

Please use the following citation when referencing this work:

McGill, R. J., Styck, K. M., 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. doi: 10.1177/0731948715618504

Critical Issues in Specific Learning Disability Identification: What We Need to Know

About the PSW Model

Ryan J. McGill

Texas Woman's University

Kara M. Styck

University of Texas at San Antonio

Ronald S. Palomares

Texas Woman's University

Michael R. Hass

Chapman University

Author notes

Ryan J. McGill, Department of Psychology and Philosophy, Texas Woman's University, P.O. Box 425470, Denton, TX 76204.

Kara M. Styck, Department of Educational Psychology, University of Texas at San Antonio, 501 W. Cesar Chavez Blvd., San Antonio, TX 78207.

Ronald S. Palomares, Department of Psychology and Philosophy, Texas Woman's University, P.O. Box 425470, Denton, TX 76204

Michael R. Hass, College of Educational Studies, Chapman University, One University Dr., Orange, CA 92668.

Correspondence concerning this article should be addressed to Ryan J. McGill, Department of Psychology and Philosophy, Texas Woman's University, P.O. Box 425470, Denton, TX 76204. E-Mail: rmcgill@twu.edu

Abstract

As a result of the upcoming Federal reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA), practitioners and researchers have begun vigorously debating what constitutes evidence-based assessment for the identification of specific learning disability (SLD). This debate has resulted in strong support for a method that appraises an individual's profile of cognitive test scores for the purposes of determining cognitive processing strengths and weaknesses, commonly referred to as patterns of strengths and weaknesses or PSW. Following the Fuchs and Deshler model (2007), questions regarding the psychometric and conceptual integrity of the PSW model are addressed. Despite the strong claims made by many PSW proponents, the findings by this review demonstrate the need for additional information in order to determine whether PSW is a viable alternative to existing eligibility models and worthy for large scale adoption for SLD identification. Implications for public policy and future SLD research are also discussed.

Keywords: specific learning disability, PSW, cognitive assessment, diagnostic validity

Critical Issues in Specific Learning Disability Identification: What We Need to Know

About the PSW Model

There has been a vigorous debate within the field of school psychology regarding the best procedures for determining eligibility for special education and related services under the category of specific learning disability (SLD) since the passage of the final regulations for the current iteration of the Individuals with Disabilities Education Improvement Act (IDEA) in 2004 (e.g., Hale, Naglieri, Kaufman, & Kavale, 2004; Kavale & Flanagan, 2007). Prior to IDEA (2004), federal regulations emphasized the primacy of the ability-achievement discrepancy model for the identification of SLD. Public Law 94-142 (i.e., the Education for All Handicapped Children Act of 1975) mandated SLD eligibility determination:

based on (1) whether a child does not achieve commensurate with his or her age and ability when provided with appropriate educational experiences, and (2) whether the child has a severe discrepancy between achievement and intellectual ability in one or more of seven areas relating to communication skills and mathematical abilities.

These concepts are to be interpreted on a case by case basis by the qualified evaluation team members. The team must decide that the discrepancy is not primarily the results of (1) visual, hearing, or motor handicaps; (2) mental retardation; (3) emotional disturbance; or (4) environmental, cultural, or economic disadvantage. (Federal Register, 1977, 42, p. 65082)

In contrast to previous legislation, IDEA (2004) permitted state education agencies (SEA) to select between the discrepancy method and several alternatives by specifying that state adopted SLD eligibility criteria “must *not* [emphasis added] require the use of a severe discrepancy between intellectual ability and achievement...” (IDEA, 34 C. F. R. §

300.307(a)(1)). Further stating, SEA “*must* [emphasis added] permit the use of a process based on the child’s response to scientific, research-based intervention” (IDEA, 34 C. F. R. § 300.307(a)(2)) and “*may* [emphasis added] permit the use of other alternative research-based procedures for determining whether a child has a specific learning disability” (IDEA, 34 C. F. R. § 300.307(a)(3)). The former alternative to the discrepancy method is a process commonly referred to as response-to-intervention (RTI), whereas the latter alternative opened the door for other research-based procedures to be used in SLD eligibility determinations.

The paradigm shift away from sole use of the discrepancy method was likely the result of a confluence of several factors spanning three decades (Aaron, 1997), including psychometric evidence of poor diagnostic convergence between different discrepancy formulas (e.g., Francis et al., 2005; Mellard, Deshler, & Barth, 2004; Reynolds, 1984) and research findings suggesting that individuals with discrepancies were able to benefit significantly from direct academic interventions (e.g., McMaster, Fuchs, Fuchs, & Compton, 2005; Vellutino, Scanlon, & Lyon, 2000). Although the discrepancy model has lost credibility, federal regulations codified in IDEA (2004) permitting alternative methods for determining SLD eligibility has provided momentum for continued debate over the role of cognitive testing in the identification of SLD.

Ahearn (2008) surveyed SEA representatives regarding SLD identification procedures after the enactment of IDEA (2004) and reported that six states required use of RTI while simultaneously prohibiting the discrepancy method for making SLD determinations (i.e., Colorado, Delaware, Georgia, Indiana, Iowa, and West Virginia), 10 states permitted use of all three methods described in federal regulations (i.e., Alabama, Arkansas, Florida, Kansas, Michigan, Nebraska, New Hampshire, Ohio, Oregon, and South Carolina), and 26 states allowed either the use of RTI or the discrepancy method for the identification of SLD (the Arizona SEA

representative declined to respond to the survey). Consequently, the discrepancy method and RTI may be the most widely used procedures for diagnosing SLD in school-based settings nationwide.

However, alternatives to the discrepancy model for SLD identification have also been proposed that emphasize the role of strengths and weaknesses in cognitive processing as measured by individually administered standardized intelligence tests, collectively referred to as the patterns of strengths and weaknesses approach (PSW; e.g., Hale, Flanagan, & Naglieri, 2008; Reynolds & Shaywitz, 2009). Johnson, Humphrey, Mellard, Woods, and Swanson (2010) reported moderate to large effect sizes in cognitive processing differences between students identified as SLD and typically achieving peers in a meta-analysis of 32 empirical studies. The authors concluded that measures of cognitive processing should be included in the evaluation and identification of SLD on the basis of their results. Moreover, a recent survey of 58 professional experts with experience in SLD assessment, regarding procedural best practices for diagnosing SLD, resolved that an “approach that identifies a pattern of psychological processing strengths and weaknesses, and achievement deficits consistent with this pattern of processing weaknesses, makes the most empirical and clinical sense” (Hale et al., 2010, p. 225).

Federal regulations require SLD diagnostic procedures to be “research-based.” Consequently, the PSW approach to cognitive test score interpretation must be accompanied by adequate evidence of reliability and validity (American Educational Research Association [AERA], American Psychological Association [APA], & The National Council on Measurement in Education [NCME], 2014). Given the nature and importance of the task(s) in which PSW is being prescribed, additional information is needed to determine the degree to which it is a viable alternative to existing eligibility models. The resultant purpose of the present review is to address

a series of questions regarding the psychometric and conceptual composition of the PSW model following the model of Fuchs and Deshler (2007):

1. How is a cognitive weakness defined?
2. Are PSW models diagnostically valid?
3. Are factor-based scores from cognitive tests suitable for individual decision-making?
4. Do school psychologists have adequate training to implement PSW with integrity?

How is a Cognitive Weakness Defined?

To date, several models have been proposed that operationalize PSW, including: (a) the concordance/discordance model (C/DM; Hale & Fiorello, 2004), (b) the Cattell-Horn-Carroll operational model (CHC; Flanagan, Alfonso, & Mascolo, 2011), and (c) the discrepancy/consistency model (D/CM; Naglieri, 2011). It is beyond the scope of the present review to provide a detailed account of the procedures involved for each PSW method. However, it is noteworthy that each of these PSW models shares at least three core assumptions as it relates to the diagnosis of SLD despite underlying differences in theoretical orientation: (a) evidence of cognitive weaknesses must be present, (b) an academic weakness must also be established, and (c) there must be evidence of spared cognitive-achievement abilities (Flanagan & Alfonso, 2015).

Problems Associated With Operationalizing a Weakness in PSW Models

Lack of a Uniform Method for Defining PSW. Establishing a cognitive processing weakness is the *sine qua non* of the PSW model. Yet, there is considerable disagreement regarding exactly how processing weaknesses should be operationalized (Flanagan, Ortiz,

Alfonso, & Dynda, 2006; Flanagan, Fiorello, Ortiz, 2010; Hale & Fiorello, 2004; Hale, Wycoff, & Fiorello 2011; Naglieri, 1999). A survey of contemporary SLD diagnostic research reveals that there is little consensus on exactly how PSW should be defined and that different diagnostic models of SLD identification are likely to identify different students (Lyon & Weiser, 2013). Some models (e.g., CHC) use a normative psychometric approach wherein an individual's cognitive-achievement scores are compared to the performance of a sample of same aged peers from the general population. In the CHC model, a *weakness* is defined as performance that falls below the average range (e.g., standard scores ranging from 90-110) whereas a *deficit* is defined as performance that falls greater than one standard deviation below the mean (e.g., standard scores at or below 85). Despite the weakness/deficit distinction, Flanagan and colleagues (2011) suggest that practitioners may utilize both thresholds for determining whether a particular score is indicative of an intra-individual weakness. Practitioners are also encouraged to utilize clinical judgment in order to determine whether strict adherence to the aforementioned thresholds is appropriate. To wit, "Some children who struggle academically may not demonstrate academic weaknesses or deficits on standardized, norm-referenced tests of achievement... Therefore, it is not important to assume that a child with a standard score of 90 in broad reading is okay" (Flanagan, Alfonso, & Mascolo, 2011, p. 245). However, it is important to highlight that probability distribution theory determines that the use of such a high cut-off (e.g., 90) is likely to result in approximately 25% of the population being identified as having a cognitive weakness on any given measure (Decker, Schneider, & Hale, 2012), which is larger than the estimated proportion of school-aged children and adolescents diagnosed with SLD nationwide (Kena et al., 2014).

In contrast, the C/DM model utilizes an intra-individual parametric approach to identify cognitive weaknesses that is denoted by statistically significant differences between a cognitive strength score and a cognitive weakness score using the standard error of the difference (SED) statistic. However, the SED formula is heavily mediated by the reliability of the constituent measures. Thus, if the reliability coefficients of each of two reference indices are .90 or higher, differences of only 3-5 standard score points are required for a cognitive weakness to be indicated. As a result, Hale and Fiorello (2004) encourage practitioners to seek convergent evidence from multiple data sources to make eligibility determinations. This advice is prescient given that the use of SED is likely to result in high initial false positive decisions, given that many composite and subtest scores have corresponding alpha levels that exceed .90.

Failure to Report PSW Base Rates. Flanagan and Alfonso (2015) have recently revised the CHC model in the form of the Dual Discrepancy/Consistency (DDC) model. The DDC model explicitly uses the Cross-Battery Assessment System (X-BASS; Flanagan, Ortiz, & Alfonso, 2015) proprietary software program for calculating cognitive strengths and weaknesses. According to Flanagan and Alfonso (2015), the X-BASS program uses elements of previously released software programs associated with the cross-battery assessment model (e.g., Flanagan, Ortiz, & Alfonso, 2013). The X-BASS program can be used to derive statistically significant ($p \leq .05$) critical values for cognitive weaknesses—users input their battery of cognitive-achievement scores and specify *a priori* which scores correspond to required elements of the DDC model. If the observed score from the specified cognitive weakness exceeds the critical value derived from the program algorithm, it is considered to demarcate a statistically significant cognitive weakness. The X-BASS also includes an additional computational element for synthesizing inputted scores into a pseudo general ability composite with a corresponding index that reports

the likelihood that an individual presents with otherwise spared cognitive abilities aside from the observed cognitive weakness. While the X-BASS software potentially provides users with a more statistically rigorous method for determining strengths and weaknesses associated with the DDC model, additional psychometric information (e.g., base rates in clinical and non-clinical populations) is needed to determine the diagnostic value of the scores provided by the program.

Poor Stability of Observed Difference Scores. The D/CM model utilizes a hybrid approach in which a cognitive weakness is assessed using a combination of intra-individual (i.e., ipsative analysis) and normative appraisals. Ipsative analyses (Davis, 1959) involve making intra-individual comparisons by subtracting each individual cognitive score from the mean of the profile of scores, the resulting difference score is then determined to be significant if it exceeds *a priori* statistical thresholds. In contrast to other PSW methods, ipsative analysis of cognitive test scores has been thoroughly investigated (e.g., Glutting, Watkins, & Youngstrom, 2003; Watkins, 2000). The resulting difference score is a transformation of the original score. Consequently, interpretation is difficult because the test score has been removed from its norm-referenced anchor point and the underlying psychometric properties of the transformed score are unknown (Glutting et al., 2003). Ipsative profiles have also consistently demonstrated poor stability and diagnostic accuracy at the subtest and composite levels (e.g., McDermott, Fantuzzo, & Glutting, 1990; McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992; McDermott & Glutting, 1997; Watkins, 2000). For example, Watkins (2000) investigated the degree to which Naglieri's (2000) relative weakness (i.e., ipsative assessment), cognitive weakness (i.e., at least one factor score below a cut-score and a relative weakness present), and cognitive and academic weakness (i.e., both relative and cognitive weakness present concurrently with a low normative academic achievement test score) measured by the Wechsler Intelligence Scale for Children-Third Edition

(WISC-III; Wechsler, 1997) and Cognitive Assessment System (CAS; Naglieri & Das, 1997) could accurately distinguish between students placed in special education programs and students placed in regular education programs. Results indicated only 24-51% of students in special education displayed a relative weakness, 17-32% of students in special education displayed a cognitive weakness, and 14-28% of students in special education displayed a cognitive and academic weakness suggesting these cognitive profiles are not commonly observed in students with disabilities. Furthermore, rates of false positive cases were similar indicating that relative weaknesses, cognitive weaknesses, and cognitive and academic weaknesses are also frequently observed in the cognitive profiles of students without disabilities (i.e., 27-42%, 10-13%, and 4-6% respectively). Yet, ipsative analysis at the composite score level continues to be recommended for the identification of SLD (e.g., Dehn, 2014; Naglieri, 2011) based upon the belief that the development of more advanced cognitive measures that are based upon theoretically derived models of intellectual abilities (e.g., CHC) allow users to overcome the aforementioned shortcomings of ipsative and similar types of intra-individual profile analysis (Flanagan, Ortiz, & Alfonso, 2013). Additional psychometric examinations of the potential diagnostic accuracy of ipsative profiles generated from recently revised cognitive measures (e.g., CAS2, WISC-V) would permit more relevant inferences as to the potential efficacy of the D/CM method.

Are PSW Models Diagnostically Valid?

There is scientific consensus that SLD is a neurological disorder with an etiology demarcated by specific neurocognitive deficits that impair an individual's ability to benefit from traditional academic instruction without significant psychoeducational remediation (Reynolds & French, 2005). Although proponents of PSW suggest that such methods provide a mechanism for

determining the specific cognitive deficits responsible for an individual's achievement difficulties as well as critical information for designing individualized treatment plans that take into consideration the dynamic nature of cognitive-achievement relationships (e.g., Hajovsky et al., 2014, Kaufman et al., 2012; McGrew & Wendling, 2010), this information is not sufficient for establishing the diagnostic utility of these procedures (McFall & Treat, 1999; Wiggins, 1973).

Determining the accuracy of PSW as a diagnostic sign requires the computation of sensitivity and specificity statistics from a 2 x 2 contingency table that cross-tabulates decisions made from PSW with those from a gold standard diagnosis (a sample diagnostic contingency array is provided in Table 1). Sensitivity is the proportion of individuals with SLD who exhibit significant PSW (i.e., true positive decisions). Specificity represents the proportion of individuals without SLD who are not marked with psychoeducational profiles that contain significant PSW (i.e., true negative decisions). Diagnostic accuracy is often represented as the rate of "hits" (i.e., true positive and true negative decisions) in a selected population. Steubing, Fletcher, Branum-Martin, and Francis (2012) investigated the diagnostic accuracy of several PSW models using simulated WISC-IV assessment data and reported high diagnostic specificity (i.e., true negative decisions) across all models with only 1-2% of the population meeting SLD criteria. However, this high specificity came at the cost of an inflated Type II error rate (i.e., false positive decisions) calling into question the ability of PSW methods for improving upon the diagnostic accuracy of existing models (e.g., discrepancy model, RTI) for correctly identifying children and adolescents *with* SLD. The authors concluded that the inflated error was likely the result of transforming a continuous scale of measurement into a binary model in which scores at or below a cut-point are considered to be indicative of a disorder. Problems associated with artificially

dichotomizing continuous data have long been known (Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002). However, it is important to note that multidisciplinary evaluation teams must make “yes/no” eligibility decisions according to IDEA (2004) and some degree of decision error will undoubtedly occur when cut-scores are used to diagnose SLD as a result. In consideration of this problem, it is important to highlight that all cognitive-achievement measures contain measurement error and thus true scores are bound to fluctuate around arbitrary cut-off points with repeated diagnostic testing (Macmann et al., 1989). According to Fletcher, Stuebing, Morris, & Lyon (2013), the imposition of a different classification model (e.g., PSW) may shift the focus on underlying casual variables but does nothing to address these more fundamental psychometric concerns that apply to all diagnostic models that utilize these data.

More explicit investigations of individual PSW models have only recently begun to emerge within the technical literature. C/DM procedures have been employed by researchers in a series of recent studies examining the potential diagnostic utility of the model for identifying SLD in a multitude of academic domains. Kubas et al. (2014) used C/DM diagnostic procedures with a referred sample of 283 children and adolescents from the United States and Canada. Results indicated that specific cognitive processing subtypes distinguished between those identified as having math SLD via the C/DM procedure and those that were not, with different cognitive skills mediating math performance across groups. Similar results have been replicated with respect to use of the C/DM model for diagnosing SLD in written expression (Fenwick et al., 2015) and for the CHC model in reading (Feifer et al., 2014). However, a recent investigation by Miciak and colleagues (2014) found that the utilization of different assessment measures in the C/DM model resulted in poor classification agreement ($k = .29$), suggesting that test selection in PSW is not arbitrary and may result in inconsistent diagnostic decisions across practitioners and

educational agencies. Moreover, Miciak, Fletcher, Stuebing, Vaughn, and Tolar (2014) administered multiple cognitive measures designed to measure CHC-related broad abilities to non-responders of a Tier 2 reading intervention program in order to test the reliability and validity of C/DM and CHC methods. Diagnostic overlap in which an individual met SLD criteria according to both PSW models fluctuated between 13.6% and 62.1% across different iterations of the models, indicating that model choice likely impacts diagnostic decision.

Absence of Aptitude-Treatment Interaction. As previously mentioned, in addition to its potential promise as an identification model, Hale et al. (2010) suggested “processing assessment could also lead to more effective individualized interventions for children who do not respond adequately to intensive interventions in an RTI approach,” (p. 229) implying the presence of an aptitude-treatment-interaction. This assumes that unique broad cognitive abilities differentially predict deficits in specific academic abilities. However, latent variable modeling studies have provided inconsistent evidence in support of this contention. Multiple investigations have reported that broad cognitive abilities have direct effects on reading and math achievement beyond the general factor (Benson, 2008; Floyd, Keith, Taub, & McGrew, 2007; Hajovsky et al., 2014; Taub, Keith, Floyd, & McGrew, 2008; Vanderwood, McGrew, & Flanagan, Keith, 2001), but the predictive effects associated with those abilities are relatively small (Beaujean, Parkin, & Parker, 2014; Glutting, Watkins, Konold, & McDermott, 2006; Oh et al., 2004; Parkin & Beaujean, 2012). Whereas there is emerging evidence (e.g., Compton et al., 2012; Fuchs, Hale, & Kearns, 2011) that individuals with SLD may present with discrepant cognitive profiles when compared to normal controls, these cognitive correlates rarely mediate academic intervention outcomes (Fletcher et al., 2011; Miciak et al., 2014; Stuebing et al., 2014).

Are Factor-Based Scores from Cognitive Tests Suitable for Individual Decision-Making?

The interpretation of subtest score profiles for diagnostic decision-making has long been advocated in the technical literature despite suggestions that PSW is a new and revolutionary approach to cognitive test interpretation and SLD identification (Flanagan, Ortiz, Alfonso, & Dynda, 2006; Hale, Wycoff, & Fiorello, 2011). Over 70 years ago, Rapaport et al. (1945) proposed an interpretive framework that provided clinicians with a step-by-step process for analyzing intra-individual cognitive strengths and weaknesses based upon the belief that individual variations in cognitive test performance served as evidence for the presence of a variety of clinical disorders and a multitude of related approaches have been subsequently developed (e.g., Kaufman, 1994; Naglieri, 2000; Priftera & Dersh, 1993).

However, Gnys, Willis, and Faust (1995) characterized the belief that subtest scatter distinguishes individuals with SLD from those without disabilities as an *illusory correlation*—the false belief that two variables are related (Chapman & Chapman, 1967) due to the preponderance of evidence indicating the poor ability of subtest scatter to accurately discriminate between clinical and non-clinical subgroups (e.g., Kavale & Forness, 1984; Macmann & Barnett, 1997; Watkins, 2000). Nevertheless, Pfeiffer, Reddy, Kletzel, Schmelzer, and Boyer (2000) surveyed 354 nationally certified school psychologists regarding their use and perceptions of profile analysis for the diagnosis of SLD and reported approximately 70% of respondents believed the information obtained from profile analysis was clinically meaningful and 89% of respondents declared that they used profile analysis for making diagnostic decisions. Yet, Fletcher et al. (2013) argued, “It is ironic that methods of this sort [PSW] continue to be proposed when the basic psychometric issues are well understood and have been documented for many years” (p.40). Hale and colleagues (2010) counter this critique by arguing that previously documented shortcomings of profile analysis do not hold for contemporary PSW models as a

result of the development of more sophisticated measurement instruments and advances in cognitive and neuropsychological theory (e.g., McGrew, 2009). While associated genetic and neuro imaging studies are important, they are not instructive for evaluating the clinical utility of the PSW models.

Poor Reliability of Latent Factor Scores

PSW models encourage clinician's to interpret factor-based scores (e.g., broad ability indexes and composites) as primarily indicating orthogonal broad cognitive abilities. However, investigators have reported that many measures of specific abilities contain large proportions of *g* variance and relatively little unique variance that can be attributed to the abilities purported to be estimated by those measures (e.g., Canivez, 2011; Canivez & Watkins, 2010; Dombrowski & Watkins, 2013; Styck & Watkins, in press; Watkins, 2006). Given the multidimensional nature of intelligence, additional research is needed to provide clinicians with a procedure for disentangling the many sources of construct irrelevant variance that contaminate measures of broad abilities (Canivez, in press; Reise, Bonifay, & Haviland, 2013). In the absence of such a procedure, the generation of reliable and valid inferences from cognitive profile data can at best be presently described as aspirational (Watkins, 2000). According to Beaujean, Parkin, and Parker (2014), multidimensionality is not the problem per se, the problem occurs when interpretations of individual cognitive abilities and their related composites "fails to recognize that Stratum II factors derived from higher order models are not totally independent of *g*'s influence" (p. 800). As a result, it may be more useful to re-conceptualize broad cognitive abilities as different "flavors" of *g* (Carroll, 1993) in contrast to more discrete indicators of unique abilities.

Not surprisingly, complimentary studies assessing the incremental predictive effects of cognitive test scores have consistently found that observed factor-based scores rarely account for meaningful proportions of achievement scores after controlling for the strong predictive effects of the general intelligence score (e.g., Canivez, 2013; Glutting et al., 2006; McGill, 2015; McGill & Busse, 2015). Frazier and Youngstrom (2007) suggest that the inability of factor-based broad ability measures to consistently provide incremental predictive effects beyond the full scale IQ (FSIQ) score may be due to a variety of reasons including but not limited to: a) poorly defined measures; b) measures that lack adequate measurement specificity; and c) potential overfactoring of cognitive test batteries. For example, independent factor analytic investigations using more conservative procedures for factor extraction and variance partitioning (e.g., Canivez, 2008; Dombrowski, 2013; Strickland, Watkins, & Caterino, 2015) have generally failed to replicate the reported structure in many contemporary test manuals. This is not to suggest that there are not unique cognitive abilities apart from general intelligence that may be useful for practitioners to examine when determining whether an individual has SLD. However, practitioners must be mindful of the limitations of working with observed level data when interpreting such information (Schneider, 2013).

Lack of Information Regarding Long-Term Stability of Observed Factor Scores

Another potential threat to valid factor-based score interpretation, as advocated in PSW models, involves the stability of these scores over test-retest intervals. Although the results of test-retest studies are often reported in technical and interpretive manuals for cognitive measures, the results of these studies often do not provide users with information as to the long-term stability of cognitive test scores. For example, in the technical manual for the Kaufman Assessment Battery for Children-Second Edition (KABC-II; Kaufman & Kaufman, 2004),

adjusted average stability coefficients for the CHC-related cognitive composites ranged from .76 (Visual Processing) to .95 (Crystallized Ability) for ages 7-18. Although these coefficients generally meet established standards for evidence of temporal stability (e.g., Wasserman & Bracken, 2013), it is important to highlight that these coefficients were obtained over a relatively short re-test interval (12-56 days) with a relatively small sample of participants ($n = 205$). Although short-term stability is important, demonstrating the long-term stability (e.g., 1-3 years) of cognitive scores is critical as these scores provide the basis for potential long-term placements in special education and related intervention programs (Cronbach & Snow, 1977).

Recently, Watkins and Smith (2013) reported that 29-44% of WISC-IV composite scores (i.e., Verbal Comprehension Index, Perceptual Reasoning Index, Processing Speed Index, and Working Memory Index) for a sample of 344 students referred for special education evaluations demonstrated differences ≥ 10 standard score points across a mean test-retest interval of 2.84 years. These results suggest that an individual's pattern of cognitive scores is not stable over time. Future research is needed to replicate this work with other measures of cognitive ability. It also is important to note that scores from intelligence tests are mediated by error that is unique to each testing situation and is not stable over time (see Figure 1). Additionally, these effects may be exacerbated when different examiners assess the same individual across re-examination periods (see McDermott, Watkins, & Rhoad, 2014). Many PSW models (e.g., C/DM) encourage practitioners to engage in ongoing data-based decision making over time in order to protect against the well-known psychometric deficiencies of point in time profile analysis. However, more research is needed to determine if the individual strengths and weakness generated from PSW assessment are sufficiently stable to permit the diagnostic inferences as required in current federal and state regulations.

Do School Psychologists have the Training Necessary to Implement a PSW Model with Integrity?

In contrast to actuarial methods (i.e., the discrepancy model), PSW emphasizes the clinical judgment of the assessor in searching for patterns of strengths and weaknesses and identified areas of low achievement (Schultz, Simpson, & Lynch, 2012). According to Flanagan et al. (2012) “SLD identification is complex and requires a great deal of empirical and clinical knowledge on the part of practitioners” (p. 666). Thus, school psychologists must have advanced training and preparation in the theory of cognitive abilities, causal cognitive-achievement relationships, and advanced psychometrics and test interpretation in order for PSW to be implemented with fidelity in educational settings (Fiorello, Hale, & Wycoff, 2012; Weiner, 1989). However, Decker, Hale, and Flanagan (2013) argued that school psychology training programs do not adhere to evidence-based assessment practices due to: (a) an overemphasis of interpreting the FSIQ score; b) a gradual decline in the breadth and scope of the cognitive assessment training sequence as a result of RTI implementation, and (c) poor training in linking cognitive test results to evidence-based interventions. As a result, Decker et al. conclude that the cognitive assessment training sequence in many school psychology programs must be overhauled if the promise of cognitive testing is to be realized with contemporary measures, regardless of the method being used for SLD identification.

These criticisms are especially prescient given the multitude of diagnostic errors that are possible when engaging in interpretive procedures with multiple sources of data that place a premium on clinical judgment (Dawes, Faust, & Meehl, 1989; Garb, 2005; Watkins, 2009), notwithstanding standardized administration and scoring errors commonly identified on intelligence test protocols (Styck & Walsh, in press). According to Gambrill (2005), clinical

predictions often overlook a number of confounding causes, such as “the play of chance and misleading effects of small biased samples” (p. 453). A consistent theme across all of the PSW models examined in the present review was the need for users to integrate multiple sources of data in order to enhance the ecological and treatment validity of diagnostic decisions.

Accordingly, Fiorello and colleagues (2012) suggest that “CHT [C/DM] avoids many of the difficulties of this process [profile analysis] by confirming or disconfirming hypotheses with further data collection, including further testing of psychological processes beyond a single cognitive test” (p. 487). Prescriptive statements such as these in education are rarely justified and require adherence to high standards of empirical evidence (Marley & Levin, 2011). It is not clear how such default statements provide users with the appropriate guidance for avoiding errors in clinical judgment given that strengths and weaknesses are endemic in the population. As stated by Flanagan et al. (2011), “Most individuals have statistically significant strengths and weaknesses in their cognitive ability and processing profiles... Therefore, statistically significant variation in cognitive and neuropsychological functioning in and of itself must not be used as *de facto* evidence of SLD” (p. 242). While we agree with Flanagan and colleagues that ridged adherence to any one specific interpretive heuristic is likely to lead to method bias (Cook & Campbell, 1979), additional research is needed to determine how users are to effectively integrate multiple PSW data sources in a manner that facilitates improved diagnostic decision-making.

Conclusion and Future Directions

Similar to Dombrowski and Gischlar (2014), we argue that it is beneficial and important to evaluate proposed models of SLD identification in relationship to established ethical codes (American Psychological Association, 2004; National Association of School Psychologists,

2010) psychometric standards (e.g., AERA, APA, & NCME, 2014), and relevant educational laws (i.e., IDEA, 2004). Fiorello, Flanagan, and Hale (2014) claim that PSW: a) is the only empirically supported SLD identification model that comports with the statutory definition of SLD, b) when used in concert with an RTI-based pre-referral intervention system, is more likely to result in correct identification of SLD, and c) provides users with information relevant for developing individualized interventions. Although evidence for the efficacy of the PSW model is presently accumulating, we believe that wholesale endorsement of these claims at the present time is premature.

The purpose of this review was to raise a series of psychometric and conceptual questions that have yet to be addressed within the empirical literature regarding the PSW model. This information is vital to establishing evidence-based procedures for SLD treatment and assessment given the futility of previous attempts at validating ATI (Fletcher et al., 2013). Most of the research cited by PSW proponents has investigated relationships between specific cognitive variables and achievement (e.g., McGrew & Wendling, 2010) or the degree to which SLD subtypes can be identified using these procedures (e.g., Carmichael, Fraccaro, Miller, & Maricle, 2014; Fiorello, Hale, & Snyder, 2006). As stated by Miciak et al. (2014), “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). Nor does it obviate the fundamental measurement issues, common to all diagnostic models, which complicate SLD classification (Macmann et al., 1989). In sum, at this early stage in research on the PSW model, additional empirical evidence is needed to determine the degree to which PSW is a more robust and valid alternative to existing procedures, or simply a better mousetrap.

References

- Aaron, P. G. (1997). The impending demise of the discrepancy formula. *Review of Educational Research, 67*, 461-502. doi: 10.3102/00346543067004461
- Ahearn, E. M. (2008, July). State eligibility requirements for specific learning disabilities. *inForum* (Brief Policy Analysis). Retrieved from <http://www.projectlorum.org/docs/StateEligibilityRequirementsforSpecificLearningDisabilities.pdf>
- American Educational Research Association, American Psychological Association, & National Council on Measurement on Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychological Association. (2004). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060-1073.
- Beaujean, A. A., Parkin, J., & Parker, S. (2014). Comparing Cattell-Horn-Carroll factor models: Differences between bifactor and higher order factor models in predicting language achievement. *Psychological Assessment, 26*, 789-805. doi: 10.1037/a0036745
- Benson, N. (2008). Cattell-Horn-Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment, 26*, 27-41. doi: 10.1177/073428290701424
- Canivez, G. L. (2008). Orthogonal higher-order factor structure of the Stanford-Binet Intelligence Scales for children and adolescents. *School Psychology Quarterly, 23*, 533-541. doi: 10.1037/a0012884
- Canivez, G. L. (2011). Hierarchical factor structure of the Cognitive Assessment System: Variance partitions from the Schmid-Leiman (1957) procedure. *School Psychology Quarterly, 26*, 305-317. doi: 10.1037/a0025973

- Canivez, G. L. (2013). Incremental validity of WAIS-IV factor index scores: Relationships with WIAT-II and WIAT-III subtest and composite scores. *Psychological Assessment, 25*, 484-495. doi: 10.1037/a0032092
- Canivez, G. L. (in press). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements*. Gottingen, Germany: Hogrefe Publishers.
- Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV): Exploratory and higher-order factor analyses. *Psychological Assessment, 22*, 827-836. doi: 10.1037/a0020429
- Carmichael, J. A., Fraccaro, R. L., Miller, D.C., & Maricle, D. E. (2014). Academic achievement and memory differences among specific learning disabilities subtypes. *Learning Disabilities: A Multidisciplinary Journal, 20*, 8-17. Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5150>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72*, 193-204.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249-253. doi: 10.1177/014662168300700301
- Compton, D. L., Fuchs, L. S., Fuchs, D., Lambert, W., & Hamlett, C. (2012). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities, 45*, 79-85. doi: 10.1177/002221941039012

- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design analysis issues for field settings*. Boston: Houghton & Mifflin.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Davis, F. B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology, 50*, 162-170. doi: 10.1037/h0044024
- Dawes, R. M., Faust D., & Meehl, P. E. (1989). Clinical versus actuarial judgement. *Science, 243*, 1668-1674. doi: 10.1126/science.2648573
- 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. doi: 10.1002/pits.21675
- Decker, S. L., Schneider, W. J., & Hale, J. B. (2012). Estimating base rates of impairment in neuropsychological test batteries: A comparison of quantitative models. *Archives of Clinical Neuropsychology, 27*, 69-84. Doi: 10.1093/arclin/acr088
- Dehn, M. J. (2014). *Essentials of processing assessment* (2nd ed.). Hoboken, NJ: John Wiley.
- Dombrowski, S. C. (2013). Investigating the structure of the WJ-III Cognitive at school age. *School Psychology Quarterly, 28*, 154-169. doi: 10.1037/spq0000010
- 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. doi: 10.1080/15377903.2013.869786
- Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher-order factor analysis of the WJ-III full-test battery: A school-aged analysis. *Psychological assessment, 25*, 442-455. doi: 10.1037/a0031335

- 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. Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5151>
- Fenwick, M. E., Kubas, H. A., Witzke, J. W., Fitzer, K. R., Miller, D. C., Maricle, D. E., Harrison, G. L., Macoun, S. J., & Hale, J. B. (2015). Neuropsychological profiles of written expression learning disabilities determined by Concordance-Discordance Model criteria. *Applied Neuropsychology: Child*. Advance online publication. doi: 10.1080/21622965.2014.993396
- Fiorello, C. A., Hale, J. B., & Snyder, L. E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools*, 43, 835-853. doi: 10.1002/pits.20192
- Fiorello, C. A., Flanagan, D. P., & Hale, J. B. (2014). Response to the special issue: The utility of the pattern of strengths and weaknesses approach. *Learning Disabilities: A Multidisciplinary Journal*, 20, 87-91. Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5154>
- Fiorello, C. A., Hale, J. B., & Wycoff, K. L. (2012). Cognitive hypothesis testing. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 484-496). New York: Guilford.
- Flanagan, D. P., & Alfonso, V. C. (2015, February). *The SLD pattern of strengths and weaknesses and the WJ-IV*. Mini-skills workshop presented at the annual meeting of the National Association of School Psychologists, Orlando, FL.
- Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2011). A CHC-based operational definition of

- SLD: Integrating multiple data sources and multiple data-gathering methods. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 233-298). Hoboken, NJ: John 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. doi: 10.1002/pits.20501
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2015). *Cross-Battery Assessment Software System (X-BASS)*. Hoboken, NJ: John Wiley.
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). Hoboken, NJ: John Wiley.
- Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. (2006). Integration of response-to-intervention and norm-referenced tests in learning disability identification: Learning from the Tower of Babel. *Psychology in the Schools, 43*, 807-825. doi: 10.1002/pits.20190
- Fletcher, J. M., Stuebing, K. K., Barth, A. E., Denton, C. A., Cirino, P. T., . . . Vaughn, S. (2011). Cognitive correlates of inadequate response to reading intervention. *School Psychology Review, 40*, 3-22. Retrieved from <http://www.nasponline.org>
- Fletcher, J. M., Steubing, 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). New York: Guilford Press.
- Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll cognitive abilities and their effects on reading decoding skills: *g* has indirect effects, more specific abilities have direct effects. *School Psychology Quarterly, 22*, 200-233. doi:

10.1037/1045-3830.22.2.200

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. doi:

10.1177/00222194050380020101

Frazier, T. W., & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence, 35*, 169-182. doi: 10.1016/j.intell.2006.07.002

Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 22*, 129-136. doi: 10.1111/j.1540-5826.2007.00237.x.

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. doi:

10.1177/0022219411400019

Gambrill, E. (2005). *Critical thinking in clinical practice: Improving the quality of judgements and decisions* (2nd ed.). Hoboken, NJ: John Wiley.

Garb, H. N. (2005). Clinical judgement and decision making. *Annual Review of Clinical*

Psychology, 1, 67-89. doi: 10.1146/annurev.clinpsy.1.102803.143810

Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *Journal of Special Education, 40*, 103-114. doi: 10.1177/00224669060400020101

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). New York: Guilford.
- Gnys, J. A., Willis, W. G., & Faust, D. (1995). School psychologists' diagnoses of learning disabilities: A study of illusory correlation. *Journal of School Psychology, 33*, 59–73.
- Hajovsky, D., Reynolds, M. R., Floyd, R. G., Turek, J. J., & Keith, T. Z. (2014). A multigroup investigation of latent cognitive abilities and reading achievement relations. *School Psychology Review, 43*, 385-406. Retrieved from <http://www.nasponline.org>
- 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. doi: 10.1177/073194871003300310
- Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York: Guilford.
- Hale, J. B., Flanagan, D. P., & Naglieri, J. A. (2008). Alternative research-based methods for IDEA (2004) identification of children with specific learning disabilities. *Communiqué, 36*(8), 14-15. Retrieved from <http://www.nasponline.org>
- Hale, J. B., Naglieri, J. A., Kaufman, A. S., & Kavale, K. A. (2004). Specific learning disability classification in the new Individuals with Disabilities Education Act: The danger of good ideas. *The School Psychologist, Winter*, 6-13.
- 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*

- identification (pp. 173-201). Hoboken, NJ: John Wiley.
- Individuals with Disabilities Education Improvement Act. Public Law 108-446, 108th Cong., 118 Stat. 2647 (2004) (enacted).
- Individuals with Disabilities Education Improvement Act Final Regulations. 34 C.F.R. § 300.1 *et seq.*
- Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disabilities Quarterly*, 33, 3-18. doi: 10.1177/073194871003300101
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: John Wiley.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children-Second Edition*. Circle Pines, MN: American Guidance Service.
- Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence*, 40, 123-138. doi: 10.1016/j.intell.2012.01.009
- Kavale, K. A., & Flanagan, D. P. (2007). Ability-achievement discrepancy, response to intervention, and assessment of cognitive abilities/processes in specific learning disability identification: Toward a contemporary operational definition. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 130-147). New York: Springer.
- Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler Scale profiles and recategorizations: Patterns of parodies? *Learning Disabilities Quarterly*, 7, 136-156. doi: 10.2307/1510314

- Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., & Kristapovich, P. (2014). *The Condition of Education 2014* (NCES 2014-083). U.S. Department of Education, National Center for Education Statistics. Washington, DC.
- Kubas, H. A., Drefs, M. A., Poole, J. M., Schmid, A. D., Holland, S. C., & Fiorello, C. A. (2014). Cognitive and academic profiles associated with math disability subtypes. *Learning Disabilities: A Multidisciplinary Journal*, *20*, 31-44. Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5152>
- Lyon, G. R., & Weiser, B. (2013). The state of science in learning disabilities: Research impact on the field from 2001 to 2011. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of the learning disabilities* (2nd ed., pp. 118-151). New York: 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. Doi: 10.1037//1082-989X.7.1.19
- 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. doi: 10.1037/h0088959
- Macmann, G. M., Barnett, D. W., Lombard, T. J., Belton-Kocher, E., & Sharpe, M. N. (1989). On the actuarial classification of children: Fundamental studies of classification agreement. *Journal of Special Education*, *23*, 127-149. doi: 10.1177/002246698902300202
- Marley, S. C., & Levin, J. R. (2011). When are prescriptive statements in educational research justified? *Educational Psychology Review*, *23*, 197-206. doi: 10.1007/s10648-011-9154-y
- 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. doi: 10.1177/073428299000800307
- 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. *The Journal of Special Education*, 25, 504-526. doi: 10.1177/002246699202500407
- McDermott, P. A., & Glutting, J. J. (1997). Informing stylistic learning behavior, disposition, and achievement through ability subtests—Or, more illusions of meaning? *School Psychology Review*, 26, 163-176. Retrieved from <http://www.nasponline.org>
- McDermott, P. A., Watkins, M. W., & Rhoad, A. (2014). Whose IQ is it?—Assessor bias variance in high-stakes psychological assessments. *Psychological Assessment*, 26, 207-214. Doi: 10.1037/a0034832
- McFall, R. M., & Treat, T. A. (1999). Quantifying information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, 50, 215-241. doi: 10.1146/annurev.psych.50.1.215
- McGill, R. J. (2015). Incremental criterion validity of the WJ-III COG clinical clusters: Marginal predictive effects beyond the general factor. *Canadian Journal of School Psychology*, 30, 51-63. doi: 10.1177/0829573514560529
- McGill, R. J., & Busse, R. T. (2015). Incremental validity of the WJ III COG: Limited predictive effects beyond the GIA-E. *School Psychology Quarterly*, 30, 353-365. doi: 10.1037/spq0000094
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1-10. doi: 10.1016/j.intell.2008.08.004

- McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the last 20 years of research. *Psychology in the Schools, 47*, 651-675. doi: 10.1002/pits.20497
- McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*, 445-463. doi: 10.1177/001440290507100404
- Mellard, D. F., Deshler, D. D., & Barth, A. (2004). LD identification: It's not simply a matter of building a better mousetrap. *Learning Disability Quarterly, 27*, 230-242. doi: 10.2307/1593675
- 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. Doi: 10.1037/spq0000037
- 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 weaknesses framework. *School Psychology Quarterly, 30*, 321-334. doi: 10.1037/spq0000091
- Miciak, J., Stuebing, K. K., Vaughn, S., Roberts, G., Barth, A. E., & Fletcher, J. M. (2014). Cognitive attributes of adequate and inadequate responders to reading intervention in middle school. *School Psychology Review, 43*, 407-427. Retrieved from <http://www.nasponline.org>
- Naglieri, J. A. (1999). *Essentials of CAS assessment*. New York: John Wiley.
- Naglieri, J. A. (2000). Can profile analysis of ability test scores work? An illustration using the

- PASS theory and CAS with an unselected cohort. *School Psychology Quarterly*, *15*, 419-433. doi: 10.1037/h0088798
- 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). Hoboken, NJ: John Wiley.
- Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System*. Chicago, IL: Riverside.
- Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive Assessment System* (2nd ed.). Austin, TX: Pro-Ed.
- National Association of School Psychologists. (2010). National Association of School Psychologists principles for professional ethics. *School Psychology Review*, *39*, 302-319.
- Oh, H., Glutting, J. J., Watkins, M. W., Youngstrom, E. A., & McDermott, P. A. (2004). Correct interpretation of latent versus observed abilities: Implications from structural equation modeling applied to the WISC-III and WIAT linking sample. *Journal of Special Education*, *38*, 159-173. doi: 10.1177/00224669040380030301
- Parkin, J. R., & Beaujean, A. A. (2012). The effects of Wechsler Intelligence Scale for Children—Fourth Edition cognitive abilities on math achievement. *Journal of School Psychology*, *50*, 113-128. doi: 10.1016/j.jsp.2011.08.003
- Pfeiffer, S. I., Reddy, L. A., Kletzel, J. E., Schmelzer, E. R., & Boyer, L. M. (2000). The practitioner's view of IQ testing and profile analysis. *School Psychology Quarterly*, *15*, 376-385. doi: 10.1037/h0088795
- Prifitera, A., & Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning-disabled, and ADHD samples. *Journal of Psychoeducational Assessment Monograph Series*, *43-55*.

- Rapaport, D., Gil, M., & Schafer, R. (1945). *Diagnostic psychological testing: The theory, statistical evaluation, and diagnostic application of a battery of tests* (Vol. 1). Chicago: Yearbook Medical.
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95*, 129-140. doi: 10.1080/00223891.2012.725437
- Reynolds, C. R. (1984). Critical measurement issues in learning disabilities. *Journal of Special Education, 18*, 451-476. doi: 10.1177/002246698401800403
- Reynolds, C. R., & French, C. L. (2005). The brain as a dynamic organ of information processing and learning. In R. C. D'Amato, E. Fletcher-Janzen, & C. R. Reynolds (Eds.), *Handbook of school neuropsychology* (pp. 86-119). Hoboken, NJ: John Wiley.
- Reynolds, C. R., & Shaywitz, S. A. (2009). Response to intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 34*, 130-145. doi/10.1037/a0016158
- Schneider, W. J. (2013). What if we took our models seriously? Estimating latent scores in individuals. *Journal of Psychoeducational Assessment, 31*, 186-201. doi: 10.1177/0734282913478046
- Schultz, E. K., Simpson, C. G., & Lynch, S. (2012). Specific learning disability identification: What constitutes a pattern of strengths and weaknesses? *Learning Disabilities: A Multidisciplinary Journal, 18*, 87-97. Retrieved from <http://www.ldaamerica.org>
- Stuebing, K. S., Barth, A. E., Trahan, L. H., Reddy, R. R., Miciak, J., & Fletcher, J. M. (2014). Are child cognitive characteristics strong predictors of response to intervention? A meta-analysis. *Review of Educational Research*, Advance online publication. doi:

10.3102/0034654314555996

Steubing, 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. Retrieved from <http://www.nasponline.org>

Strickland, T., Watkins, M. W., & Caterino, L. C. (2015). Structure of the Woodcock Johnson III Cognitive Tests in a referral sample of elementary school students. *Psychological Assessment, 27*, 689-697. doi: 10.1037/pas0000052

Styck, K. M., & Walsh, S. M. (in press). Evaluating the prevalence and impact of examiner errors on the Wechsler scales of intelligence: A meta-analysis. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000157

Styck, K. M., & Watkins, M. W. (in press). Structural validity of the WISC-IV for students with learning disabilities. *Journal of Learning Disabilities*. Advance online publication. doi: 10.1177/0022219414539565

Taub, G. E., Keith, T. Z., Floyd, R. G., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly, 23*, 187-198. doi: 10.1037/1045-3830.23.2187

United States Office of Education. (1977). Rules and regulations implementing Public Law 94-142 (The Education for All Handicapped Children Act). *Federal Register, 42*.

Vanderwood, M. L., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2001). The contribution of general and specific cognitive abilities to reading achievement. *Learning and Individual Differences, 13*, 159-188. doi: 10.1016/s1041-6080(02)00077-8

Vellutino, F. Scanlon, D., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate

- and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*, 33, 223-238. doi: 10.1177/002221940003300302
- Wasserman, J. D., & Bracken, B. A. (2013). Fundamental psychometric considerations in assessment. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology: Assessment psychology* (2nd ed., Vol. 10, pp. 50-81). Hoboken, NJ: John Wiley.
- Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly*, 15, 465-479. doi: 10.1037/h0088802
- Watkins, M. W. (2006). Orthogonal higher-order structure of the WISC-IV. *Psychological Assessment*, 18, 123-125. doi: 10.1037/1040-3590.18.1.123
- Watkins, M. W. (2009). Errors in diagnostic decision making and clinical judgment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 210-229). Hoboken, NJ: John Wiley.
- Watkins, M. W., & Smith, L. G. (2013). Long-term stability of the Wechsler Intelligence Scale for Children-Fourth Edition. *Psychological Assessment*, 25, 477-483. doi: 10.1037/a0031653
- Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment*, 53, 827-831. doi: 10.1207/s15327752jpa5304_18
- Wechsler, D. X. (1997). *Wechsler Intelligence Test for Children—Third Edition*. San Antonio, TX: Psychological Corporation.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

Table 1

Sample PSW Diagnostic Contingency Table

Outcome of the PSW Model	Condition (e.g., known SLD diagnoses)	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	True Positive	False Positive
<i>Negative</i>	False Negative	True Negative

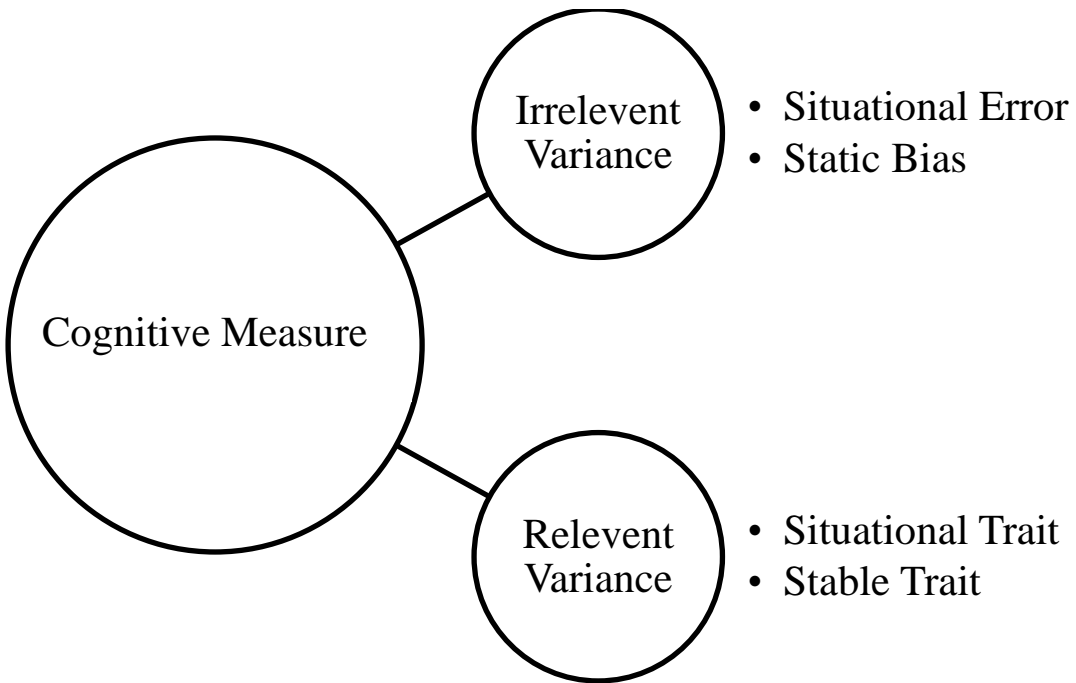


Figure 1. Sources of Influence on an Individual's Performance on a Cognitive Measure.