–770.
- Hale, J. B., Wycoff, K. L., & Fiorello, C. A. (2011). RTI and cognitive hypothesis testing for identification and intervention of specific learning disabilities: The best of both worlds. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability*

CHAPTER 14

- identification* (pp. 173–201). Wiley.
- Hallahan, D. P., & Mercer, C. D. (2002). Learning disabilities: Historical perspectives. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 1–67). Erlbaum.
- Hallahan, D. P., Pullen, P. C., & Ward, D. (2013). Brief history of the field of learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 15–32). Guilford Press.
- Hammill, D. D., & Larsen, S. C. (1974). The effectiveness of psycholinguistic training. *Exceptional Children, 41*, 429–436.
- Harris, A. J., & Shakow, D. (1937). The clinical significance of numerical measures of scatter on the Stanford–Binet. *Psychological Bulletin, 34*, 134–150.
- Head, H. (1926). *Aphasia and kindred disorders of speech*. Cambridge University Press.
- Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). *The ABCs of CBM* (2nd ed.). Guilford Press.
- Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33*, 3–18.
- Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. Wiley.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. Wiley.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children—Second Edition*. American Guidance Service.
- Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement* (3rd ed.). NCS Pearson.
- Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.). Wiley.

CHAPTER 14

Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2016). *Intelligent testing with the WISC-V*. Wiley.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly*, 7, 136–156.

Kavale, K. A., Kauffman, J. M., Bachmeier, R. J., & LeFever, G. B. (2008). Response-to-intervention: Separating the rhetoric of self-congratulation from the reality of specific learning disability identification. *Learning Disability Quarterly*, 31, 135–150.

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Sage.

Kranzler, J. H., Floyd, R. G., Benson, N., Zaloski, B., & Thibodaux, L. (2016a). Classification agreement analysis of cross-battery assessment in the identification of specific learning disorders in children and youth. *International Journal of School and Educational Psychology*, 4, 124–136.

Kranzler, J. H., Floyd, R. G., Benson, N., Zaloski, & Thibodaux, L. (2016b). Cross-battery assessment pattern of strengths and weaknesses approach to the identification of specific learning disorders: Evidence-based practice or pseudoscience? *International Journal of School and Educational Psychology*, 4, 146–157.

Kirk, S. A. (1962). *Educating exceptional children*. Houghton Mifflin.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1961). *Illinois Test of Psycholinguistic Abilities*. University of Illinois Press.

Lewandowski, L. J., & Lovett, B. J. (2014). Learning disabilities. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (3rd ed., pp. 625–669). Guilford Press.

Lichtenstein, R. (2014). Best practices in identifying learning disabilities. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative*

CHAPTER 14

- decision making* (pp. 331–354). National Association of School Psychologists.
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2006). Why questionable psychological tests remain popular. *Scientific Review of Alternative Medicine, 10*, 6–15.
- Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.
- Lyon, G. R., Fletcher, J. M., Fuchs, L. S., & Chhabra, V. (2006). Learning disabilities. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (3rd ed., pp. 512–591). Guilford Press.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19–40.
- Macmann, G. M., & Barnett, D. W. (1997). Myth of the master detective: Reliability of Interpretations for Kaufman’s “intelligent testing” approach to the WISC-III. *School Psychology Quarterly, 12*, 197–234.
- MacMillan, D. L., Gresham, F. M., & Bocian, K. M. (1998). Discrepancy between definitions of learning disabilities and school practices: An empirical investigation. *Journal of Learning Disabilities, 31*, 314–326.
- Maki, K. M., Burns, M. K., & Sullivan, A. (2017). Learning disability identification consistency: The impact of methodology and student evaluation data. *School Psychology Quarterly, 32*, 254–267.
- Maki, K. M., Burns, M. K., & Sullivan, A. L. (2018). School psychologists’ confidence in learning disability identification decisions. *Learning Disability Quarterly, 41*, 243–256.
- Maki, K. M., Floyd, R. G., & Roberson, T. (2015). State learning disability eligibility criteria: A comprehensive review. *School Psychology Quarterly, 30*, 457–469.
- Mann, L. (1979). *On the trail of process*. Grune & Stratton.
- Mather, N., & Schneider, D. (2015). The use of intelligence tests in the diagnosis of specific

CHAPTER 14

- reading disability. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 415–434). Springer.
- 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.
- McDermott, P. A., Fantuzzo, J. W., Glutting, J. J., Watkins, M. W., & Baggaley, A. R. (1992). Illusions of meaning in the ipsative assessment of children's ability. *Journal of Special Education*, 25, 504–526.
- McDermott, P. A., Goldberg, M. M., Watkins, M. W., Stanley, J. L., & Glutting, J. J. (2006). A nationwide epidemiologic modeling study of LD: Risk, protection, and unintended impact. *Journal of Learning Disabilities*, 39, 230–251.
- McGill, R. J. (2016). Invalidating the full scale IQ score in the presence of significant factor score variability: Clinical acumen or clinical illusion? *Archives of Assessment Psychology*, 6(1), 49–79.
- McGill, R. J. (2018). Confronting the base rate problem: More ups and downs for cognitive scatter analysis. *Contemporary School Psychology*, 22, 384–393.
- McGill, R. J., & Busse, R. T. (2017). When theory trumps science: A critique of the PSW model for SLD identification. *Contemporary School Psychology*, 21, 10–18.
- McGill, R. J., Conoyer, S. J., & Fefer, S. (2018). Elaborating on the linkage between cognitive and academic weaknesses: Using diagnostic efficiency statistics to inform PSW assessment. *School Psychology Forum*, 12, 118–132.
- McGill, R. J., & Dombrowski, S. C. (2018). Factor structure of the CHC model for the KABC-II: Exploratory factor analyses with the 16 core and supplemental subtests. *Contemporary*

CHAPTER 14

- School Psychology*, 22, 279–293.
- McGill, R. J., Dombrowski, S. C., & Canivez, G. L. (2018). Cognitive profile analysis in school psychology: History, issues, and continued concerns. *Journal of School Psychology*, 71, 108–121.
- McGill, R. J., Styck, K. S., Palomares, R. S., & Hass, M. R. (2016). Critical issues in specific learning disability identification: What we need to know about the PSW model. *Learning Disability Quarterly*, 39, 159–170.
- McGrady, H. J. (2002). A commentary on “Empirical and Theoretical Support for Direct Diagnosis of Learning Disabilities by Assessment of Intrinsic Processing Weaknesses.” In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 623–651). Erlbaum.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive–achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, 47, 651–675.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46, 806–834.
- Miciak, J., Fletcher, J. M., Stuebing, K. K., Vaughn, S., & Tolar, T. D. (2014). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly*, 29, 21–37.
- Miciak, J., Taylor, W. P., Denton, C. A., & Fletcher, J. M. (2015). The effect of achievement test selection on identification of learning disabilities within a patterns of strengths and

CHAPTER 14

- weaknesses framework. *School Psychology Quarterly*, 30, 321–334.
- Miciak, J., Taylor, W. P., Stuebing, K. K., & Fletcher, J. M. (2018). Simulation of LD identification accuracy using a patterns of processing strengths and weaknesses method with multiple measures. *Journal of Psychoeducational Assessment*, 36, 21–33.
- Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Korne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS One*, 9(7), 1–8.
- Monroe, M. (1932). *Children who cannot read*. Chicago: University of Chicago Press.
- Naglieri, J. A. (1999). *Essentials of CAS assessment*. Wiley.
- Naglieri, J. A. (2011). The discrepancy/consistency approach to SLD identification using the PASS theory. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 145–172). Wiley.
- Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, 13, 248–277.
- O'Donnell, P. S., & Miller, D. N. (2011). Identifying students with specific learning disabilities: School psychologists' acceptability of the discrepancy model versus response to intervention. *Journal of Disability Policy Studies*, 22, 83–94.
- Orton, S. T. (1925). "Word-blindness" in school children. *Archives of Neurology and Psychiatry*, 14, 581–615.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9, 105–119.
- Peterson, K. M. H., & Shinn, M. H. (2002). Severe discrepancy models: Which best explains school identification practices for learning disabilities? *School Psychology Review*, 31, 459–476.

CHAPTER 14

- Rapaport, D., Gil, M., & Schafer, R. (1945). *Diagnostic psychological testing: The theory, statistical evaluation, and diagnostic application of a battery of tests* (Vol. 1). Yearbook Medical.
- Restori, A. F., Gresham, F. M., & Cook, C. R. (2008). “Old habits die hard”: Past and current issues pertaining to response-to-intervention. *California School Psychologist, 13*, 67–78.
- Reynolds, C. R. (1984–1985). Critical measurement issues in learning disabilities. *Journal of Special Education, 18*, 451–476.
- Reynolds, C. R., & Shaywitz, S. E. (2009). Response to intervention: Ready or not? Or, from wait-fail to watch-them fail. *School Psychology Quarterly, 24*, 130–145.
- Riley-Tillman, T. C., & Burns, M. K. (2009). *Evaluating educational interventions: Single-case design for measuring response to intervention*. Guilford Press.
- Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry, 16*, 181–197.
- Schneider, W. J., & Kaufman, A. S. (2017). Let’s not do away with comprehensive cognitive assessments just yet. *Archives of Clinical Neuropsychology, 32*, 8–20.
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73–163). Guilford Press.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock–Johnson IV*. Riverside.
- Schrank, F. A., Stephens-Pisecco, T. L., & Schultz, E. K. (2017). *The WJ IV core-selective evaluation process applied to identification of a specific learning disability* (Woodcock–Johnson IV Assessment Service Bulletin No. 8). Houghton Mifflin/Harcourt.
- Schrank, F. A., & Wendling, B. J. (2018). The Woodcock–Johnson IV: Tests of cognitive abilities, tests of oral language, tests of achievement. In D. P. Flanagan & E. M.

CHAPTER 14

- McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 383–451). Guilford Press.
- Schroeder, M., Drefs, M. A., & Cormier, D. C. (2017). The messiness of LD identification: Contributions of diagnostic criteria and clinical judgment. *Canadian Psychology, 58*, 218–227.
- Schultz, E. K., & Stephens-Pisecco, T. L. (2015). Core-Selective Evaluation Process: An efficient and comprehensive approach to identify students with SLD using the WJ IV. *The Dialog, 44*(2), 5–12.
- Scruggs, T. E., & Mastropieri, M. A. (2002). On babies and bathwater: Addressing the problems of identification of learning disabilities. *Learning Disability Quarterly, 25*, 155–168.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine, 326*, 145–150.
- Smith, C. B., & Watkins, M. W. (2004). Diagnostic utility of the Bannatyne WISC-III pattern. *Learning Disabilities Research and Practice, 19*, 49–56.
- Sotelo-Dynega, M., & Dixon, S. G. (2014). Cognitive assessment practices: A survey of school psychologists. *Psychology in the Schools, 51*, 1031–1045.
- Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*, 735–749.
- Stephens-Pisecco, T. L., & Schultz, E. K. (2019). *Core-Selective Evaluation Process (C-SEP): Overview and procedures*. Authors.
- Streiner, D. L. (2002). Breaking up is hard to do: The heartbreak of dichotomizing continuous data. *Canadian Journal of Psychiatry, 47*, 262–266.
- Streiner, D. L. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests.

CHAPTER 14

- Journal of Personality Assessment*, 81, 209–219.
- Stuebing, K. K., Barth, A. E., Molfese, P. J., Weiss, B., & Fletcher, J. M. (2009). IQ is not strongly related to response to reading instruction: A meta-analytic interpretation. *Exceptional Children*, 76, 31–51.
- Stuebing, K. K., Fletcher, J. M., Branum-Martin, L., & Francis, D. J. (2012). Evaluation of the technical adequacy of three methods for identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review*, 41, 3–22.
- Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities. A meta-analysis. *American Educational Research Journal*, 39, 469–518.
- Taylor, W. P., Miciak, J., Fletcher, J. M., & Francis, D. J. (2017). Cognitive discrepancy models for specific learning disabilities identification: Simulations of psychometric limitations. *Psychological Assessment*, 29, 446–457.
- Truscott, S. D., Baumgart, M. B., & Rogers, K. M. (2004). Financial conflicts of interest in the school psychology assessment literature. *School Psychology Quarterly*, 19, 166–178.
- U.S. Office of Education. (1976). Proposed rulemaking. *Federal Register*, 41(230), 52404–52407.
- U.S. Office of Education. (1977). Assistance to the states for education of handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register*, 42(250), 65082–65085.
- VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, 45, 225–256.
- Van Norman, E. R., Maki, K. E., Burns, M. K., McComas, J. J., & Helman, L. (2018).

CHAPTER 14

- Comparison of progress monitoring data from general outcome measures and specific subskill mastery measures for reading. *Journal of School Psychology*, 67, 178–189.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601–638.
- Ventura County SELPA. (2017). The Ventura County SELPA pattern of strengths and weaknesses model for specific learning disability eligibility procedural manual. Retrieved from www.venturacountyselpa.com
- Wanzek, J., Vaughn, S., Scammacca, Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after Grade 3. *Review of Education Research*, 83, 163–195.
- Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly*, 15, 465–479.
- Watkins, M. W. (2003). IQ subtest analysis: Clinical acumen or clinical illusion? *Scientific Review of Mental Health Practice*, 2, 118–141.
- Watkins, M. W. (2005). Diagnostic validity of Wechsler subtest scatter. *Learning Disabilities: A Contemporary Journal*, 3, 20–29.
- Watkins, M. W., & Canivez, G. L. (2004). Temporal stability of WISC-III subtest composite: Strengths and weaknesses. *Psychological Assessment*, 16, 133–138.
- Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997a). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools*, 34, 309–319.

CHAPTER 14

- Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997b). Prevalence and diagnostic utility of the WISC-III SCAD profile among children with disabilities. *School Psychology Quarterly*, 12, 235–248.
- Watkins, M. W., Kush, J. C., & Schaefer, B. A. (2002). Diagnostic utility of the WISC-III learning disability index. *Journal of Learning Disabilities*, 35, 98–103.
- Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). NCS Pearson.
- Williams, J., & Miciak, J. (2018). Adoption costs associated with processing strengths and weaknesses methods for learning disabilities identification. *School Psychology Forum*, 12, 17–29.
- Williams, J. L., Miciak, J., McFarland, L., & Wexler, J. (2016). Learning disability identification criteria and reporting in empirical research: A review of 2001–2013. *Learning Disabilities Research and Practice*, 31, 221–229.
- Wilson, V. L. (1987). Statistical and psychometric issues surrounding severe discrepancy. *Learning Disabilities Research*, 3, 24–28.
- World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Author.
- Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, 39, 204–221.
- Youngstrom, E. A., & Van Meter, A. (2016). Empirically supported assessments of children and adolescents. *Clinical Psychology: Science and Practice*, 23, 327–347.
- Ysseldyke, J. E., Thurlow, M., Graden, J., Wesson, C., Algozzine, B., & Deno, S. L. (1983). Generalizations from five years of research on assessment and decision making: The University of Minnesota Institute. *Exceptional Education Quarterly*, 4(1), 75–93.

CHAPTER 14

Zimmerman, I. L., & Woo-Sam, J. M. (1985). Clinical applications. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 873–898).

Wiley.

Zirkel, P. A. (2013). A comprehensive evaluation of the Supreme Court's Forest Grove decision? *Journal of Psychoeducational Assessment, 31*, 313–317

Appendix LD Identification Assessment Case Study

Background

Matthew Smith, an 8-year-old second-grade student, is experiencing significant reading and behavioral difficulties at school. Mathew was born prematurely at 34 weeks gestation. Background information did not reveal any immediate adverse development delays other than skipping the babbling stage preceding talking. Matthew attended a Montessori preschool, and reports from the school did not reveal academic or behavioral struggles at that time. Upon primary school entry in kindergarten, Matthew was observed to struggle with basic reading skills during early literacy development and task persistence when assigned independent work. These difficulties have persisted to the present time period. Matthew's pediatrician referred Matthew to a child psychiatrist for his struggles with attention-deficit/hyperactivity disorder (ADHD). Ms. Smith thought it was a good idea to obtain a non-school-based evaluation to address the reading difficulties prior to the meeting with the psychiatrist.

Psychological Report—Confidential

Reason for Referral

Matthew was referred for a comprehensive evaluation (1) to gain insight into his present level of functioning, (2) to ascertain diagnostic impressions, and (3) to determine treatment recommendations and accommodations that might be appropriate for him. Background information and teacher reports reveal that Matthew struggles with all aspects of reading. Specifically, Matthew experiences difficulty with word decoding, oral reading fluency, and reading comprehension. Additionally, Matthew's faces difficulties with attention, hyperactivity, and distractibility.

Assessment Methods and Sources of Data

Reynolds Intellectual Assessment Scale—Second Edition (RIAS-2)

CHAPTER 14

Woodcock–Johnson Tests of Achievement—Fourth Edition (WJ-IV)

Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)

Behavior Assessment System for Children, Third Edition (BASC-3)

Child Development Questionnaire (CDQ)

ADHD Rating Scale, Fifth Edition

Teacher Interview

Parent Interview

Child Interview

Classroom Observations

Review of Academic Records

Review of School Records

Review of Pediatrician's Records

Comment: Background information reveals that there are additional concerns related to attention and suspicion that Matthew may have ADHD. Therefore, additional attention rating scales have been added to the assessment battery to shed insight on potential comorbidity with LD.

Background Information and Developmental History

Matthew Smith is an 8-year-old child in the second grade at the Cherry Hill Public School District (CHPS). Matthew faces difficulty with paying attention, organization, and remaining on task. Background information also revealed difficulty with reading comprehension. Ms. Smith, Matthew's mother, indicates that Matthew has been diagnosed by his pediatrician with ADHD. Ms. Smith noted that Matthew's pediatrician offered the diagnosis based on his observation and a computerized test on which Matthew scored poorly on a measure of attention. Matthew is not presently taking any medication for the management of his symptoms, but Ms. Smith reports that he is scheduled for a psychiatric evaluation in mid-July at the ADHD Clinic

CHAPTER 14

within the Children's Hospital of Philadelphia. She explained that she sought this psychological evaluation to gain a better sense of Matthew's overall functioning prior to the psychiatric evaluation.

Prenatal, Perinatal, and Early Developmental History

Ms. Smith experienced gestational diabetes during her pregnancy with Matthew. Matthew was born at 34 weeks weighing 6 pounds, 5 ounces. He had a 2-week stay in the neonatal intensive care unit (NICU) for jaundice. Matthew's early developmental milestones were generally attained within normal limits, with the exception of skipping the babbling stage that precedes learning to talk. Matthew attended a Montessori school and was determined to be developing at an age-expected level with no behavioral, social, or academic concerns noted.

Medical

Ms. Smith indicated that Matthew is not currently taking any medications but suffers from occasional migraines (averaging six times per month). Matthew will be evaluated by a child psychiatrist in mid-July. Matthew's hearing and vision are intact. Ms. Smith reported no prior incidence of head injury, accident, or major infection.

Cognitive, Academic, and Language Functioning

Background information and teacher reports indicate that Matthew struggles with all aspects of reading. Ms. Smith commented that Matthew's reading is below grade level, noting that he rarely stays focused on the lesson or during independent reading time. Teacher reports confirmed difficulties with these academic areas. Background information suggests that Matthew's guided reading level is approximately two grades below where he is expected to be at this point in the school year. Background reports indicated that Matthew's mathematics progress is at grade level. Results of Matthew's state benchmark testing for last spring were as follows: reading (5th percentile), written expression (8th percentile) and mathematics (25th percentile).

CHAPTER 14

Ms. Jones indicated that Matthew has a good understanding of basic mathematics skills but struggles when required to sustain attention in multistep math problems. She noted that his writing progress is slightly below that of other students his age.

Comment: Teacher reports and review of academic records indicate that Matthew is presently functionally impaired in reading and that his performance in mathematics is at expected levels, indicating these deficits are specific in nature. These deficits began to occur almost immediately at the onset of early literacy instruction during kindergarten, persisted since that time during the developmental period, and have persisted for longer than 6 months (DSM-5 criteria).

Social–Emotional and Behavioral Functioning

Ms. Jones indicated that Matthew is a pleasant but active and impulsive child. Every once in a while, he becomes upset and stomps the ground, throws his pencil, or slams his book on the table. Ms. Jones explained that he usually calms down after being redirected, but he can sulk for a long time. She noted that Matthew needs to pay attention to whole-class instruction, stay focused on independent work, keep his body under control by not bothering those around him, and listen to and follow the teacher’s instructions. Ms. Jones noted that Matthew has been sent to the ReSet room numerous times, and this sometimes seems to help his behavior. Both Ms. Jones and Ms. Smith report that Matthew has many friends, but they are sometimes annoyed by his tendency to intrude into their activities, interrupt what they are saying, and avoid waiting for his turn. Both Ms. Smith and Ms. Jones are concerned that Matthew may become alienated from his peers at school.

Comment: Although there are clearly concerns with reading, it is presently unclear whether these deficits are due to LD or are a collateral deficit of Matthew’s attention difficulties; it is common for children with ADHD to also present with academic deficits consistent with LD.

Strengths

CHAPTER 14

Matthew's strengths include helping out around the classroom when asked by a teacher. He has also been described as a child with good artistic ability, a good sense of humor, and one who enjoys life. Matthew enjoys playing video games, soccer, and basketball.

Summary

Matthew struggles with symptoms of inattentiveness, hyperactivity, and impulsivity. He also struggles with all aspects of reading. Matthew enjoys helping in the classroom and is motivated to do well at school. He struggles with both academic and behavioral functioning at school.

Interview Results

Parent Interview

Ms. Smith was contacted to ascertain impressions of Matthew's academic, behavioral and social-emotional functioning at school. Ms. Smith commented first regarding Matthew's behavioral issues. She noted that he has been diagnosed with ADHD by his pediatrician, who suggested both psychological and psychiatric evaluations. Ms. Smith explained that the wait for the psychological evaluation at CHOP was over a year, so she sought out services for that elsewhere. Ms. Smith indicated that her concerns are "more behavioral than academic." Ms. Smith indicated that Matthew has a few instances of not listening, not controlling his body, and being put in time out from recess. She explained that Matthew is a generally kind child but struggles to control his activity level and exuberance for life. Commenting next on his social progress, Ms. Smith indicated that socially Matthew is OK but she has noticed that his friends are beginning to become frustrated by his tendency to jump into and intrude in their activities at inappropriate times. Commenting next on Matthew's academic progress, Ms. Smith noted that Matthew struggles with reading, and his handwriting can be sloppy. She discussed Matthew's areas of strengths and needs, noting that he can focus for long periods if he is interested in a topic

CHAPTER 14

or activity. For instance, Ms. Smith explained that Matthew can play video games for hours on end. She also indicated that Matthew likes to draw and play outside. She noted that math is one of his strengths. She also mentioned that Matthew is very athletic and loves to play all types of sports.

Child Interview

Matthew was interviewed to ascertain impressions of his overall social, emotional, and behavioral progress. Matthew stated that he enjoys school, especially science, art, and recess. When asked about his academic progress, Matthew explained that reading is sometimes difficult for him, but mathematics and writing are good. Matthew next discussed his behavior at school. He explained that he sometimes gets into trouble for “doing something bad.” When asked to elaborate, Matthew indicated that he once went to the principal’s office for making noises with his throat. He did not discuss any additional behavioral incidents. When asked about his mood or feelings, Matthew noted that he is generally a happy child who enjoys going outside and playing sports.

Teacher Interview

Ms. Jones, Matthew’s second-grade teacher, was interviewed regarding Matthew’s academic, behavioral, emotional, and social functioning. She explained that Matthew struggles with reading, and she is really concerned about this. Ms. Jones indicated that any language-based topic is difficult for Matthew. She explained that Matthew had regressed in reading this past summer and had forgotten many basic reading skills from the prior year. Ms. Jones noted that phonological skills are a problem for Matthew. She also mentioned that Matthew has low sight-word knowledge and places at the primer level. Ms. Jones explained that Matthew’s writing skills are below those of his peers, but not as low as his reading abilities. She explained that Matthew can demonstrate some level of competence if a teacher is able to sit with him and coach

CHAPTER 14

him. Commenting on Matthew's behavioral progress, Ms. Jones noted that Matthew struggles with distractibility, remaining on task, loss of focus, and a high activity level. She explained that Matthew generally gets along with other children in the classroom, but other children have recently become frustrated by his tendency to interrupt them, to avoid waiting his turn, and to intrude into their activities. Ms. Jones noted that Matthew is an athletic child, and this supports his social development at school. However, Ms. Jones explained that Matthew sulks when denied his own way but eventually comes around and regains his focus. She indicated that Matthew is motivated to do well in school despite his academic difficulties.

Comment: Additional corroborating information through teacher report indicates that Matthew is functionally impaired in reading.

Teacher Interview

Ms. Mia Riley, reading specialist, was asked to furnish her impressions of Matthew's progress in school. Matthew sees the reading specialist three times per week for 30 minutes. Ms. Riley indicated that Matthew is eager to learn and willing to try whatever is put in front of him; however, Matthew struggles with phonological awareness and with sight-word decoding. Ms. Riley comments that this impacts his comprehension of written text. Additionally, she stated that Matthew struggles with spelling and writing at an expected grade level, but this is not as much a concern as his reading capabilities. Ms. Riley explained that she has been working on fostering Matthew's basic understanding of phonemic awareness skills.

Comment: Clinicians in private practice settings may not have the resources to facilitate targeted interventions for children who are referred for LD evaluations and explicitly evaluate a student's response to instruction. However, many students who experience academic difficulties (especially in reading) are provided with some form of targeted assistance beyond the general education classroom, thus permitting inferences pertaining to RTI. Here we have evidence that

CHAPTER 14

Matthew has been provided with supplemental interventions through the reading specialist three times a week, which is roughly equivalent to a Tier 2 RTI intervention. Given the persistence of Matthew's difficulties, it is clear that this intervention has not been successful at remediating his reading problems (DSM-5 criteria).

Observations

Classroom Observation

Matthew was observed for 15 minutes in Ms. Jones's classroom. The observation occurred during a reading activity in which students were instructed on how to make connections between books with a partner. During this whole-group instruction, Matthew was observed to sit attentively and listen to Ms. Jones. When the activity shifted and students were asked to partner with another student to share their connections, Matthew again complied with this request. Impressions of the observation were that Matthew was on task and compliant with teacher requests.

Observation during Assessment

Matthew was very compliant during the beginning of assessment, though he struggled in both cognitive and achievement tests. Matthew grew frustrated during the Passage Comprehension subtest of the WJ-IV and received encouragement for his efforts on this subtest. Matthew responded well to encouragement and was engaged in the subtest. He needed two breaks during the assessment session. Matthew was oriented to person, place, and time. He denied any feeling of suicidal or homicidal ideation. Test results are considered a valid representation of Matthew's abilities.

Cognitive and Academic Functioning

Reynolds Intellectual Assessment Scale—Second Edition (RIAS-2)

CHAPTER 14

Matthew was administered the RIAS-2, which is an individually administered measure of intellectual functioning, normed for individuals between ages 3 and 94 years. The RIAS-2 contains several individual tests of intellectual problem solving and reasoning ability that are combined to form a Verbal Intelligence Index (VIX) and a Nonverbal Intelligence Index (NIX). The subtests that comprise the VIX assess verbal reasoning ability, along with the ability to access and apply prior learning in solving language-related tasks. Although labeled the VIX, it is also a reasonable approximation of crystallized intelligence. The NIX comprises subtests that assess nonverbal reasoning and spatial ability. Although labeled the NIX, it also provides a reasonable approximation of fluid intelligence and spatial ability. These two indices of intellectual functioning are then combined to form an overall Composite Intelligence Index (CIX). By combining the VIX and the NIX into the CIX, a strong, reliable assessment of general intelligence (g) is obtained. The CIX measures the two most important aspects of g according to recent theories and research findings: reasoning or fluid abilities and verbal or crystallized abilities.

The RIAS-2 also contains subtests designed to assess verbal memory and nonverbal memory. Depending on the age of the individual being evaluated, the Verbal Memory subtest comprises a series of sentences, age-appropriate stories, or both, read aloud to the examinee. The examinee is then asked to recall these sentences or stories as precisely as possible. The Nonverbal Memory subtest comprises the presentation of pictures of various objects or abstract designs for a period of 5 seconds. The examinee is then shown a page containing six similar objects or figures and must discern which object or figure was previously shown. The scores from these subtests are combined to form a Composite Memory Index (CMX), which provides a strong, reliable assessment of working memory and may also provide indications as to whether a more

CHAPTER 14

detailed assessment of memory functions may be required. In addition, the high reliability of the Verbal and Nonverbal Memory subtests allows direct comparison of them to each other.

Each of these indices is expressed as an age-corrected standard score that is scaled to a mean of 100 and a standard deviation of 15. These scores are normally distributed and can be converted to a variety of other metrics if desired.

Following are the results of Matthew's performance on the RIAS-2.

| | Composite IQ | Verbal IQ | Nonverbal IQ | Memory Index |
|---------------------------|--------------|-----------|--------------|--------------|
| RIAS-2 Index | 84 | 79 | 100 | 97 |
| Percentile | 14 | 8 | 50 | 42 |
| Confidence interval (95%) | 79–90 | 73–87 | 45–57 | 91–103 |

On testing with the RIAS-2, Matthew attained a CIX of 84. On the RIAS-2, this level of performance falls within the range of scores designated as low average and exceeded the performance of 14% of individuals who were Matthew's age. His Verbal IQ (Standard Score = 79; 8th %ile) was in the below-average range and exceeded 8% of individuals who were Matthew's age. Matthew's Nonverbal IQ (Standard Score = 93; 32nd %ile) was in the average range, exceeding 32% of individuals Matthew's age. Matthew earned a CMX of 97, which falls

CHAPTER 14

within the average range of working memory skills and exceeds the performance of 42 out of 100 individuals Matthew's age.

Comment: The RIAS-2 was administered to rule out intellectual deficiency (ID) as a potential explanation for Matthew's reading difficulties. Interpretation is focused mostly on the CIX, as that is the most reliable and valid score on the RIAS-2, and little, if any consideration is given to the interpretation of subscale scores, although they are reported to comport with ethical test standards. These results indicate that Matthew has enough ability to benefit from the instructional environment; thus, cognitive ability is not considered to be the reason he is struggling to learn how to read.

Woodcock–Johnson Tests of Achievement IV (WJ-IV)

The WJ-IV is an achievement test used to measure reading, writing, and mathematics skills. The Reading Index includes letter and word identification, vocabulary, and comprehension skills. The Writing Index includes spelling, writing fluency, and simple sentence writing. The Mathematics Index includes calculation, practical problems, and knowledge of mathematical concepts and vocabulary.

Matthew obtained the following scores in each of the areas of measurement:

| | <u>Standard</u> | <u>Percentile</u> | <u>95% CI</u> | <u>Descriptive classification</u> |
|---|-----------------|-------------------|----------------|-----------------------------------|
| <i>Broad Reading Composite</i> | 79 | 8 | <1–6 | Below average |
| Letter–word ID | 89 | 23 | 2–8 | Low average |
| Passage comprehension | 79 | 8 | 2–11 | Below average |
| Word attack | 79 | 8 | 1–31 | Below average |
| Oral reading | 87 | 20 | <1–4 | Low average |
| Sentence reading fluency | 77 | 7 | <1–20 | Below average |
| <i>Broad Written Language</i> | 87 | 17 | 9–22 | Low average |
| Sentence writing fluency | 83 | 12 | 5–40 | Low average |
| Writing samples | 82 | 11 | 5–45 | Low average |
| Spelling | 92 | 30 | 2–19 | Average |
| <i>Broad Mathematics Composite</i> | 93 | 32 | 9–31 | Average |
| Math facts fluency | 95 | 38 | 3–38 | Average |
| Applied problems | 99 | 48 | 4–33 | Average |

CHAPTER 14

| | | | | |
|-------------|----|----|-------|---------|
| Calculation | 90 | 25 | 23–53 | Average |
|-------------|----|----|-------|---------|

Standardized achievement test results revealed below-average performance across the broad reading composite, with low average performance on the broad written language performance. Matthew scored in the average range on the broad mathematics composite.

Comment: The below-average reading scores on the WJ-IV provide evidence of low achievement and that Matthew's reading difficulties are also a normative deficit. At this point, both of the dual-deficit criteria are satisfied.

Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)

The CTOPP-2 is a standardized test of phonological processing that yield three composite scores: (1) Phonological Awareness, (2) Phonological Memory, and (3) Rapid Naming. The Phonological Awareness composite measures a student's ability to access the phonological structure of oral language. The Phonological Memory composite measures the ability to code information phonologically for temporary storage in working or short-term memory. The Rapid Naming Composite measures a student's ability to retrieve phonological information from memory and to complete a sequence of operations quickly and repeatedly. Matthew's performance across the three index composite areas was as follows:

| | Scaled score | %ile | Description |
|------------------------|--------------|------|---------------|
| Phonological Awareness | 79 | 7 | Below average |
| Phonological Memory | 100 | 50 | Average |
| Rapid Naming | 78 | 8 | Below average |

CHAPTER 14

Matthew's profile on the CTOPP-2 revealed a child who falls within the below-average range on the Phonological Awareness and Rapid Naming subtests, and in the average range on the Phonological Memory subtest. The current test administration appears to provide an accurate estimate of Matthew's present phonological processing.

Comment: The CTOPP-2 provides direct assessment of one of the most important early literacy subskills: phonological awareness. These results indicate that Matthew has a normative deficit in this skill. This deficit is likely the reason for his reading difficulties. Intervention efforts should target phonics development to remediate this skill area.

Social–Emotional and Behavioral Functioning

Behavior Assessment System for Children, Third Edition (BASC-3)

The BASC-3 is an integrated system, designed to facilitate the differential diagnosis and classification of a variety of emotional and behavioral conditions in children. It possesses validity scales and several clinical scales that reflect different dimensions of a child's personality. *T*-scores between 40 and 60 are considered average. Scores greater than 70 ($T > 70$) are in the clinically significant range and suggest a high level of difficulty. Scores in the at-risk range (*T*-score 60–69) identify either a significant problem that may not be severe enough to require formal treatment or the potential to develop a problem that needs careful monitoring. On the Adaptive Scales, scores below 30 are considered clinically significant, while scores between 31 and 39 mean that the child is at risk.

CHAPTER 14

| | Ms. Jones | | Ms. Smith | |
|--------------------------|-----------------|------------|-----------------|------------|
| | <i>T</i> -score | Percentile | <i>T</i> -score | Percentile |
| <u>Clinical scales</u> | | | | |
| Hyperactivity | 85** | 99 | 72** | 95 |
| Aggression | 59 | 82 | 58 | 80 |
| Conduct Problems | 55 | 74 | 51 | 55 |
| Anxiety | 43 | 26 | 54 | 67 |
| Depression | 47 | 51 | 59 | 85 |
| Somatization | 46 | 46 | 45 | 44 |
| Attention Problems | 73** | 99 | 75** | 96 |
| Learning Problems | 73** | 99 | 73** | 98 |
| Atypicality | 43 | 19 | 45 | 23 |
| Withdrawal | 55 | 74 | 53 | 70 |
| Adaptability | 33* | 3 | 32* | 5 |
| Social Skills | 40 | 20 | 40 | 22 |
| Leadership | 47 | 43 | 45 | 40 |
| Study Skills | 38* | 15 | — | — |
| Functional Communication | 43 | 25 | 44 | 2 |

Composite scores

| | | | | |
|------------------------|-----|----|----|----|
| Externalizing Problems | 63* | 88 | 54 | 67 |
| Internalizing Problems | 44 | 30 | 45 | 32 |

CHAPTER 14

| | | | | |
|---------------------|-----|----|----|----|
| School Problems | 67* | 94 | — | — |
| Behavioral Symptoms | 56 | 75 | 54 | 66 |
| Index | | | | |
| Adaptive Skills | 38* | 12 | 40 | 14 |

Note. For clinical scales, *at-risk rating ($T = 60-69$); **clinically significant rating ($T = 70+$). For composite scores, *at-risk rating ($T = 30-39$); **clinically significant rating ($T < 30$). —Not applicable/assessed on Parent Rating Scale.

These results indicate at-risk elevations on Externalizing Problems, Adaptive Skills, and the School Problems composite, and clinically significant elevations on the Hyperactivity and Attention Problems clinical scales, with an at-risk rating on the Adaptive Skills and Study Skills clinical scales.

ADHD Rating Scale–5

The ADHD Rating Scale–5 comprises ADHD symptoms based on DSM-5 diagnostic criteria. In general, scores between the 85th and 93rd percentiles are considered above average or “at-risk” for symptom cluster compared to the normative sample. Scores above the 93rd percentile are generally considered clinically significant. Matthew received the following scores:

| Scale | Teacher percentile | Parent percentile |
|---------------------------|-------------------------------|-------------------------------|
| Hyperactivity–Impulsivity | 95th (clinically significant) | 95th (clinically significant) |
| Inattention | 94th (clinically significant) | 95th (clinically significant) |
| Combined | 97th (clinically significant) | 97th (clinically significant) |

The ratings on the ADHD Rating Scale–5, across both parent and teacher versions, converged to suggest clinically significant elevations across all scales.

Diagnostic Impressions

Multiple data sources and methods of assessment inform the conceptualization of Matthew's cognitive, academic, social–emotional, and behavioral functioning, including whether he meets criteria for any diagnostic category. Details in support of these findings are offered below.

Cognitive and Academic Functioning

Matthew's on measures of cognitive ability was low average (Composite IQ = 84, 14th %ile; VIQ = 79, 8th %ile; NIQ = 100, 50th %ile). Matthew's performance on the WJ-IV Achievement was below average in reading, low average in writing, and average in mathematics. Matthew's performance on the CTOPP-2 was below average on a measure of phonological processing and rapid naming. Matthew's performance on standardized measures of academic achievement is consistent with teacher reports in which Matthew was noted to struggle with reading. Matthew is performing at grade level in mathematics on both measures of academic achievement and classroom progress.

Social and Emotional Functioning

Matthew is described as a child who struggles with attention, loss of focus, distractibility, impulsivity, and hyperactivity. He has a classification of ADHD, combined type (314.01). This is consistent with BASC-3 results in which he scored in the clinically significant range on the Inattention and Hyperactivity clinical scales. It is also consistent with scores on the ADHD rating scales. Background information and standardized behavior rating scales revealed that Matthew sometimes disregards classroom rules and teacher requests, and needs structure and support for

CHAPTER 14

these difficulties. He sometimes sulks when he does not get his own way. His peers are becoming frustrated by his tendency to intrude at inappropriate times and in inappropriate ways into their activities. Matthew can be a helpful child when a teacher requests his assistance.

Diagnostic Impression Summary

Multiple methods and sources of evaluation including the dual-deficit academic model of LD supported by clinical judgment suggest Matthew has a classification of specific learning disorder (SLD; 315.0) with impairment in reading. Matthew also meets criteria for a diagnosis of ADHD, combined presentation (314.01). It is recommended that Ms. Smith present this report to Matthew's school district, as he will likely qualify for special education support and receive an individualized education plan (IEP) and a Section 504 plan for his difficulties with ADHD.

Comment: Matthew currently meets all of the criteria for LD classification in both "hybrid" approaches and DSM-V. There is information to indicate that he has been provided access to appropriate instruction and intervention supports, and although he has been diagnosed with ADHD, there is evidence to indicate that a separate LD diagnosis is warranted. The most compelling factor is the specific nature of his deficits in reading, with no concomitant weaknesses in mathematics or writing. If his deficits were predominately due to attention, we would expect to see those deficits also manifest in other academic area. Additionally, the results from the CTOPP-2 indicate that Matthew is currently deficient in an important subskill associated with early literacy development.

Summary and Recommendations

Matthew's overall cognitive ability falls within the low-average range. Matthew's performance on measures of academic achievement (WJ-IV) was in the below-average to average range. He qualifies for a diagnosis of SLD (315.0) with impairment in reading. Matthew also experiences a high level of activity, distractibility, impulsivity, and inattention. He qualifies

CHAPTER 14

for a diagnosis of ADHD, combined presentation (314.01). The following home- and school-based recommendations will benefit Matthew. Matthew's school district should consider the diagnostic impressions included in this report and consider offering him special education support for his learning and behavioral difficulties, both of which are adversely impacting his educational performance.

Strategies for Difficulties with Attention, Distractibility, Overactivity, Impulsivity, and Loss of Focus

The following recommendations might be beneficial for Matthew:

1. *Direct contingency management at home and school and clinical cognitive-behavioral therapy to help supplement these efforts.* Both caregivers and teachers should implement home- or school-based contingencies incorporating (a) positive incentives such as praise or individualized reward programs for targeted behavior; (b) negative consequences such as reprimands, response cost, or time out; or (c) combinations of positive and negative contingencies. This approach has been found to be effective in home and school settings. Caregivers specifically may benefit from parent behavioral management training to better understand how to properly implement a direct contingency management program in the home. Experienced teachers generally have the tools to implement, but consultation with the teacher may be warranted.

2. *Additional school-based strategies for ADHD*

- Seating.* Matthew should continue to sit in a location where there are minimal distractions.
- Provision of directions by teachers.* When Matthew's teachers interact with him, he should be encouraged to repeat and explain instructions to ensure understanding. The provision of directions to Matthew will be most effective when the teacher makes eye contact, avoids multiple commands, is clear and to the point, and permits repetition of directions when asked for or needed.

CHAPTER 14

- Positive reinforcement and praise for successful task completion.* Matthew's teachers should provide positive reinforcement and immediate feedback for completion of desired behaviors or tasks. Initially, praise and reinforcement should be offered for successful effort on a task or behavior, regardless of quality of performance.
 - Time on task.* Communicate to Matthew how long he will need to engage in or pay attention on a particular task. Open-ended expectations can be distressing to any child, let alone one with attentional difficulties.
 - Prepare student discreetly for transitions.* Furnish Matthew with verbal prompts and visual cues that a new activity or task is about to start. This should be accomplished discreetly so as to avoid student embarrassment.
 - Recess time.* Matthew should be permitted to participate in recess. Recess should not be a time to complete unfinished classwork or homework.
 - Extended time, teacher check-ins, and frequent breaks.* Matthew should be permitted additional time to complete academic tasks and projects. Matthew's teachers should also consider review of classwork as Matthew progresses on an assignment or project to assist Matthew in avoiding careless mistakes. More frequent breaks than what is typical may also reduce careless mistakes and help to maintain focus.
 - Daily behavior report card.* Key targeted behavior is monitored by the teacher, with home reinforcement contingent on positive teacher report.
3. *Behaviorally oriented social skills training.* This approach has received preliminary empirical support for social skills difficulties in children with ADHD, but not in all studies. Specific behavioral deficits such as intruding in an activity without asking, interrupting other children, and difficulty waiting one's turn in social exchange should be the initial targets for intervention.

CHAPTER 14

The goals of therapy should be conveyed to Matthew's teacher, so that she may reinforce prosocial behaviors or correct behaviors that are targets of intervention.

4. *Child psychiatric evaluation.* Matthew will benefit from the forthcoming child psychiatric evaluation at the Children's Hospital of Philadelphia to determine whether he would benefit from psychotropic medication as an adjunct to his behavioral intervention at school and home.

5. *Support for difficulties with reading comprehension, phonological awareness, sight-word recognition, word decoding, and reading fluency.* Matthew struggles with all aspects of reading, including word decoding, phonological/phonemic awareness, reading fluency, and reading comprehension. Matthew requires special education support for his reading difficulties, as noted below.

- *Phonological awareness and sight-word knowledge skills.* Matthew will benefit from continued intervention with basic phonemic awareness skills, such as emphasizing instruction on basic rhymes (*ack, ame, all, ake*). Matthew would be well served to increase his familiarity with reading fundamentals through a focus on words via alliteration lessons (e.g., tongue twisters), a personal dictionary of sight words (i.e., most frequently used words), and word family study (e.g., *neat, beat, heat; noise, poise, choice*).

- *Reading fluency.* Matthew should practice oral reading fluency. Accordingly, Matthew will benefit from repeated reading of the passage until an appropriate grade-level fluency rate is attained. The research literature suggests that improvements in oral reading fluency via repeated passage reading generalizes to improvements in overall reading ability.

- *Reading comprehension.* Matthew struggles with the comprehension of written text and will benefit from prereading and organizational strategies that attempt to improve this skill area.

Following are a few suggestions that will likely benefit Matthew:

- Before reading, preview the text by looking at the title and illustrations.

CHAPTER 14

- Encourage the creation of a possible story from the illustrations.
- Make predictions about the story based on story features prior to reading the story.
- During reading, generate questions about the story that are directly related to the text and that require thinking beyond the text.
- After reading, spend time reflecting on the material and relating it to experiences and events the child has encountered.
- After reading, have Matthew engage in the reading material using text summarizing.

CHAPTER 14

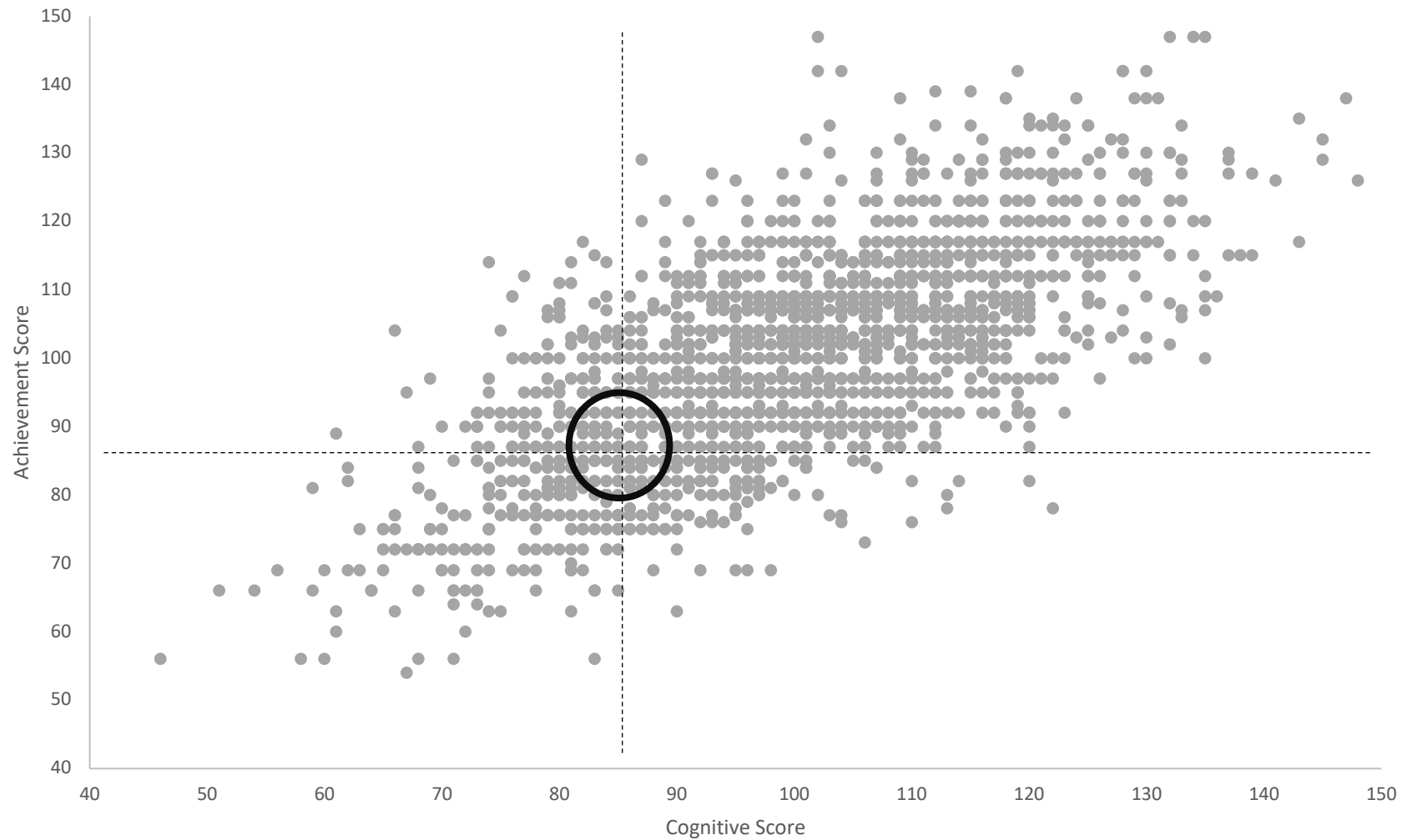


Figure 14.1. Scatter plot of simulated cognitive–achievement test scores illustrating the cutoff point logic employed in some PSW models. Decision-making error is likely to be exacerbated at the intersection of the thresholds due to measurement error in the scores. The lower left-hand quadrant contains individuals who may have LD, as they present with deficits in both cognitive ability and achievement. All other quadrants contain individuals who do not present with a confirmatory PSW pattern.

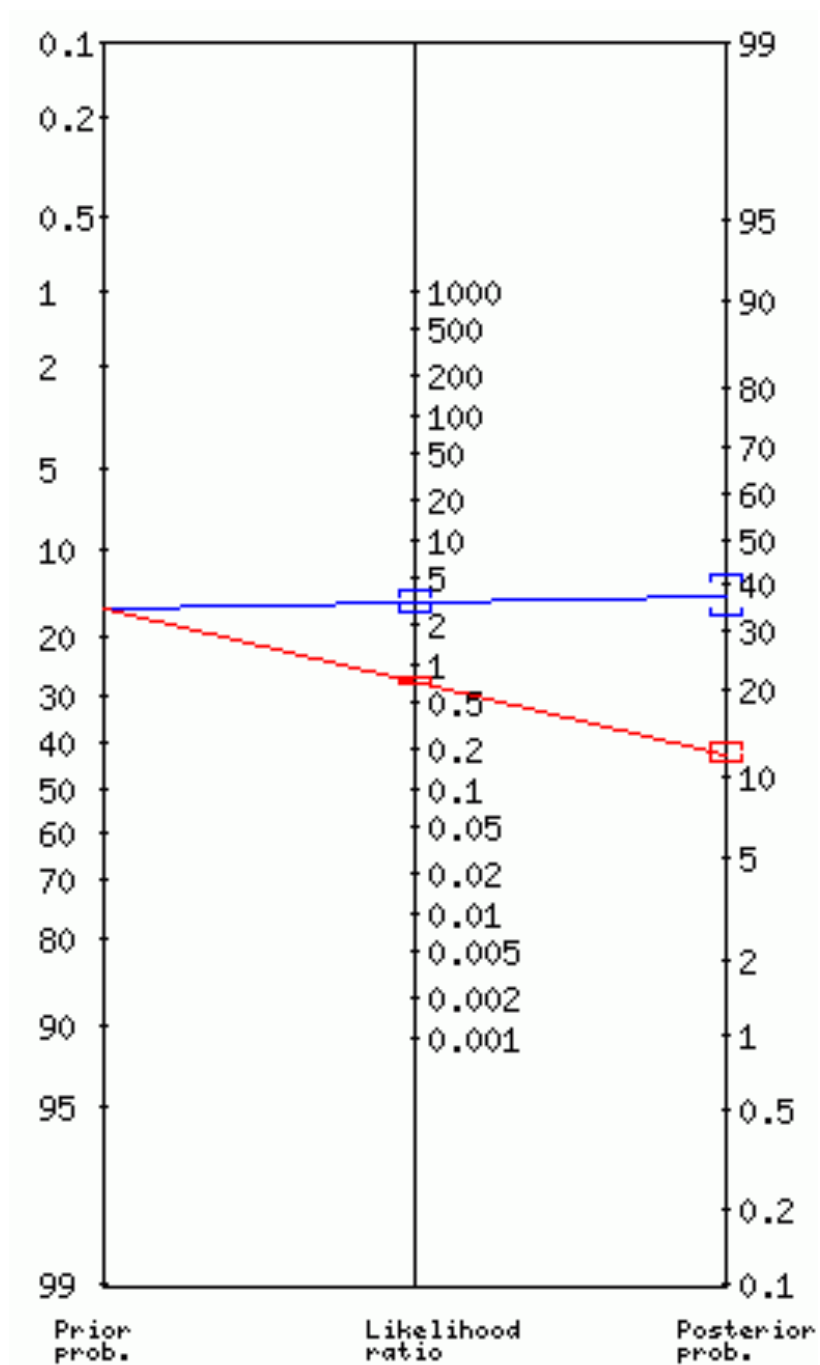


Figure 14.2. Probability nomogram used to combine prior probability (i.e., base rate [15%]) with the likelihood ratio (+3.13) to estimate revised posterior probability (37%). Using the clinical decision-making guidelines outline in Youngstrom and Van Meter (2013), additional assessment is needed to provide focused targeted for LD.