



Core-Selective Evaluation Process Handbook™

Updated and Revised 2024



Tammy L. Stephens & Edward K. Schultz

Page Intentionally
Left Blank

Core-Selective Evaluation Process (C-SEP™)



Overview and Procedures Updated and Revised

Tammy L. Stephens, Ph.D.
Edward K. Schultz, Ph.D.

Page Intentionally
Left Blank

Table of Contents

TABLE OF CONTENTS	3
ABBREVIATIONS	7
FORWARD	11
ORIGINS OF C-SEP™	15
DEFINITION	15
CONTEXTUAL BACKGROUND	18
SLD IDENTIFICATION.....	24
DEFINITION OF PATTERNS OF STRENGTHS AND WEAKNESSES.....	26
CORE-SELECTIVE EVALUATION PROCESS™	29
A LEGALLY DEFENSIBLE PSW MODEL.....	29
LANGUAGE.....	32
THE FOUR PS.....	37
<i>Publisher Guidance</i>	37
<i>Policy</i>	39
<i>Professional Judgment</i>	40
<i>Practices</i>	43
CRITICAL STEPS OF C-SEP™	45
REVIEW	47
RESPONSE-TO-INTERVENTION (RTI) AND C-SEP™	50
PRELIMINARY ASSESSMENT OF EXCLUSIONARY FACTORS.....	51
MULTIPLE SOURCES OF DATA (MSD)	52
INFORMATION OBTAINED PRIOR TO TESTING.....	54
PLAN	57
A FOCUSED/PURPOSEFUL EVALUATION	58
TESTING HYPOTHESIS & FOCUSED REFERRAL QUESTION (FRQ)	59

Table of Contents

CHOOSING CORE AND SELECTIVE TESTS	61
LIMITATIONS OF INDIVIDUAL NORM-REFERENCED TESTS	62
ASSESS	65
DATA OBTAINED DURING FORMAL TESTING	66
DATA OBTAINED FROM THE SCORE REPORT.....	67
<i>Standard Scores & Percentile Ranks</i>	<i>67</i>
<i>Criterion-Based Interpretations Relative Proficiency Index (RPI)</i>	<i>68</i>
<i>Review of Base Rate and Intra-Individual Performance</i>	<i>69</i>
<i>Analysis of Task Demands</i>	<i>70</i>
INTELLECTUAL DEVELOPMENT	73
NORM-REFERENCED TESTS (NRTs).....	75
DECIDE	83
MANAGEMENT AND ANALYSIS OF DATA TO IDENTIFY PSWs	84
<i>Pattern Seeking Techniques.....</i>	<i>84</i>
<i>Assessing Individual Data Sources.....</i>	<i>86</i>
<i>Inter-Normative & Intra-Relative Measures.....</i>	<i>88</i>
<i>Final Rule-Out of Exclusionary Factors</i>	<i>90</i>
<i>Merging the Available Data.....</i>	<i>90</i>
<i>Systematic Data Analysis Using the Chain of Evidence Technique.....</i>	<i>91</i>
<i>LINK 1</i>	<i>91</i>
<i>LINK 2</i>	<i>92</i>
<i>LINK 3</i>	<i>92</i>
<i>Cross-Validation of the Data</i>	<i>93</i>
MANAGING INCONSISTENCIES IN THE DATA.....	93
OTHERWISE NORMAL PROFILE AND C-SEP™	94
DECISION-MAKING AND PROFESSIONAL JUDGMENT	95
REPORT WRITING.....	99
DOS AND DON'TS.....	100

Table of Contents

<i>Timelines</i>	101
<i>Strengthening Legal Defensibility</i>	101
<i>Avoid Legalese</i>	102
<i>Ensure Accessibility</i>	104
<i>Avoid Shortcuts</i>	105
STRUCTURE AND CONTENT	105
<i>Structuring a Report</i>	106
<i>Document Assessment Procedures</i>	107
<i>Integrate and Synthesize Data</i>	107
<i>Include Impacts and Needs Statements</i>	108
<i>Offer Meaningful Recommendations</i>	109
SPECIAL POPULATIONS	111
C-SEPT™ FOR DYSLEXIA ASSESSMENTS.....	113
C-SEPT™ AND CLD/EL EVALUATIONS	119
FAQs	127
C-SEPT™ CASE ILLUSTRATIONS	131
CASE STUDY: LORI (FULL)	133
CASE STUDY: OLIVIA (DNQ)	139
CASE STUDY: TIMOTHY (ACHIEVEMENT)	143
CASE STUDY: VIRGINIA (INTELLECTUAL DEVELOPMENT & ACHIEVEMENT).....	145
REFERENCES	149
C-SEPT™ LEGAL ALIGNMENT	157
RESOURCE LINKS	161
C-SEPT™ QUICK START (TESTING PROCEDURES)	163

Page Intentionally
Left Blank

Abbreviations

ACH	Achievement
ADHD	Attention Deficit Hyperactive Disorder
ARD	Admission, Review, Dismissal
BICS	Basic Interpersonal Communication Skills
C/DM	Concordance-Discordance Method
CALP	Cognitive Academic Language Proficiency
CBA	Curriculum-Based Assessment
CBM	Curriculum-Based Measurement
CELF	Clinical Evaluation of Language Fundamentals
CFR	Code of Federal Regulations
CHC	Cattell-Horn-Carroll
CLD	Culturally and Linguistically Diverse
CLI	Comparative Language Index
CMB	Curriculum-Based Measurement
COG	Cognitive
C-SEP	Core Selective Evaluation Process
C-TOPP	Comprehensive Test of Phonological Processing
D/CM	Discrepancy/Consistency Method
DA	Dynamic Assessment
DD/C	Dual Discrepancy/Consistency
DIBS	Dynamic Indicators of Basic Skills
DNQ	Does Not Qualify
EF	Exclusionary Factors
EL	English Learner
ESSA	Every Student Succeeds Act

Abbreviations

FAQ	Frequently Asked Questions
FBA	Functional Behavioral Analysis
FIE	Full and Individualized Evaluation
FIIE	Full and Individualized Initial Evaluation
FRQ	Focused Referral Question
FSIQ	Full Scale Intelligence Quotient
GAI	Gross Ability Index
Gen. Ed	General Education
GORT	Gray Oral Reading Test
HLS	Home Language Survey
ID	Intellectual Disability
IDA	International Dyslexia Association
IDEA	Individuals with Disabilities Education Act
IDEIA	Individuals with Disabilities Education Improvement Act
IEP	Individualize Education Plan
IQ	Intelligence Quotient
KTEA	Kaufman Test of Educational Achievement
LD	Learning Disability
LEP	Limited English Proficient
LEA	Local Education Agency
LEP	Limited English Proficiency
MSD	Multiple Sources of Data
MSDW	Multiple Sources of Data Worksheet
NCLB	No Child Left Behind
NIH	National Institutes of Health
NRT	Norm-Referenced Test
OL	Oral language
PJ	Professional Judgment

Abbreviations

PLAAFP	Present levels of academic achievement and functional performance
PM	Progress monitoring
PR	Percentile Rank
PSW	Pattern of Strengths and Weaknesses
RIOT	Review, Interview, Observation, Test
ROI	Rate of Improvement
RPI	Relative Proficiency Index
RTI	Response-to-Intervention
SBOE	State Board of Education
SD	Standard Deviation
SES	Socioeconomic Status
SLD	Specific Learning Disabilities
SLP	Speech-Language Pathologist
SS	Standard Score
SST	Student Support Team
WIAT	Wechsler Individual Achievement Test
WISC	Wechsler Intelligence Scales for Children
WJ	Woodcock-Johnson
WMLS	Woodcock-Muñoz Language Survey
WRAT	Wide Range Achievement Test
XBA	Cross-Battery Assessment

Page Intentionally
Left Blank

Forward

We would like to dedicate this book to all the diligent educational evaluators that are serving students.

~ Tammy L. Stephens, Ph.D. and Edward K. Schultz, Ph.D.

Page Intentionally
Left Blank

Defining the Core Selective Evaluation Process™ (C-SEP™)

“Work Smarter, Not Harder”

“Method Matters”

Page Intentionally
Left Blank

Origins of C-SEP™

Core-Selective Evaluation Process™ (C-SEP™), a comprehensive pattern of strengths and weaknesses (PSW) model, was first introduced to the field in 2015 in response to what Dr. Schultz and I describe as assessment malpractice witnessed through our interactions with evaluators across the United States. Dr. Schultz and I have spent decades working in the field of special education as teachers, evaluators, and university professors. During this time, we have witnessed the use of different methods for determining a Specific Learning Disability (SLD), many of which overemphasized standard score discrepancies with little emphasis on multiple sources of data (MSD). Consequently, most evaluators were using what we refer to as a “standard protocol” approach when evaluating SLD.

Specifically, the standard protocol approach places heavy focus on standard scores and seldom considered other sources of data when establishing a pattern of strengths-and-weaknesses (PSW). Additionally, the standard protocol method requires a large amount of testing. In most cases, every student referred for SLD evaluation (regardless of the referral question or data already collected) was administered over thirty subtests/tests. This testing regiment included being evaluated in areas where there was no indication or a history of academic problems. Such practice is extremely time consuming and costly, and is exhausting for the student, the latter of which raises questions about the validity and reliability of the results acquired. Furthermore, over-emphasis on norm-referenced test (NRT) scores while undervaluing other data sources does not allow the evaluator to truly establish a PSW.

C-SEP™ was conceptualized due to a need for an SLD method that had a targeted use of NRT, integrated MSD to plan a more purposeful assessment, and aligned with policy (a legally- defensible approach). For example, a triangulation of MSD is necessary to truly establish a PSW. In fact, special education policy mandates that “a variety of assessment tools and strategies be used to gather relevant information” (34 CFR §300.304). We believe that C-SEP™ offers solutions to many of the weaknesses that we have observed.

Definition

Core-Selective Evaluation Process™ (C-SEP™) is defined as a PSW approach to SLD identification that involves a targeted, purposeful assessment process that integrates MSD with NRTs

and professional judgment. Conceptualized following the release of major revisions and improvements of individualized norm-referenced tests of cognition, language, and achievement, in most cases, less formal testing is required. By utilizing tests for their intended purposes (e.g., using selective testing practices and fully exploiting the data collected from NRTs to include qualitative information, error analysis, testing observations, and scores), the evaluator obtains a more thorough understanding of the learner. When conducting a comprehensive assessment using C-SEP™, the evaluator organizes and analyzes the referral data to establish a focused referral question (FRQ). The FRQ and student's background are then used to determine which NRT to use as part of the assessment.

“Work Smarter, Not Harder,” “Method Matter,” “C-SEP™ Does Not Fear the Split,” and “Beyond the Score” are popular descriptors of the C-SEP™ method. Using C-SEP™, the core set of tests are initially administered, and results are merged with other sources of data to determine if the referral question can be answered. Selective (or “Core-Selective”) testing can be utilized whenever additional information is needed to answer the referral question and fully understand the learner. Four key pillars of C-SEP™ serve as the building blocks of a comprehensive assessment, these include: *Publisher guidance* (what does the examiner's manual recommend?), *Policy* (what does the law mandate?), *Professional Judgment* (what does the evaluator's expertise say about the data collected?), and *Practices* (what are best practices in SLD assessment?). More information about the four Pillars of C-SEP™ will be provided later in this manual.

By comparison to other PSW models, C-SEP™ is an assessment method that collects and analyzes multiple sources of data—it is **not** a testing model that focuses on a single test, and it does not require over-testing. See Box 1 for an explanation of testing versus assessment. C-SEP™ can be used with different test instruments depending on the expertise and professional judgment of the evaluator and the individual needs of the student being assessed. In this frame, test selection is proportional to the complexity of the assessment, which is based on the presenting problem or referral question, the evaluator's professional judgment regarding the existing data and the assessed student (Schrack, Stephens-Pisecco, & Schultz, 2017). The evaluator is thus at liberty to select test instruments, provided they are used according to publisher guidance, they meet the individual needs

of the student being tested (e.g., age, cultural background), and are suited for testing the area(s) of suspected disability, as required by law. Therefore, C-SEP™ adheres to CFR 300.304(c) which mandates that non-discriminatory methods be used when assessing students that are culturally and linguistically diverse (CLD) (IDEA, 2004).

Box 1. Testing versus Assessment

Testing versus Assessment:	<i>Testing</i> involves the administration of a test to obtain a product: a score. In this frame, testing (and the score it generates) is but one component of an assessment. <i>Assessment</i> is the process of gathering multiple sources of data. These data can include, but are not limited to, observations, recollections, previous test results, work samples, parent/teacher input, and professional judgment. Assessment is therefore much broader than testing.
-----------------------------------	---

Within the C-SEP™ framework, PSW emerges through the integration of MSD with core and selective norm-referenced cognitive, language, and achievement test results (Schultz & Stephens, 2015). When implementing C-SEP™, MSD is collected, organized, and analyzed in a manner to both understand the learner as well as to steer the assessment process and inform the testing plan. C-SEP™ holds that only through the integration of all MSD can PSW emerge. Within this context, C-SEP™'s emphasis on careful analysis and synthesis of MSD, and its explicit utilization of PSW for determining eligibility, provides examiners with a unique, valid, and legally defensible approach for assessing SLD.

When analyzing data using C-SEP™, equal weight is placed on all valid MSD, whether it was collected in the past or collected during the on-going evaluation. In this context, the model recognizes that a student's existing data contains valuable information that can be leveraged for demonstrating a pattern. More specifically, if the evaluator carefully analyzes and considers the MSD available and the referral question, they will notice a pattern emerge and can thereby strategically target their NRT at *all* the areas of suspected disability (34 CFR § 300.304(c)(4)). Once selected NRTs have been administered and analyzed, their findings should be integrated and triangulated with the previously existing data to identify a true pattern of PSW whereby an informed decision can be made about whether an SLD exists. MSD is leveraged to help focus the referral question, to create a pattern of

strengths and weaknesses, to make an informed determination about a student's eligibility, and to develop individualized recommendations to meet the student's individual needs.

Contextual Background

A Specific Learning Disability (SLD) includes one of the thirteen special education eligibility categories under IDEA (2004). The SLD category makes up most individuals served under IDEA (2004). Box 2 provides a definition of SLD according to federal law (34 CFR 300.8(c)(10)). While the definition of SLD has not changed in decades, the methods used for identifying whether an SLD exists have changed dramatically.

Box 2. Definition of Specific Learning Disability (SLD)

Specific Learning Disability (SLD):	Is a DISORDER in one or more of the basic psychological processes involved in understanding or in using LANGUAGE, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. The term does not include learning problems that are primarily the result of visual, hearing, or motor disabilities; of intellectual disability; of emotional disturbance; or of environmental, cultural, or economic disadvantage. (34 CFR 300.8(c)(10))
--	---

It is important for evaluators to have a clear understanding of what SLD is and what it is not. Such knowledge will help them to make informed and appropriate decisions. Table 1 provides a list of SLD characteristics by comparison to characteristics that are not reflective of SLD. While not comprehensive, this table is useful for underscoring differences.

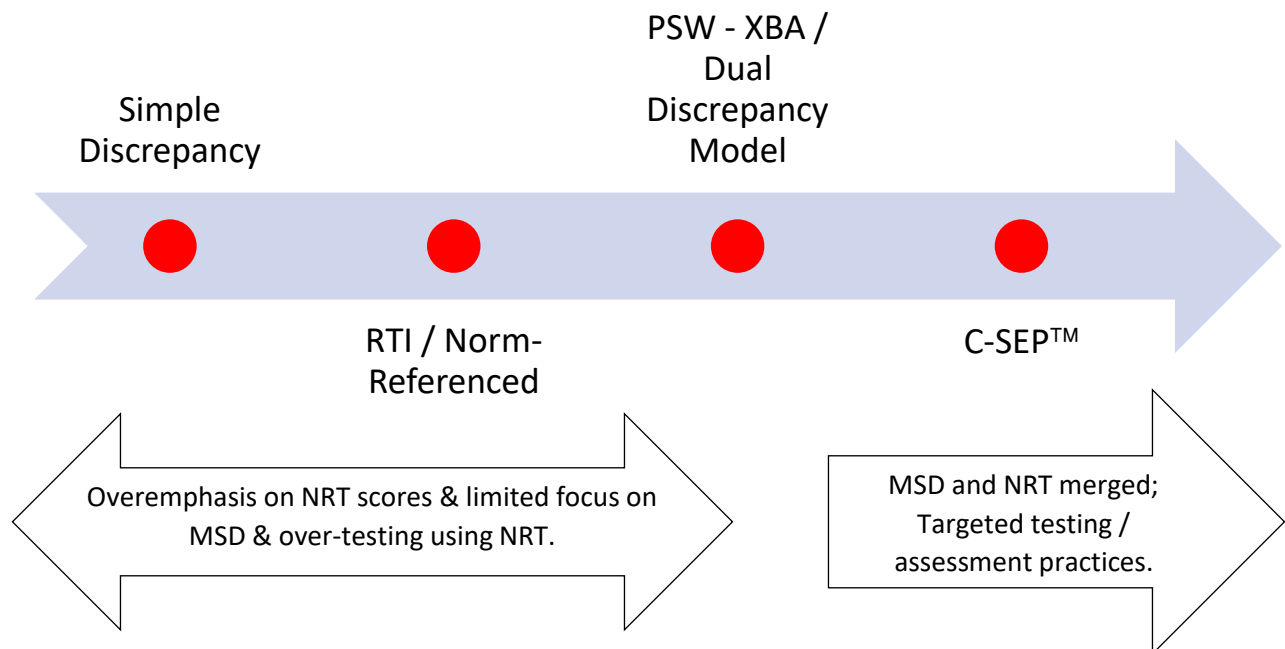
Table 1. What is SLD?

SLD is...	SLD is not...
Explained by a neurologically based processing deficit or deficits.	Characterized by low or below average cognitive abilities with little or no areas of strength.
Sometimes appears in conjunction with other disability conditions (sensory, language, behavioral).	Characterized by processing weakness(es) that are not linked with academic weakness(es).
An educational disability.	Explained primarily by low or below average cognitive abilities, another disability category, or an exclusionary factor.
A disability category under Federal IDEA regulations.	Explained by external factors such as instructional or environmental variables.
A learning difficulty that ranges in type and directly impacts on academic performance.	Primarily explained by another disability, and/or condition (e.g., emotional disturbance, intellectual disability, and so forth).
An impairment requiring a comprehensive and individual evaluation by a team of experts to ensure all SLD federal, state, and district standards and criteria are met.	Solely a medical or mental health diagnosis.
An educational classification in which a student meets the criteria for SLD, so much so that s/he cannot profit in the general education curriculum without special education support.	An automatic entitlement for students with any academic difficulties. An automatic default category when a student demonstrates lack of progress in the general education setting. Applied when a student exhibits a pattern of strengths and weaknesses but does not require special education support to benefit from the general education curriculum.

Assessment Practices for SLD

Historically, most methods for identifying SLD have focused on standard score discrepancies, which has resulted in an overemphasis on standard scores (cut scores) when making eligibility decisions. Figure 1 provides a continuum of assessment practices utilized when assessing for SLD over the past two decades. It likewise provides a brief overview of the weaknesses of each model identified. The figure is followed by additional information on two PSW models with a denotation of their pros and cons.

Figure 1. Assessment Practices Continuum



Simple Discrepancy Model: When using the simple discrepancy model, the determination of a SLD was based on whether a specific discrepancy existed between the full-scale IQ (FSIQ) and one or more areas of achievement. If a specified discrepancy (e.g., 16-point discrepancy) between the FSIQ and a standard score obtained in an area of achievement (e.g., Basic Reading, Reading Comprehension, Math Calculation, Written Expression) existed and the student had an educational need for specialized instruction, the student would be found to have a SLD in that area. The Simple Discrepancy Model has pros and cons which are outlined below:

Pros: Simple to implement and explain.

Cons: Wait to fail model, overemphasis on NRT scores, overemphasis on FSIQ, limited consideration placed on other sources of data, and the results did not inform instructional needs.

Response-to-Intervention (RTI) and NRTs: Response-to-Intervention (RTI) is a multi-tier approach that is designed to make early identifications of students with learning and behavior needs whereby they are then offered support. The RTI process begins with high-quality instruction and universal screening of all children in the general education classroom. Struggling learners are provided with interventions at increasing levels of intensity to accelerate their rate of learning. Student progress or rate of improvement (ROI) is monitored through frequent assessments using curriculum-based measurements (CBM). Data obtained from the student's RTI is merged with other MSD, including data obtained from NRTs.

Pros: Integration of data demonstrating a student's RTI and NRT results are integrated within the assessment process to obtain a comprehensive understanding of PSW.

Cons: The quality of RTI implementation, intervention fidelity, and data collection vary from school to school resulting in a variability in the quality of data obtained. Overuse of norm-referenced tests resulting in over-testing students.

Box 3. Common Features of a Comprehensive PSW Approach

1. Multiple sources of data are collected **over time** using a variety of assessment tools and strategies.
"Over time" suggests that at least one year's worth of data be collected to identify patterns and trends in behavior and/or academic performance. Nonetheless, the more information that is collected over longer periods of time is going to produce a more robust and insightful pattern.
2. Data analysis is grounded in pattern seeking techniques.
3. Predictive and treatment validity are used.
4. Evidence-based and logical decision making is applied (Schultz, Simpson, & Lynch, 2012).

Pattern of Strengths and Weaknesses (PSW) Models:

A PSW model is a "third method" approach to identify a specific learning disability (SLD) through the identification of patterns of strengths and weaknesses in cognition, achievement, or both. The most prevalent PSW approaches include the Concordance-Discordance Method (C/DM; Fiorello,

Hale, & Snyder, 2006), the Discrepancy/Consistency Method (D/CM; Naglieri, 1999; Naglieri, 2011), the Dual Discrepancy/Consistency (DD/C; Flanagan, Ortiz, & Alfonso, 2013), Dehn's processing approach (Dehn, 2014), and C-SEP™ (Schultz & Stephens, 2015). Table 2 provides a comparison of the features of each PSW model.

Table 2. The Three Most Common PSW Models

Cross-Battery Assessment (XBA)	Dehn's PSW Model	Core-Selective Assessment Process™ (C-SEP™)
• Grounded in CHC Theory • Seven cognitive clusters • Dual Discrepancy / Consistency Model • Domain specific cognitive weakness • Unexpected underachievement • Empirical link between cognitive and academic deficit. • Overemphasis on standard score discrepancies • XBASS Software	• Discrepancy / Consistency Model • Intra-individual processing weakness • Processing weakness links to academic weakness • Psychological processes • Overemphasis on standard score discrepancies • Psychological Processing Analyzer	• Grounded in Contemporary cognitive theory (e.g., CHC or Pass Theory). • Intellectual Development is measured. • Targets assessment and testing – Less emphasis on norm-referenced tests (NRT). • Data from NRT is fully exploited and integrated into the assessment (e.g., RPI, CALP, testing observations, error analysis, task demand analysis, and so forth). • Language is assessed for all students. • Task demand analysis conducted to further understand strengths and weaknesses. • PSW emerges through integration of: • Multiple Sources of Data • Standardized cognitive, achievement, oral language assessment. • Professional judgment • Valid data analysis

All PSW models share certain features and processes. For instance, they generally agree that a student with SLD has the following characteristics: a) they are not responsive to appropriate traditional and supplemental instruction; b) they exhibit a disorder of basic psychological processes that directly impacts the identified area of academic concern; c) they exhibit a pattern of strengths in the psychological processes that are linked to academic performance; and d) they have identifiable deficits that are not the primary result of exclusionary factors. However, of the PSW approaches identified, C-SEP™ is more comprehensive and differs from the other PSW models in the manner NRTs are used and the way PSWs are determined. Table 2 provides a general comparison of C-SEP, XBA, and Dehn's model.

Box 4. Weaknesses of other PSW models

C/DM, D/CM, and DD/C, and Dehn share several important and interconnected weaknesses. Foremost, they are constructed around a "Standard Protocol Approach." This requires a set number and type of tests be administered regardless of the referral question (e.g., SLD Basic Reading, Math Calculations) or type of referral (e.g., initial evaluation, re-evaluation). Such practices often require a student to be evaluated in areas where no disability is suspected. This practice might be considered a misuse of time and resources, and it can be taxing on the student being examined.

Second, these models often do not utilize the normative scores derived from a test publisher and frequently do not require evaluators to examine and consider scores in conjunction with other data sources. Instead, test scores are entered into a software program and findings are generated. Unfortunately, the scores generated by such software are sometimes manipulated and based on contrived norms. Equally concerning, the software-generated scores may only provide a snapshot of student performance since they are not taken into consideration with other multiple sources of data.

Third, and connected to the previous, some PSW models tend to base their findings exclusively on the standard scores of tests administered. After entering the scores into the software, a profile is generated. Thereafter, the profile frequently becomes the primary reference point when making an SLD eligibility determination. In instances where other data is considered, it may only be used to support the findings generated by the software.

Finally, some PSW models minimize the role of professional judgment. This means that the experience and insight possessed by a trained evaluator is muted in leu of a software program that only processes and considers a limited amount of the total data available on a student. While such practices may expedite decision-making and free evaluators from the challenge of collecting, integrating, and examining data, it may not always demonstrate a pattern because it overlooks MSD while marginalizing the knowledge and experience that evaluators and teachers possess.

Without question, PSW methods are beneficial in that they help answer complex referral questions. Nonetheless, many PSW methods are likewise limited in their comprehensiveness, they can be inefficient, they sometimes lack precision, and they are not always implemented in a manner that adheres to statutory requirements. During our critical analysis of the three PSW models, Schultz and Stephens (2008) identified several of their limitations which are highlighted in Box 4. Despite these weaknesses, evaluators continue to use these models to identify SLD.

SLD Identification

SLD identification is mandated and regulated by federal and state law. 34 CFR §300.304, as one section of applicable federal law, provides concise standards for SLD evaluation procedures. Among its numerous requirements are the “use of a variety of assessment tools and strategies to gather relevant functional, developmental, and academic information about the child, including information provided by the parent” (34 CFR § 300.304(b)(1)); that are administered in a manner that meets publisher guidance (34 CFR § 300.304(c)(1)(v)); that “the child is assessed in all areas related to the suspected disability, including, if appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities (34 CFR § 300.304(c)(4)); and that “the evaluation is sufficiently comprehensive to identify all of the child's special education and related service needs” (34 CFR § 300.304(c)(6). While it is beyond the scope of this manual to address every applicable federal and state statute, we highly recommend evaluators familiarize themselves with all applicable laws, not just those highlighted here, and to adhere to them.

According to Federal standards, when conducting a full and individual initial evaluation (FIE), the Local Education Agency (LEA) must use a variety of assessment tools and strategies to gather relevant information on (34 CFR § 300.304):

- Functional (e.g., daily living skills, social skills, communication skills, behavior)
- Developmental (e.g., milestones)
- Academic information (reading, spelling, writing, math calculations and problem solving, science, social studies)
- Parent input

The relevancy of this information implies that it measures one of the characteristics listed, it is relevant to identify the PSW and any educational needs the student has, and it is useful for establishing student goals and objectives.

Such data should be acquired through school records, parents, teachers, and a variety of professions. The law equally spells out other rules. For instance, the LEA (34 CFR § 300.304):

- May not use a single measure or assessment as the sole criterion.
- Must evaluate the child in all areas related to the suspected disability.
- Must use technically sound instruments that are racially and culturally unbiased; administered in the student's native language; used for the purpose for which they are valid and reliable; and are administered as designed by trained professionals.
- Must use the assessment data to assist the IEP/ARD committee in determining the educational needs of the child.
- Must evaluate in a manner that is sufficiently comprehensive to identify the child's special education and related service(s) needs.
- Must review existing relevant evaluations and data to determine what additional data are needed.

After the individualized and comprehensive evaluation has been conducted, four conditions must be satisfied for a child to meet SLD eligibility requirements under federal law (34 CFR § 300.309):

1. MSD must show that appropriate instruction has been provided, whether it is demonstrated prior to or as part of the referral. Supporting evidence may be obtained through Response-to-Intervention (RTI) systems or another type of supplemental instruction.
2. The student must be found to not perform satisfactorily for the child's age or does not meet state-approved grade-level standards (IDEA, 2004). Multiple measures should be used to determine underachievement (e.g., Curriculum-Based Measurement [CBM], Curriculum-Based Assessment [CBA], state-wide assessments, grades, work samples).
3. The evaluation must eliminate other factors that may account for a student's academic skills weakness and learning difficulties. These exclusionary factors include vision/hearing, motor disabilities, intellectual disability (ID), social/emotional or psychological disturbance, environmental or economic disadvantage, and cultural and linguistic factors (e.g., limited English proficiency). While the assessment team must rule out each of these factors as the primary cause of a student's academic and learning difficulties, the degree of influence or their contributions to the learning problems must also be addressed.
4. A pattern of strengths and weaknesses must reveal significant variance among definite areas of cognitive function (e.g., working memory and verbal comprehension); or between specific areas of cognitive function and academic achievement (Evans, Floyd, McGrew, & Leforgee, 2001; Flanagan, Fiorello, & Ortiz, 2010; Floyd, Evans, & McGrew, 2003; Floyd, Meisinger, Gregg, & Keith, 2012; McGrew & Wendling, 2010; Newton & McGrew, 2010). While "significant variance" is not always defined in regulations, a variation is deemed important and meaningful (practical) when scores differ by ~1 SD when considering confidence intervals of norm referenced tests. Such patterns can

be identified using NRTs of cognition, language, and achievement, but must also be supported by other data sources integrated into the assessment.

Significant variance means a statistically large or clinically meaningful difference. While clinically significant differences are rare, a statistically meaningful difference is suggested to be enough to impact classroom performance. When evaluating specific areas of cognitive functioning to determine a PSW, they should be linked to the failure to achieve adequately when determining SLD. In the end, a student should not be determined to have a SLD when their classroom achievement indicates a pervasive weakness, yet they do not meet the threshold of a PSW.

Definition of Patterns of Strengths and Weaknesses

Since C-SEP™ is a PSW model and federal law requires a pattern be identified to determine SLD eligibility (34 CFR § 300.309(a)(2)), it is essential to provide an operational definition of *a pattern* of strengths and weaknesses. A *pattern* is defined as a combination of qualities, acts, and tendencies forming a consistent or characteristic arrangement (Random House, 2011). Similarly, Cambridge dictionary defines a pattern as “a particular way in which something is done, is organized, or happens.” In the context of SLD identification and PSW, a pattern is a characteristic of performances and/or behaviors that develops over time. A pattern is not illustrated through a one-time snapshot acquired by a single test score (which is a student’s testing profile), but rather a pattern is identified through MSD collected over time that demonstrate the repetitive nature of a particular type of performance or behavior (either as a strength or a weakness).

To identify patterns and establish (or rule out) SLD eligibility, a PSW approach expects these criteria be established:

- There are MSD collected over time using a variety of assessment tools and strategies,
- Data analysis is grounded in the techniques of pattern seeking (McMillan & Schumacher, 2010),
- Predictive and treatment validity are utilized,
- Logical and empirical evidence is used to guide decision-making.

As emphasized above, establishing a pattern of strengths and weaknesses requires the evaluator to not over-rely on data that provides a “one-time snapshot.” A pattern is therefore not obtained by simply administering standardized cognitive, achievement, and oral language tests because the results acquired from these scores provide a profile of the student’s performance on that

test at that given time. Thus, standardized test results alone DO NOT establish a pattern. Instead, these test results should be treated as one piece of evidence that proves a pattern. It is likewise beneficial for better understanding the area of suspected weakness. A pattern only emerges from the collection and interpretation of multiple sources of data (including standardized test scores).

To establish a PSW during an individualized and comprehensive SLD evaluation, members of the team will need to:

1. Identify an academic need in / a suspected disability that is impacting one (or more) of the eight areas of eligibility found in federal guidelines for SLD (basic reading, reading comprehension, reading fluency, math calculations, math problem solving, listening comprehension, oral expression) (see 34 CFR § 300.309(a)(1)),
2. Determine if there is an area (or areas) of cognitive weakness that has a research-based link to problems in the identified academic area (see 34 CFR § 300.309(a)(2)),
3. Establish whether there are other cognitive areas which are average or above (see 34 CFR § 300.309(a)(2)), and,
4. Analyze these findings using MSD to demonstrate a pattern that rules out or confirms the presence of SLD (see 34 CFR § 300.309(a)(2)).

What becomes evident in points 2 through 4 above, and is essential to making an SLD identification, is the establishment of data-proven patterns that demonstrate the student possesses both strengths and weaknesses.

Page Intentionally
Left Blank

Core-Selective Evaluation Process™

Within the C-SEP™ framework, PSW emerges through the integration of MSD with core and selective norm-referenced cognitive, language, and achievement test results (Schultz & Stephens, 2015). C-SEP™ holds that only through the integration of all data will PSWs be demonstrated. When implementing C-SEP™, MSD is collected, organized, and analyzed in a manner to both understand the learner as well as to steer the assessment process and inform the testing plan. In this frame, C-SEP™'s emphasis on a careful analysis and synthesis of MSD, and its explicit utilization of PSW as a framework for eligibility, provides examiners with a unique and defensible approach for assessing students as it adheres to federal policy.

This chapter outlines the basic tenants and structure of C-SEP™ to give the reader a general overview. We begin by demonstrating how C-SEP™ is a legally defensible PSW model. We then examine the importance of assessing language as it is foundational for learning. Thereafter, we introduce the basic framework of the methodology, outlining the four Ps of C-SEP™ before introducing the four steps of the C-SEP™.

A Legally Defensible PSW Model

C-SEP™ is a PSW approach to SLD identification whose design reflects current federal law and state policy regarding SLD identification. The title “C-SEP™” describes the ways in which norm-referenced tests are used in the context of integrated data analysis techniques, contemporary policy, and academic research pertaining to the construct of SLD. Using C-SEP™, each of the pertinent evaluation procedures are addressed and documented as part of an individualized, comprehensive evaluation. When following the framework, data is collected, documented, and analyzed to demonstrate a pattern of strengths and weaknesses whereby a legally defensible eligibility decision is grounded in a preponderance of the data.

C-SEP™ assesses a broad set of cognitive, achievement, and oral language abilities to effectively identify all special education and related service needs, creating a comprehensive and individualized assessment. It utilizes a variety of assessment tools and strategies prior to and as part of the evaluation process. Co-normed tests, for instance, are utilized to obtain a formal

representation of the student’s ability. These scores are interpreted using the test publisher’s software and normative standards. However, this normative data is not over-weighted, as it is then merged with other available data prior to making eligibility decisions. In this context, C-SEP™ is comprehensive since it requires the integration of formal assessment results from MSD associated with all areas of suspected disability. See Box 5 for a brief overview of the model.

Box 5. C-SEP™ Overview

- Uses individualized assessment practices that are guided by the referral question.
- Uses a testing hypothesis to establish a focused referral question (FRQ).
- Uses MSD to establish underachievement, conduct preliminary rule-out of exclusionary factors, and establish initial emergence of a pattern of strengths and weaknesses.
- Selects “core” tests as the foundation of the formal evaluation.
- Considers the results of targeted “core” tests in relation to other sources of data and analysis to determine whether more “selective testing” is necessary.
- Considers “selective testing” results in conjunction with core tests and MSD.
- Fully exploits NRTs, considering all the data obtained (e.g., testing observations, error analysis, task demands analysis, RPI).
- Utilizes pattern seeking strategies.

While processing data, C-SEP™ uses data integration techniques to include authentic assessment strategies (e.g., observation, work samples, curriculum-based assessments) which are valuable measures for helping to determine educational needs. The evaluator also uses their professional judgment, supported by a preponderance of data, to determine SLD. Once decisions are made, C-SEP™ documents the data collected, their, and clearly explains why the eligibility determination was made and recommends to teachers and parents the services or accommodations that are needed. These processes adhere to federal and state law and the best practices embraced by the assessment profession.

Upon this foundation, we believe that the best place to begin to flesh out our SLD assessment methodology is to root it in existing laws and expectations. Our profession has clearly established standards that dictate how assessments should be designed, conducted, and managed. It is imperative that evaluators, schools, and districts adhere to the federal and state laws governing the

SLD identification process. As a means of presenting C-SEP™'s alignment with legal standards, Table 3 provides a side-by-side comparison of federal procedures with those of the model.

Table 3. Relationship of C-SEP™ with Federal Regulations of SLD Identification

Federal Regulations	C-SEP™
A variety of assessment tools and strategies must be used to gather relevant functional, developmental, and academic information about the child, including information provided by the parent, which may assist in determining whether a child has a disability; and this data must be used for individualized educational planning.	C-SEP™ is a strategic approach that helps collect, organize, and integrate a variety of data sources to determine if a child meets the state requirements for SLD identification.
No single measure or assessment can be used as the sole criterion for determining whether a child has a disability or for determining an appropriate educational program for the child.	C-SEP™ uses numerous measures and integrated data analysis to ensure that SLD identification is based on multiple criteria.
Technically sound instruments that may assess the relative contribution of cognitive and behavioral factors, in addition to physical or developmental factors, must be used.	C-SEP™ procedures rely on the technical specifications of the test publisher for norm-reference tests and ensure reliability and validity by integrating multiple sources of data.
Assessments are used for the purposes for which the assessments or measures are valid and reliable.	C-SEP™ procedures require adherence to the publisher's administration and interpretive guidelines. In addition, publisher software and/or tables should be used for scoring the student's performance on the tests.
Assessments are administered in accordance with the instructions provided by the producer of the assessments.	C-SEP™ procedures require adherence to the publisher's administration and interpretive guidelines. In addition, publisher software and/or tables must be used for score calculations.
Assessments and other evaluation materials, including those tailored to assess specific areas of educational need, should be utilized.	C-SEP™ procedures emphasize diagnostic precision, which results in a deeper understanding of "specific areas of educational need."
Children must be assessed in all areas related to the suspected disability, including, if appropriate, health, vision, hearing, social and emotional status,	C-SEP™ is comprehensive, as it requires the integration of formal assessment results with other

general intelligence, academic performance, communicative status, and motor abilities.	data sources to include all areas related to the suspected disability.
--	--

When evaluating a child with a disability, the evaluation must be sufficiently comprehensive to identify all the child's special education and related service needs, whether or not commonly linked to the disability category in which the child has been classified.	C-SEP™ is comprehensive, as it requires the integration of formal assessment results with multiple data sources to include all areas related to the suspected disability. In addition, C-SEP™ assesses a broad set of cognitive, achievement, and oral language abilities to effectively identify all special education and related service needs.
---	--

Assessment tools and strategies must provide relevant information that directly assists persons in determining the educational needs of the child.	C-SEP™ uses data integration techniques that incorporate authentic assessment strategies (observation, work samples, curriculum-based assessments) that directly assist in determining educational needs.
--	---

<https://framework.esc18.net/display/Webforms/ESC18-FW-Summary.aspx?FID=121&DT=G&LID=en>

With the basic legal alignment presented, we will now turn our attention to the importance of language for academic success, and consequently for inclusion in all SLD assessments.

Language

Language has a significant and broad impact on academic performance and is recognized as being crucial for academic success (Gatlin-Nash & Dockterman, 2020). Conversation, interaction, reading, and writing are all interconnected skills associated with language acquisition (Gatlin-Nash & Dockterman, 2020). In addition, language influences the development of other academic skills. According to Mazzocco and Thomas (2005), math learning starts with basic speaking activities such as counting physical objects with parents or caretakers. Children also learn concepts of greater than and less than, full and empty, and of mass or volume with bigger or smaller, often before they learn to count. None of the math concepts mentioned require numerals to understand, but language. Thus, the earliest math skills are usually learned using language and physical objects (Kolkman, Kroesbergen, & Leseman, 2013). Nevertheless, language continues to impact math skills at the higher levels. The classic example are word problems, which requires a student to not only understand

language, but also understand the connection of language with symbolic representations of numerals and math operations (Mazzocco & Thomas, 2005).

For these reasons, and unlike other SLD identification methodologies, C-SEP™ requires language be investigated for ***all*** students being referred for SLD evaluation. We set this requirement because language plays an integral part in cognition and learning. Additionally, language is embedded into the legal definition of SLD. According to 34 CFR § 300.8(c)(10), an SLD is:

a DISORDER in one or more of the basic psychological processes involved in understanding or in using LANGUAGE, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations.

The importance of evaluating the language skills of students referred for an SLD evaluation should therefore be evident. Reference to language in the definition is unsurprising, as language is essential to learning. In fact, language skills are listed as one of the five domains of school readiness. Box 6 highlights the five domains of school readiness.

Box 6. Five Domains of School Readiness (adapted from Gregory et al., 2021).

1. Language and communicative development.
 - Uses and understands language, begins and carries a discussion, understands and discusses pictures.
2. Physical and motor development.
 - Can walk a beam, cut with scissors, complete puzzles.
3. Social and emotional development.
 - Accepts responsibility for actions, takes turns, form relationships.
4. Approaches to learning.
 - Is motivated and curious, persists in tasks, uses materials innovatively.
5. Cognitive development or general knowledge.
 - Can match and sort shapes and colors, ascribes value to numbers, is aware of cause and effect.

The complex process of acquiring language naturally begins early. The first three years of life are fundamental to speech and language development, with an estimated 90% of brain growth occurring during this period (National Institutes of Health, 2021). According to the National Institutes of Health (NIH) (2021), language skills develop best in an environment filled “with sounds, sights, and

consistent exposure to the speech and language of others.” If this window for language development is missed, it becomes harder for a child to learn.

Research and theory associated with language acquisition has multiplied over the past few decades, helping to raise awareness of its complexities and inner workings. One example is the Simple View of Reading framework (Hoover & Gough, 1990), which theorizes that reading comprehension is the product of decoding and linguistic comprehension (word recognition x language comprehension = reading comprehension). A weakness in either decoding or linguistic comprehension is suggested to undermine reading comprehension. Such insight has given way to regular monitoring of young children’s language development as a means of identifying difficulties and promoting early intervention (Farrall, 2012; National Institutes of Health, 2021). For instance, developmental checklists have been devised to track the language development of infants and youth.

Another theory of language acquisition that aids assessments is the *Reading Rope* (Scarborough, 2001). This model centers on language comprehension and word recognition. Both domains are conceptualized as consisting of several subskills, referred to as strands, that become interwoven as automaticity as skills improve and reading comprehension is mastered. Using this framework, Scarborough (2001) hypothesizes that most reading disabilities are rooted in poor phonemic awareness and decoding skills. She also suggests that language comprehension deficits equally undermine reading skills in older children, with early language impairments being highly predictive of future reading impairment.

One final scholar addressed herein, that has contributed to our understanding of language acquisition is Wasowicz (2021). Among other contributions, her research challenges our notions of literacy, which she claims is not narrowly conceptualized only as reading skills but also includes writing skills. In this framework, writing is presented as a multifaceted and integrated language process, whereby Wasowicz (2021) adds new strands to Scarborough’s *Reading Rope*, by linking writing and reading. She argues that humans are inherently wired for language, but not for reading and writing.

In addition to dissecting the intricacies of language learning processes, researchers have also demonstrated how environmental factors, such as socioeconomic status (SES), impact language acquisition. Studies show that children from lower SES families, for example, have been shown to

have slower vocabulary growth by comparison to their affluent peers. Hart and Risley (1995) attribute this gap to the quality of conversations that parents have with their children. They estimate that by age 4, children from higher SES families hear approximately forty-five million words on average, while children living in poverty hear 13 million words on average (Hart & Risley, 1995). Such variations in language input have been shown to predict a child's language development, including vocabulary growth and the speed at which they can process familiar words.

The final consideration made here about language as it pertains to SLD evaluations is to address the differences between BICS and CALP. See Table 4 for a basic overview. Basic Interpersonal Communication Skills (BICS) are the language skills we use in social situations on a regular basis. BICS are the skills students use when they interact with their peers at recess, while texting on their phones, or socializing at lunch. According to Cummins (1979), BICS do not require a great deal of cognitive resources. In fact, Cummins argued that everyone can acquire BICS in a first language regardless of Intelligence Quotient (IQ), or academic aptitude, and he estimates it takes 3 to 5 years to develop.

Another contribution that Cummins (1979) made is his hypothesis that language develops on a continuum, where "social language proficiency" develops before "academic language proficiency." He called the latter Cognitive Academic Language Proficiency (CALP). CALP is more cognitively demanding due to the complexity of the vocabulary and grammar involved. This specialized language knowledge helps students listen, read, speak, and write about complicated subject matter. CALP likewise encompasses the abilities to infer, classify, compare, evaluate, and synthesize language whereby meaning can be deduced from content. Proficiency in CALP can take an estimated five to seven years. However, in instances where students lack prior experience and/or support, CALP development can take upwards of ten years to acquire these skills.

The distinction between BICS and CALP is important when assessing students for SLD, especially English Learners (ELs) (Stephens, Olvera, & Schultz, 2022). While ELs may understand conversational English and express themselves with fluency when speaking to peers (BICS), they may struggle with academic text or scientific language (CALP). Stated differently, ELs may appear fluent in conversation, while their command of CALP may hinder their ability to understand and excel in an academic setting. In this context, the task of the evaluator is to distinguish between these distinct levels of proficiency and pinpoint whether the student lacks specific language knowledge (e.g., CALP)

or has a language acquisition issue. Similar considerations, nonetheless, may also be warranted for children from low socioeconomic backgrounds whose primary language is English, as their exposure to certain language may have been limited.

Table 4. BICS and CALP

	Basic Interpersonal Communication Skills (BICS)	Cognitive Academic Language Proficiency (CALP)
Description	• Everyday use. • Social interaction. • Less specialized. • Less cognitive demanding. • Used in a social setting. • Requires an understanding of cultural and social norms, including nonverbal cues.	• Academic. • Used in the classroom for reading and writing tasks. • More cognitive demanding due to complex vocabulary and grammar structures. • Requires specialized knowledge.
Acquisition	• Can develop in 3-5 years.	• Proficiency develops in a minimum of 5-7 years.
Example	• Engage in an informal, face-to-face conversation. • Writing a social media post. • Texting. • Reading a menu.	• Defining a scientific term. • Explaining how to solve a math problem. • Comparing and contrasting art. • Summarizing a research paper.

(Table adapted from Cummins, 1979)

Various instruments are available to measure CALP or its rough equivalent. For instance, the Woodcock-Muñoz Language Survey, Third Edition (WMLS III; Woodcock, Alvarado, Ruef, & Schrank, 2017) assesses CALP by rating students on a scale of one to six. Those with a low CALP level have negligible skills while those with a higher CALP level have advanced skills. By comparison, the WMLS III does not use CALP levels, but rather uses Language Proficiency Levels. Its continuum begins with Initial Development, then Early Development, Continuing Development, Emerging Proficiency, Proficient, and finally Advanced Proficient. In the WMLS III (Woodcock, Alvarado, Ruef, & Schrank, 2017), CALP is measured indirectly. CALP can likewise be measured informally, for instance, as one

component of a classroom observation. A Classroom Language Demands worksheet (See Resource Links), or similar, might equally be used to obtain an informal understanding of a student's language ability.

This section on language, we hope, conveys the importance of always assessing a student's language skills when determining SLD eligibility. For these reasons, C-SEP™ demands language be assessed even when language specialists are not involved on the assessment team. Following the above overview, we now briefly introduce the C-SEP™ model, identifying its four standards (the Four Ps) and four steps.

The Four Ps

When utilizing C-SEP™, evaluators must consider and implement the Four Ps—pillars of the model. These include Publisher Guidance, Policy, Practices, and Professional Judgment. We describe and contextualize each in the following subsections.

Publisher Guidance

The first P refers to the guidance a test publisher provides with their instruments. Evaluators must fully understand the norm-referenced tests they are using to assess students. This requires that they carefully read the test manuals of each of the instruments selected. Publisher guidelines explain everything from how to administer the tests, for whom they are designed, as well as the types of information, scores, and comparisons that are available.

Delving into the test manuals and reviewing the technical adequacy, normative update, and potential biases associated with each test instrument helps the evaluator carefully select and utilize norm-referenced standardized assessments. To this end, evaluators must read the test manuals and investigate the norm samples utilized during test standardization to ensure that the norm samples reflect the linguistic and cultural experiences of the student they wish to assess. Once the norm sample has been verified, evaluators should then consider which tests are most appropriate for the student, as well as consider their normative makeup when interpreting the results (Holman & Stephens, 2021). For example, when assessing an EL student, evaluators must determine if there is a

limitation in the standardized assessment tools that include EL groups/subgroups in the norming samples (Barrett & Mathews, 2021; Comerchero & Krach, 2014). Even standardized tests available in Spanish may not accurately reflect ELs born and educated in the United States (Doll & Plata-Potter, 2021; Braden & Mathews, 2021). Our hypothetical example explains why evaluators must obtain a clear understanding of which population the instrument is normed to. It equally points out that evaluators must know the constructs being measured by each test and which skills may be required to perform successfully on those tests, especially when the child does not have native-like proficiency. Finally, a clear understanding of the language loadings of the tests is imperative when choosing which tests to administer to ELs. Recognizing and respecting the limitations of standardized assessments is therefore vital when assessing and interpreting results for students.

In addition to the above, test publisher guidance also broadly explains how to analyze and interpret the results obtained. C-SEP™ consequently recommends an evaluator be knowledgeable regarding the constructs that each test measures, as well as understands how the performance(s) and skills being measured can impact learning. Moreover, when interpreting test results, C-SEP™ discourages the use of secondary software programs, meaning software that is not produced or recommended by the test publisher. Test instruments are informed by Standards for Educational and Psychological Testing which are reported in the publisher test manuals, and they describe sources of measurement error, interpretive information, as well as the validity and reliability of the instrument. Taking such guidance into consideration allows for a more accurate interpretation and ensures proper use of the instrument, in comparison to external software programs.

Finally, best practices dictate that actual norms be used as they are more precise and more defensible than calculations that combine tests from different batteries, the latter of which are problematic because they were normed at contrasting times, with different populations, and derived under different conditions. Methodology that departs from testing manuals, do not follow as rigorous of a process, or inappropriately use external software can lead to measurement and interpretive errors. All these guidelines adhere to IDEA (2004), which states that tests be administered in accordance with any instructions provided by the producer of the assessments.

Policy

Regardless of the methodology utilized to identify SLD, it is essential that evaluators adhere to applicable federal, state, and local laws or policy. Among its edicts, IDEA outlines the evaluation process standards, special education eligibility categories, and demands the ruling out of other causes of learner struggles (e.g., exclusionary factors). The law is purpose driven. For instance, the ruling out of exclusionary factors (e.g., language, socioeconomic status, etc.) as being the primary cause of learner struggles is designed to minimize the overrepresentation of clusters of students within special education (e.g., low-income or ELs). Further, federal policy mandates that no single test be used to determine a disability, but rather requires that MSD be utilized when conducting a comprehensive evaluation. C-SEP™ adheres to each aspect of special education policy, requiring evaluators to understand the eligibility criteria for each disability category, to consider the impact language has on learning, to rule out exclusionary factors, to establish a pattern of underachievement, and to use MSD when making eligibility decisions, all within the confines of the law (Stephens et al., 2018).

In addition, IDEA (2004) stipulates that assessment and other evaluation materials be selected and administered so as not to be racially or culturally discriminatory. For instance, evaluators must ensure that a student's racial and cultural makeup closely matches or resembles those of the norm sample of any normed tests used. Among other issues, a lack of adequate norm representation may provide an unfair comparison and lead evaluators to over or underestimate a student's cognitive and academic abilities. To circumvent these challenges, practitioners are advised to thoroughly review the psychometric properties of the tests utilized to ensure they select the most appropriate instrument for that student. Other factors that must also be considered during an evaluation include, but are not limited to, being discriminatory based on race, gender, socio-economic status (SES), or parent education levels. We again caution that the use of inaccurate or biased assessment practices can lead to erroneous eligibility decisions and inappropriate placements. Only when considering these factors does the student receive a fair and valid assessment process.

IDEA (2004) likewise mandates that tests are administered in the child's native language, or other mode of communication, and in a form most likely to yield accurate information on what the child knows and can do academically, developmentally, and functionally, unless it is clearly not feasible to so provide or administer (34 CFR § 300.304; TEC 29.004(b)). Historically, failure to consider

language proficiency and its impact on academic learning led to an overidentification of ELs in special education (Ortiz, 2019). For instance, ELs with normal English learning developmental patterns have been misclassified as having a learning or intellectual disabilities without proper consideration of how language proficiency is foundational to accessing the academic curriculum. In C-SEP™, analysis of language proficiency and environmental factors should be followed by a cognitive assessment that is linguistically appropriate, which may require the assessment be conducted in English, the child's home language, bilingually, and/or through nonverbal modalities. To ensure proper interpretation of these assessment results, evaluators must then take the following into consideration outlined below in Box 7.

Box 7. Considerations when assessing EL Students

- (a) The child's exposure, formal and informal, to each language assessed.
- (b) Their opportunities to learn and practice English.
- (c) Their English language development support.
- (d) The duration of time the student has been learning formal English.

Most importantly, the examiner should appraise the student's academic strengths and weaknesses with peers from similar backgrounds.

Professional Judgment

IDEIA (2004) and subsequent regulations published in August of 2006 changed the way students suspected of having specific learning disabilities are evaluated and found eligible for special education services. In particular, the profession has moved away from using traditional methods for identifying SLD (e.g., simple discrepancy) and towards the use of response-to-intervention (RTI) and/or third method approaches (PSW models). This shift has increased the necessity for an evaluator to use their professional judgment as, for instance, statistical analysis alone is not the determining factor for eligibility (Schultz & Stephens, 2009). Reflecting this shift, C-SEP™ recognizes the importance of an evaluator using their professional judgment when not only interpreting data results and MSD, and ultimately when determining whether an SLD or other disability exists.

Often referred to as “clinical judgment,” professional judgment is defined as a reflective, self-corrective, purposeful thinking process which requires the professional to leverage content knowledge, context, evidence, and methods (Facione, Facione, & Giancarlo, 1997). According to Schalock and Luckasson (2005), professional judgment enhances the precision, accuracy, and integrity of the decisions made by the evaluator when it adheres to best practices. Today, the value of professional judgment is evident. As SLD identification has transitioned from a simple discrepancy method to the integration of RTI data, and closer towards a process of establishing a PSW, the importance of an evaluator’s professional judgment continues to grow. Box 8 outlines Schalock and Luckasson’s (2005) standards of professional judgment.

Box 8: Standards of Professional Judgment

Schalock & Luckasson (2005) outline 10 clinical judgment standards in the field of intellectual disability. We denote six of these characteristics which are most applicable to SLD evaluations.

- Employs research-based best practices in diagnosis, classification, and planning supports.
- Is based on the practitioner’s relevant training and experience in the field.
- Uses systematically collected and extensive data as a basis for making decisions and formulating recommendations.
- Is consistent with and augments professional standards and professional ethics.
- Employs critical thinking skills involving analysis, alignment, and synthesis.
- Aligns the functions of diagnosis, classification, and supports planning to the specific questions asked and the procedures that are used.

To enhance professional judgment, it is imperative that evaluators be highly trained in the areas of special education policy, theories of learning, instructional best practices, assessment theory and practice, as well as be skilled in the collection, analysis, and interpretation of various forms of data (Schultz & Stephens, 2009). Moreover, when assessing and interpreting results, evaluators should be knowledgeable of the instruments they are using and how to accurately interpret and report the data acquired. They must also have a foundational understanding of the variables that influence both cognitive and academic learning and must understand second language acquisition.

State special education guidance documents note the importance of evaluator’s use of professional judgment. Naturally, it may take a diagnostician several years of assessment experience

to become confident, but when professional judgment is supported by multiple data sources to demonstrate PSW (Schultz & Stephens, 2009), we expect sound eligibility determinations will be made.

We make this assumption because educational diagnosticians and other school-based evaluators are highly trained professionals with classroom teaching experience, curriculum knowledge, an understanding of learning processes, have knowledge of assessment practices, and knowledge of special education policy. Armed with their education and experience, diagnosticians possess insight that should be intelligently leveraged during the evaluation process. In this frame, we encourage evaluators to use their professional judgment by referencing their background and experience during every step of the evaluation process. This requires an evaluator to be confident in their professional judgment and avoid conventional pitfalls. Watkins (2009) identifies common errors in judgment that evaluators must be aware of throughout the evaluation process. Consult Table 5 for a list of common errors and their remedies.

Table 5. Common Errors in Judgment

Types of Errors	
Skills-Based Error: Caused by being inattentive or being distracted. This type of error commonly occurs when dealing with routine tasks (e.g., lapses in judgment or mistakes made during the scoring a test). Rule-Based Error: Caused by an oversimplification of data or processing rules when dealing with complex data as individuals try to expedite processes based upon past experience. However, this practice can ignore or discount certain factors and/or patterns (e.g., automatic misdiagnosis of a student based on a few test scores from one instrument). Knowledge-Based Error: Overconfidence or overreliance on personal knowledge when confronting a new situation or when known rules do not apply. This lack of resources and poor or incomplete knowledge (e.g., This is a novel situation, so it must be explained by x).	
Remedies	
Be on the lookout for bias and error.	Use reliable and valid tools.
Pursue continued education.	Recognize and acknowledge your limitations.
Use comprehensive checklists and guidelines	Use a scientific approach

(Watkins, 2009)

An individual's professional judgment is only as good as his/her knowledge and experience in relation to the problems encountered. It can likewise be hindered by Consequently, poor thinking, limited knowledge base, and a lack of experience will result in imprecise decision-making and outcomes (Schultz & Stephens, 2009). Such practices by educational diagnosticians can be detrimental to the students for which educational decisions are being made. Consequently, it is imperative that evaluators and other educational personnel be highly trained in the areas of educational law, instructional practices, assessment practices, and critical thinking skills.

To enhance their professional judgment, we recommend evaluators join a professional organization, subscribe to professional journals (and read them), and take every opportunity to attend conferences dealing with evaluation. We also highly encourage evaluators to attend periodic training to stay abreast of changes in the law, evaluation techniques, and testing instruments. One final method of improving professional judgment is to seek mentorship and/or use a peer review process whereby evaluators can support one another.

Practices

The last P refers to best practices. This concept describes an evaluator's leveraging of scientifically proven and effective practices, as well as observing the professional standards that govern and regulate the field. By law, evaluations must be "individualized" and of a typology that will yield the most accurate information. This means that in addition to using norm-referenced tests, other strategies such as dynamic assessment and curriculum-based measurement practices, task demand analysis, permanent product analysis, error analysis, as well as conducting interviews and observations should likewise be utilized (Kwiatek & Schultz, 2014). See Box 9 for a definition of error analysis and Box 10 for a definition of dynamic assessment. This combination should not only be used to make decisions of eligibility when using C-SEP™, but should likewise be leveraged to make recommendations regarding instructional practices designed to address each student's individualized learning deficits. Since C-SEP™ is an assessment process that requires the collection and analysis of multiple sources of data, evaluators will inevitably need to have a "variety of assessment tools and strategies" at their disposal.

Box 9. Error Analysis

Error analysis is a technique used to review a student's work to identify patterns of errors. The technique provides rich information regarding a student's strengths and weaknesses. Moreover, it can help the evaluator observe the types of errors the student makes, thus providing insight into their needs.

Another aspect of adhering to best practice refers to adherence to the standards and protocols that are put into place by the model of assessment an evaluator selects to use. If evaluators are following an RTI model, for instance, then we encourage RTI protocol to be followed. By comparison, if an evaluator is using C-SEP™, then they should adhere to its protocol and standards. While an evaluator might be tempted to randomly select practices out of convenience, this lack of fidelity may compromise the value of the model being used. In turn, lack of fidelity is likely to undermine the results acquired using a particular model.

Box 10. Dynamic Assessment

Dynamic Assessment (DA) is defined here as a collection of assessment practices in which instruction is provided during the assessment to understand the process of learning instead of just the product of assessment that is obtained from static assessments (e.g., norm-referenced tests, permanent products; Cho et al., 2020). DA is usually used in conjunction with curriculum-based measures (CBM) in a test-teach-test sequence to ascertain a student's ability to learn (Barrera, 2006). DA data should be included as part of the evaluation and can take place prior to the evaluation, such as in an RTI setting or during the evaluation. A similarly useful DA practice is the "testing-the limits" assessment technique in which an examiner explores processes and gains further insight into learning by changing the modality, providing brief teaching, or in the case of EL, adjusting the language demands after standard administration of a norm-referenced tests (Gonzalez et al., 2013). These latter techniques should be done purposely, yet they should not alter the standard scores obtained. Instead, such DA methods are useful for obtaining insight and understanding process.

Finally, our profession is overseen by numerous agencies that establish professional standards and best practices within the field. Their contributions run the gamut from setting professional qualification standards to establishing ethical responsibilities. At the national level, one thinks of organizations such as the American Psychological Association or the National Association of Psychologists. Similar professional organizations that support and grow evaluator's professional

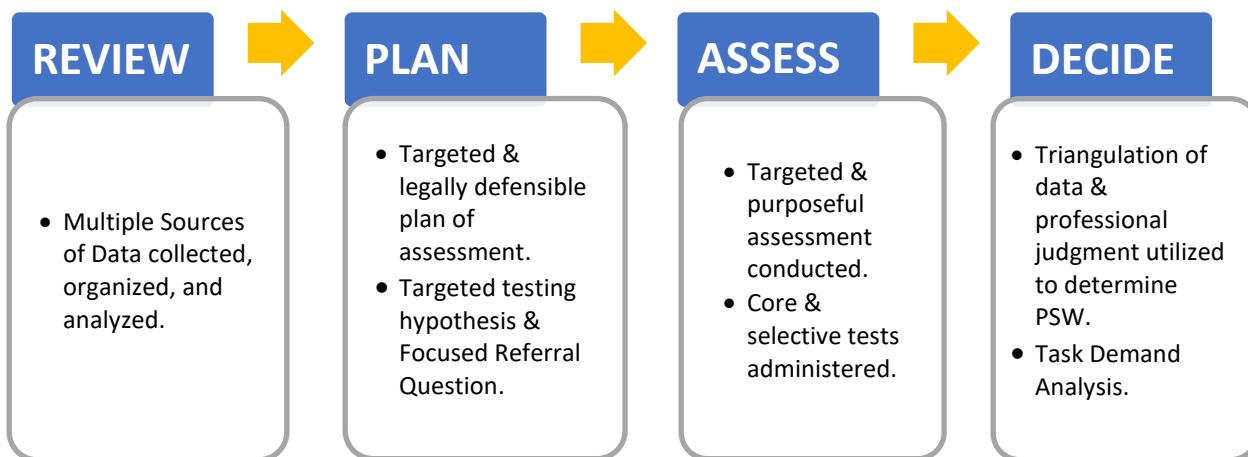
judgment are also organized at the state level. We encourage evaluators to become members of such associations, browse their websites regularly, speak with their colleagues and other professionals who are knowledgeable. Most importantly, adhere to their standards and best practices.

After this examination of the four Ps, we now turn to introducing the four primary steps of our model.

Critical Steps of C-SEP™

When utilizing C-SEP™, an evaluator follows a series of four consecutive steps: Review, Plan, Assess, and Decide. Each step serves as the precursor for the next, mapping the evaluator's progress through the gathering of MSD and to building a comprehensive understanding of the learner. At this point, we will limit our introduction to providing the reader with a general overview of how the C-SEP™ method progresses using Figure 2. We do this since the next four chapters of this manual will clearly define and flesh out each step.

Figure 2. Steps of C-SEP™



Page Intentionally
Left Blank

Review

Federal (34 CFR § 300.304(b)(1)) law dictates when determining whether a student meets SLD criteria, “a variety of assessment tools and strategies” must be considered. Consequently, collection and analysis of MSD through a variety of methods is a core practice of C-SEP™. We believe that information gathered at each stage of the evaluation process provides insight into the student’s strengths and weaknesses (Schrank, Stephens-Pisecco, & Schultz, 2017). Moreover, all data is weighed equally and interpreted in relation to other data sources. When collecting, merging, and examining the data, evaluators must use their professional judgment and data analysis skills to guide the assessment process and make sound eligibility decisions.

Box 11. Examples of MSD

RTI Interventions & Progress Monitoring	Language History
In-Class Tests	Student Interview
Grades Over Time	Educational Records Review
Norm or Criterion Referenced Tests	Observation in Classroom in Area of Struggle
Statewide Assessment	Observation in Classroom in Area of Strength
Teacher Input	Testing Observation
Parent Information	Discipline/Behavior Data
Work Samples	Special Ed Records (if applicable)
District Benchmarks	Attendance Records
Vision/Hearing Screening	

The first step, the data collection stage, is *Review* (Stephens-Pisecco et al., 2018). During this step, data is collected (or taken from the referral packet), organized, and analyzed with the aim of conducting an individualized assessment and testing plan. If the existing data proves insufficient for making an informed decision about the eligibility of the student, then the evaluator will need to collect additional information using various methodologies such as interviews with the student or parent, observations, or retrieve grades across multiple school years, benchmark scores, and so forth. The actions of *Review* are essential to developing a testing strategy and a testing hypothesis if formal data must be collected to determine student eligibility. We recommend that data collected during the *Review* process be organized using the “*Multiple Sources of Data Worksheet*” (See Resource Links).

See Box 11 for examples of MSD that is valuable at this stage. While processing and analyzing MSD, at any stage, evaluators should do the following:

- Review the referral concerns.
- Review educational records.
- Review instruction and instructional response to scientific, research-based intervention in the area(s) of suspected disability.
- Establish failure to meet age- or grade-level state standards in one of eight areas when provided appropriate instruction.
- Assess language and any implications language has on learning.
- Review existing data associated with the reason for referral and any data associated with that concern that is available.
- Conduct a preliminary rule-out of exclusionary factors.
- Start narrowing the focus of data-collection and analysis onto the areas of concern by mapping the student's individual pattern of academic strengths and weaknesses.

When using C-SEP™, data collection should focus on filling gaps in the existing data and answering outstanding questions associated with the pattern(s) identified. Efficiency is achieved by allowing the reason for the referral and the area(s) of concern to guide the assessment process.

Among the information gathered and analyzed at this time, the evaluator should conduct a preliminary overview of exclusionary factors, the student's educational background and experiences, and any existing evaluation/assessment data (from previous evaluations). C-SEP™ also requires that expressive and receptive language be informally or formally assessed for all students being considered for special education services. Student records should likewise be reviewed as they provide a historical precedent for the referral question. Identifying and examining documented patterns of learning and behavioral difficulties through a variety of resources over time can help establish ecological validity of the current assessment (Hass & Leung, 2021). Examiners should equally consider resources like the Home Language Survey and other parental feedback forms, which help LEAs make important determinations about a student's behavior, language, and cultural background (U.S. Department of Education, Office of English Language Acquisition, 2017). Such information will likewise be essential if formal testing is required.

When working with MSD, the following criteria should be met:

- Academic underachievement must be determined using multiple sources. Historical data (e.g., RTI, state testing, work samples, grades over time, etc.) are collected and analyzed to identify

weaknesses, validate the referral question, and establish a preliminary pattern of strengths and weaknesses. The data is then Reviewed and used to Plan formal evaluation.

- Standard scores obtained from norm-referenced testing should be used to understand the relationship between cognitive and language constructs (how they influence each other). Standard scores should, however, never be used as the sole determiner of a PSW.
- Statistical analysis should be conducted using norms and software/tables from the test publisher. The data can be viewed through multiple lenses. For instance, data collected from various sources (e.g., district benchmarks, unit tests, RTI data, grades, teacher, and parent reports, etc.) can be analyzed along with standardized testing results. The data should be interpreted beyond the standard scores (e.g., What does the relative proficiency index [RPI] indicate about the student's ability to complete the work? Are RPI results reflective of what the RTI data and teacher reported? What does language proficiency indicate about the student's ability to complete academics?). All decisions must be supported by multiple forms of data collected over time (further establishing ecological validity).

When evaluating data, whether pre-existing or collected during the assessment, be certain to pay close attention to trend lines and the number of data points. Never conjecture or generalize based on a limited number of data points, but instead collect additional data to obtain a comprehensive picture of patterns of strengths and weaknesses.

As the *Review* stage advances, the evaluator should acquire a more complex understanding of the student's academic struggles. With this greater understanding from the review of existing MSD, a formal and targeted testing plan, if needed, can be developed. This is accomplished by:

- Using the data-driven insight acquired during the *Review* to clarify the referral question that will guide the creation of a testing hypothesis and focused referral question (FRQ).
- Building recommendations associated with the data available hitherto.

The *Review* stage is critical to forming a comprehensive understanding of the student being evaluated, yet acquiring this knowledge efficiently and effectively requires a collective effort. When working through the referral process, it is essential that the evaluator collaborates closely with parents, teaching staff, caregivers, and related services providers to gather, synthesize, and consider MSD thoroughly and effectively. In this manner, the evaluator and other team members streamline and strengthen the assessment process, whereby team members and stakeholders become more knowledgeable of the student and the referral concern(s). Culturally responsive rapport building and interviews with stakeholders are foundational to team cohesiveness as well as for clearly defining and adequately targeting the assessment to a student's individual needs.

As the *Review* stage concludes, the evaluator will know whether additional information needs to be collected. If do, the evaluator can use the existing data and early findings to determine what formal testing is necessary to answer the focused referral question and to make a data-driven determination of eligibility. Once all the existing data has been reviewed during this stage, the evaluator will have to determine:

- Which questions, if any, are left unanswered?
- Which areas of concern require more investigation and data?
- What, if any, additional data is needed to better understand the learner and plan a targeted assessment?

Answers to these questions and others will guide the development of a targeted assessment and testing plans.

Response-to-Intervention (RTI) and C-SEP™

Regardless of the identification model an evaluator or school district uses to identify SLD, by statute, academic intervention, and supplemental instruction (e.g., tutoring, special classes, etc.) must be attempted to remediate the academic difficulty (Schultz & Stephens, 2017). Hence, RTI in some form is used in most schools across the country. In fact, the re-authorizing of No Child Left Behind (NCLB), now titled Every Student Succeeds Act (ESSA), contains language that allows funding to be used for these efforts (ESSA, 2017). Due to its popularity, evaluators need to obtain ongoing/past progress monitoring assessment data and use this information to help make instructional decisions and determine whether the child is adequately responding to intervention (Kwiatek & Schultz, 2014). Under the C-SEP™ framework, this data will initially be explored during the Review stage.

Accordingly, Schultz and Stephens (2017) call for the data collected during the RTI process (or any other instructional activity which occurs prior to referral) to be incorporated into each evaluation. This formative and archival data is nonetheless but one of those that can be acquired from a variety of assessment tools and strategies. That said, RTI data provides valuable information concerning the student's academic strengths and weaknesses and is valuable for focusing an evaluation. For these reasons and others, C-SEP™ advocates RTI, in addition to other MSD, be used to establish a pattern of strengths and weaknesses and determine eligibility.

Examples of data available from the RTI process include:

- Curriculum-based measurement (CBM) and progress monitoring (PM) charts,
- Task demand analysis,
- An analysis of the types of interventions implemented within RTI,
- Work samples,
- Benchmark scores,
- State testing results,
- Observation data,
- Parent information,
- Student interview.

While it is ideal to include RTI data into an individualized assessment, there will be times when RTI may be scarce or even unavailable. This can partially be explained by the fact that schools and districts implement RTI differently, so RTI data might not be available at times. This reality is why C-SEP™ requires evaluators draw on other MSD.

Preliminary Assessment of Exclusionary Factors

Prior to referring and/or finding a student eligible for special education services under the category of Specific Learning Disability (SLD), the Individuals with Disabilities Education Act (IDEA, 2004; 2006) requires educators and assessment personnel consider and rule-out a set of exclusionary factors as being the reason the student is struggling academically (Stephens et al., 2013). Specifically, IDEA requires that educational personnel rule out the **primary cause** of academic weakness to be the result of one or more exclusionary factors. See Box 12 for more information.

Box 12. Exclusionary Factors

According to federal law (34 CFR § 300.309), **Exclusionary Factors** should be considered prior to and as part of an individualized and comprehensive evaluation. To ensure that underachievement in a student suspected of having a specific learning disability is not due to lack of appropriate instruction in reading or mathematics, the following information must be collected:

- data that demonstrates the student was provided appropriate instruction in reading and/or mathematics within general education settings and delivered by qualified instructors; and,
- data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal evaluation of student progress during instruction.

Codifying the requirement to consider and eliminate exclusionary factors in an evaluation, IDEA (2006) determines that SLD is:

...The child exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, State-approved grade-level standards, or intellectual development, that is determined by the group to be relevant to the identification of a specific learning disability, using appropriate assessments, consistent with §§ 300.304 and 300.305; and (3) The group determines that its findings under paragraphs (a)(1) and (2) are not primarily the result of -

- (i) A visual, hearing, or motor disability;*
- (ii) Mental retardation;*
- (iii) Emotional disturbance;*
- (iv) Cultural factors;*
- (v) Environmental or economic disadvantage; or*
- (vi) Limited English proficiency.*

(b) To ensure that underachievement in a child suspected of having a specific learning disability is not due to lack of appropriate instruction in reading or math, the group must consider, as part of the evaluation described in §§ 300.304 through 300.306.

As they consider and rule out each Exclusionary Factor as being the primary cause of academic struggle, evaluators should support their rule out with multiple sources of data. Should certain data be unavailable in the referral packet needed to rule out a specific exclusionary factor, the evaluator will have to collect and document this data through multiple sources. Within the C-SEP™ framework, Exclusionary Factors undergo a preliminary rule out during the *Review* stage but are again revisited and finally ruled out during the *Decide* stage once all the data has been collected. The consideration and ruling out process should be thoroughly documented and supported by specifically referencing the data available (34 CFR § 300.311). We recommend using the Exclusionary Factors Checklist (Stephens & Moon, 2020), and other sources to document your consideration and rule out.

For consideration criteria and documentation of exclusionary factors, see Resource Links: *Exclusionary Factors Checklist* (Stephens & Moon, 2014; adapted from Stephens et al., 2013)

Multiple Sources of Data (MSD)

As has been repeated throughout this manual, one central tenant of C-SEP™ is the collection, organization, integration, and interpretation of multiple sources of data at each step of the evaluation process (*Review, Plan, Assess, Decide*). Multiple sources of data are vital because they help construct

a comprehensive picture of a student’s academic history, establish underachievement, identify the initial emergence of the PSW, support and clarify the referral question, help guide the assessment and testing plans (e.g., determine which core and selective tests are needed), rule out exclusionary factors as being the primary cause of academic struggle, clarify whether additional data is needed to determine SLD, and ultimately, and are essential for determining whether a child is suspected of having an SLD according to federal law. For a sample of MSD according to RIOT, see Table 6.

Table 6. The RIOT Model and MSD (Leung, 1993; Wright, 2010)

RIOT Component	Explanation
Review of the Records	The collection of past and current data, including, but not limited to student grades, attendance, state test results, work samples, interventions.
Interviews	Acquired from questionnaires, meetings, correspondence, telephone calls with teachers, parents, relatives of students, support staff, and other professionals.
Observations	Classroom observations provided by teachers, support staff, or acquired from the evaluator themselves. The objective is to observe and document the student performance in a classroom setting, centering on characteristics such as on-task behavior, academic skills, study habits and so forth.
Tests	These include scores and observational data acquired from formal testing, CBMs, or using dynamic assessment. This data is merged and considered in combination with the above.

Adapted from Leung (1993) and Wright (2010)

However, the question new evaluators often ask us is: “How much data is enough?” Neither federal nor state law offers clarity on the subject. We suggest that the evaluator should collect multiple sources of data, acquired over time, which provide a clear picture of each of a child’s strengths or weaknesses. For instance, passing Math grades over a three-year period, meeting expectations on state mathematics assessments during that same period, teacher feedback, and a parent interview are collectively good indicators that a hypothetical student does not have a weakness in math. Nonetheless, in the end, the evaluator will have to use their professional judgment to determine when they have reached the threshold that sufficiently demonstrates and explains the weakness.

Another valuable recommendation we can offer to evaluators collecting data is to keep RIOT in mind. This abbreviation points out certain data types that should not be overlooked.

Information Obtained Prior to Testing

Data collected **prior to** testing is crucial for establishing an educational need, ruling out exclusionary factors as a primary reason for academic struggles, and for establishing an initial understanding of the student’s individual PSW. Table 7 contains a non-exhaustive list of data that should be collected, organized, and analyzed during the **Review** step.

Table 7. MSD for Review Step

<i>Type of Data</i>	<i>Guiding Questions</i>
Student Background Information	What has been the student’s school experience thus far? When did the student begin struggling? Has the student struggled in school for many years? What types of interventions were tried? How did the student respond? Has the student been retained? Has the student had a previous referral (Speech or DNQ).
Work Samples	Were the work samples provided relevant to the suspected area of disability? Were the work samples truly work samples and not worksheets (e.g., multiple choice)? Did the work samples provided include areas of strengths and weaknesses? Were a variety of work samples collected?
Benchmark Testing	Does the student have a history of low performance on benchmarks throughout their academic experience? Benchmarks include District practice tests for state testing, CBM benchmarking, end of unit tests.
Progress Monitoring/RTI Data	How long was the student in RTI? What interventions were tried? How long were they implemented? Were the interventions implemented with integrity? How often was the student’s response assessed? What was the Rate of Improvement (ROI)?
Language Information	Is the student an EL? What language does the student speak at home, with parents? Siblings? What language does the student speak on the playground? What is the student’s dominant language? How long has the student been exposed to English? In what language are the tv shows or videos watched? How does language impact learning? Is language the primary reason for the student’s academic difficulty? What is the student’s proficiency level (WMLS-III, Logramos, or TELPAS)? What is the student’s

	CALP or proficiency levels? How are the holidays celebrated at home (acculturation)?
Attendance Records	Has the child missed many school days during the current school year? What was the student's attendance like during earlier years? Does the child have a history of tardiness to school or classes? If so, consider whether the student had an opportunity to learn the skills?
Report Cards	Do past report card(s) confirm a history of strengths and weaknesses in the suspected area(s) of disability?
State Testing Results	How has the student performed on state tests? Which subject areas has the student failed to meet standards? How many years has the student failed to meet standards? Which subject areas has the student met or exceeded standards? How did the student perform on the TPRI, or other district developed, state approved tests for K,1, 2?
Observational Data	Was the student observed in a classroom during instruction in an academic area of concern and in the classroom of academic strength? Was the student working independently or in a group setting? What was the student's behavior in relation to class peers? Describe the interaction between the student and the teacher.
Health Information	Has the student passed their vision and hearing screenings? If not, referral to an appropriate professional is necessary before testing can proceed. Does the student wear glasses/hearing aids? If so, they should be used during formal testing. Is there a history of health concerns (e.g., asthma, frequent absences from visiting the doctor, frequent ear infections, head injury, or premature birth)? Is the student taking medication or herbal supplements at home or at school? If so, for what reason?
Teacher & Parent Information	Are concerns expressed clearly (e.g., phonics, reading comprehension, multiplication) or in general terms (e.g., Cannot do anything.)? Has there been a reference made of a family history of educational difficulty that has not been clarified? If information is unclear, question teacher/parent for clarification.
Discipline/Behavior Records	Does the student have a history of inappropriate behavior that interferes with learning? Does the student have a behavior plan with documentation? Has the student recently experienced a traumatic event?

Some of the above data can be collected through interviews, structured or unstructured, with parents or guardians, students, teachers, and other educational or professional staff. Among these, interviews are especially useful for clarifying uncertainties or ambiguities in the data available. For this reason, we recommend conducting interviews before writing an FRQ, as the insight derived from

interviews may help evaluators narrow their focus. Similarly, observations are valuable for acquiring information on the student. Most states require at least one observation be conducted in the area of concern, while other states recommend using two or more. In addition, there are other opportunities available at later phases in the evaluation process where observations can and should be made. C-SEP™ and most test publishers, for instance, encourage evaluators to make careful observations of formal testing sessions, where taking note of student behaviors, problem solving strategies, and errors can provide further insight into a child's performance and behavior.

Plan

During the *Plan* stage, all data collected and organized prior to administering a formal assessment has been organized and analyzed in the Review Step. The existing data has also been used to establish a targeted assessment and testing plan. It is during the *Plan* stage that a working hypothesis centered on the referral concern(s) is fully developed. In association with the referral question, the evaluator must determine what additional data is needed to answer the referral question. Once identified, they will have to select the testing battery or batteries, or other collection methodology, best suited to the referral questions and student, before choosing the core tests to be administered. This information can then be used to write the focused referral question (FRQ) and *plan* the formal evaluation.

By way of illustration, while processing information during the first and second stages of a hypothetical SLD evaluation, the evaluator must determine if additional work samples or other missing pieces of data should be acquired (e.g., more RTI data may be needed). If so, this data must be collected and then merged with existing data, where it can be analyzed, and the combination used to create the testing hypothesis and focused referral question. After establishing an FRQ, the evaluator should then select the core tests that will be administered based on the data needed and the FRQ. To this end, the most appropriate test battery/batteries (e.g., language demands, socioeconomic background) will be those which meets the child's individual needs and that is designed to obtain the information required to answer the referral question based on the child's unique characteristics (e.g., age, language, culture).

As we examine this step in the C-SEP™ process in this chapter, we will:

- Define a Focused/Purposeful Evaluation.
- Provide guidance on establishing a Focused Referral Question (FRQ).
- Provide guidance for Selecting Core and Selective Tests.
- Outline limitations to Norm-Referenced Tests.
- Review the steps for the *Plan* stage.

A Focused/Purposeful Evaluation

C-SEP™ embraces evaluations that are efficient and focused. Our framework was created because of the traditional discrepancy SLD model’s requirement to administer at least two tests in every area of cognitive processing and in academics (Flanagan, Ortiz, & Alfonso, 2013). Such over-testing is no longer needed thanks to advancements in assessment research, such as Cattell-Horn-Carroll (CHC) theory, which have helped to improve our understanding of learning, testing, and how SLD is identified. See Box 13 for an overview of CHC. The fact is, contemporary testing instruments are more robust, comprehensive, and efficient than their predecessors (Schrack, Stephens-Pisecco, & Schultz, 2017). Revised test batteries are also ecologically valid and reliable, meaning that examiners must no longer administer multiple tests to measure the same construct. Contemporary instruments not only reduce the amount of testing necessary, but lend themselves to more focused, purposeful testing. This allows an evaluator to spend less time testing and more time interpreting the data and planning instruction.

Box 13. Cattell-Horn-Carroll (CHC)

The Cattell–Horn–Carroll (CHC) model currently identifies 16 broad abilities each consisting of multiple narrowly defined capacities (Schneider & McGrew, 2012). Of these abilities, six domain-independent general capacities are identified as fluid reasoning (*Gf*), short-term memory (*Gsm*), long-term storage and retrieval (*Glr*) (which includes learning efficiency [*Gl*] and retrieval fluency [*Gr*]), processing speed (*Gs*), reaction and decision speed (*Gt*), and psychomotor speed (*Gps*). In addition, four abilities address acquired knowledge: comprehension-knowledge (*Gc*), domain-specific knowledge (*Gkn*), reading and writing (*Grw*), and quantitative knowledge (*Gq*). Six further abilities concern domain-specific sensory and motor functions: visual processing (*Gv*), auditory processing (*Ga*) and olfactory (*Go*), tactile (*Gh*), kinesthetic (*Gk*), and psychomotor (*Gp*) abilities (van Aken et al., 2019, p. 1375-1376).

Fully aware of the advancements in instruments, C-SEP™ recommends a purposeful evaluation that requires the “core” cognitive, achievement, and language tests be administered as the foundation of any formal assessment. The “core” set of tests are those organized by test publishers that measure the broadest set of constructs of interest. The testing results acquired from the core are then carefully integrated with the other available MSD and considered. While analyzing the combined data, the evaluator must then decide whether “selective” testing is required in specific areas to

provide deeper insight into the student's needs and struggles. In this frame, C-SEP™ recommends that only the portions of the academic achievement assessment which assess the area(s) of suspected academic weakness be used to collect data to answer a referral question (e.g., collecting additional support). This practice is better explained using the two examples provided below:

Example 1:

A 4th grade student who has passed his Reading state tests, earns good grades in Reading, and has an educational history of progressing normally in Reading does not require an individual norm-referenced academic assessment in this area because we DO NOT suspect a disability in reading.

Example 2:

A student who did not pass the state Reading test and does not respond adequately to Reading intervention will require a Reading portion of the academic assessment. After confirming the student was provided adequate instruction and opportunities to learn, we need to investigate whether there is a cognitive, linguistic, or an alternate explanation as to why the student is struggling. In this case, administration of a norm-referenced academic assessment in Reading is necessary to gain a deeper understanding of the nature of the reading concern. However, since MSD from the same student shows no weaknesses in Mathematics, Math-centered tests from the achievement battery will not be administered.

NOTE: While the law only requires that **areas of suspected disability** be evaluated, some schools or districts may require all areas of cognitive and achievement be evaluated with a Full Individualized Initial Evaluation (FIIE). Please adhere to all relevant requirements! Nonetheless, C-SEP™'s selective testing practices and its analysis techniques of core and selective data are still applicable.

Testing Hypothesis & Focused Referral Question (FRQ)

Core to any comprehensive, individualized, and legally defensible evaluation is a focused referral. Unfortunately, however, many referral committees still operate under the false presumption that if a student is made to look low enough across the board, he will qualify as SLD under one of the available categories. This misconception of how SLD is determined helps to explain the vague and confusing reasons given for a referral being made. For instance, experienced evaluators understand it is common to receive referrals that generically make statements such as:

- He cannot read.

- She cannot do anything (all boxes checked).
- He cannot focus.
- Difficulty with reading and math.

However, when the reason for referral is unfocused, it lends to evaluators using a “standard protocol approach” to assess the student. This means evaluators have little choice than to assess in all areas to identify a potential problem area. In instances where the referral question is unfocused, evaluators risk having to test the student to the point that the student is worn out by over-testing, and evaluating a student at this magnitude is going to negatively impact the reliability of our findings.

To circumvent poor referral questions, we call on C-SEP™ users to establish a testing hypothesis whereupon a focused referral question can be designed. A testing hypothesis has multiple advantages. Firstly, it helps the evaluator determine which MSD is required to assess all areas of weakness. Secondly, it should eliminate the need to over-test students because preliminary strengths and weaknesses have been identified. Finally, designing a testing hypothesis and an FRQ allows the evaluator to be purposeful when selecting tests, centering the evaluation on suspected weaknesses and not on proven strengths. In an ideal world, the referral packets that land on an evaluator’s desk would be of a quality that would not require an FRQ to be written. We state this because the individual that refers the student for evaluation is more familiar with the child’s strengths and weaknesses than we are. Nevertheless, all things being equal, our experience has demonstrated that the time spent designing a hypothesis and FRQ pays dividends because it reduces the time required to assess a child’s known strengths.

Moreover, like any qualitative or quantitative research we engage in, we must establish a testing hypotheses statement to study the relation between two variables (e.g., level of student attendance with their academic performance; the amount of improvement and type/duration of RTI intervention). This hypothesis is usually a declarative sentence which interrelates variables. It is an educated guess as to the relationship between factors that serves as the basis of an experiment/assessment to determine whether a relationship holds true. A testing hypothesis should thus be formulated in a manner that describes the nature of the difficulty and allows for the development of an assessment plan to target and comprehensively assess all areas of suspected disability.

Once a hypothesis has been established, the evaluator can design (a) research question(s) (FRQ) to guide the data collection process. Research questions are a hypothesis asked in the form of a question. When writing the FRQ, it should be clear, focused, and provide a definite path for how testing will transpire. Equally important, all areas of suspected disability will need to be assessed (34 CFR § 300.304(c)(4)).

An example FRQ might read something like:

Are Jenny's struggles in math reasoning, specifically when asked to identify appropriate procedures necessary to solve the problem, the result of a deficit in one or more psychological processes (e.g., fluid reasoning, working memory) involved in math problem solving skills and thus consistent with the construct of SLD?

The example FRQ above shows how the evaluator has processed the existing data, ruling out other concerns (e.g., Reading and Writing) as well as other potential causes of the area of concern (e.g., language comprehension, attention, behavior). This hypothetical FRQ also incorporates precise areas of concern that helps identify the exact data that needs to be collected. A Testing Hypothesis Worksheet is a useful resource for aiding FRQ design. It was adapted from The Portland Public Schools PSW Manual (2015), and with the expressed written permission of the Portland Public Schools (See Resource Links).

Choosing Core and Selective Tests

Most contemporary cognitive and achievement tests include a “core” and “selective” or “supplemental” group of tests. Tests designated core are the most reliable and ecologically valid measures from a battery and are intended to be used as the foundation of formal testing. This structural arrangement is most evident in the Woodcock Johnson IV (WJ-IV; Schrank, Mather, & McGrew, 2014). Each battery of the WJ-IV (Cognitive, Achievement, and Oral Language) has a designated set of core tests. Nevertheless, other tests such as the WISC-V (Wechsler, 2014) are likewise organized around a broad set of measures that assesses many constructs of interest.

As an evaluator contemplates which core tests to administer, they should consult the examiner's manual for each battery being considered. Test publishers offer guidance on selecting “core” tests as well as for choosing “selective” or “supplemental” tests should they be required. The WJ-IV (Schrank, Mather, & McGrew, 2014), for instance, provides Selective Testing Tables and the

WISC-V (Wechsler, 2014) has options for ancillary and complementarity indexes. Evaluators should consult this guidance while keeping in mind the area of concern, the referral question, and the individual needs and characteristics of the student being evaluated.

!!REMINDER!! While contemporary assessments are reliable and ecologically valid, SLD should never be determined based on NRT standard scores alone. C-SEP™ users should instead integrate standard scores with multiple sources of data (e.g., RPI, RTI data, state testing, classroom work samples, report cards, attendance records, standard scores, etc.) when determining whether an SLD exists. Additionally, while we recommend a cut-score of 85 or 1 SD below the mean be used as a guide for assessment planning and interpretation, this is not a rigid threshold. The Standard Error of Measurement (SEM) and Confidence Bands should always be considered when interpreting a student's standard score.

Best Practices: Evaluators should always use their professional judgment that is influenced by support of multiple sources of data, federal and state policy, and publisher guidance when interpreting a student's strengths and weaknesses using NRTs.

Limitations of Individual Norm-Referenced Tests

As mentioned above, individualized norm-referenced assessments are only administered in specific areas of concern or when additional data is needed to make an eligibility decision and inform instruction. Once the formal testing data is collected, it should be considered in relation to other data available. Standard scores or a student's norm-referenced score report should never be the sole criteria used to establish a pattern of strengths and weaknesses. Such scores only provide a profile of the student's performance at a given time and are therefore singularly insufficient to demonstrate a pattern.

While advantageous, evaluators should likewise be aware that individual norm-referenced tests of achievement have limitations and these should be considered when selecting and interpreting results (Schrank, Stephens-Pisecco, & Schultz, 2017; Schultz & Stephens, 2015). We will list six limitations:

1. There is a lack of item density, which means scores are based on limited samples and test aspects of an area of concern with only a few items that differentiate age and grades.
2. Norm-referenced tests are not aligned with state curriculums and are not designed to assess state-approved grade-level standards which are a determinant factor in identifying SLD according to state policy (McGrew, LaForte, & Schrank, 2014).
3. Standard scores only indicate the relative standing in a norm group, and they do not describe functioning.
4. The normative makeup of these instruments does not adequately represent all students.
5. Racial and/or cultural bias may exist.
6. Tests are not pure measures of skills, meaning that there is usually an overlap of skills that are being measured. *Note:* This is only a limitation if evaluators do not recognize this fact and consider it.

Despite these and other weaknesses, NRTs and the scores they generate continue to form the backbone of other popular SLD discrepancy models. However, C-SEP™ leverages all the data acquired from NRT to gain insight into a student's weaknesses and combines this with other MSD to make sound and informed SLD decisions. We have outlined reasons why C-SEP™ weighs these scores equally in comparison to other forms of data, but one that has not been covered is that NRT scores may overestimate or underestimate a construct.

To demonstrate this, let us take a moment to view the insight offered by the Wechsler Individual Achievement Test – 4th Edition technical manual (NCS Pearson, 2020). When referring a student with a possible math disability, the test publisher warns evaluators:

... an overall subtest or composite score may overestimate or underestimate his or her math ability. For this reason, performing a skill analysis is particularly important for evaluating a student's profile of strengths and weaknesses (NCS Pearson, 2020).

Here, the test publisher is warning users of one of the weaknesses of its instrument. This warning suggests that norm-referenced assessments need to be carefully selected and administered according to the methods and procedures spelled out by test publishers.

Combined, the principal value of norm-referenced achievement tests in the C-SEP™ model is to explore the statistical relationships with underlying cognitive correlates inherent to achievement. When achievement instruments are viewed as observable and measurable academic behavior samples, users can obtain a more comprehensive understanding of the student being assessed. The

co-norming of the WJ-IV achievement, cognitive, and oral language batteries, for instance, is one set of tests an evaluator can use to analyze these relationships and inform judgment. In the same fashion, the relationship between the Weschler WISC/WIAT-III KABC/KTEA are equally well-established instruments (Raiford, Drozdick, Zhang, & Zhou, 2015).

Assess

During the ASSESS stage, core and selective testing is administered and analyzed. In this frame, C-SEP™ utilizes a targeted, purposeful testing process. This means that tests are strategically chosen based on the referral question and are centered on collecting data in areas where additional testing data is needed to make an informed decision about SLD eligibility and to guide instructional planning. By law, all areas of suspected disability will need to be assessed using a variety of assessment tools (34 CFR § 300.304). When using the C-SEP™ model, core cognitive, achievement, and oral language tests, the most reliable and ecologically valid of a test battery, are administered. Once these have been administered and interpreted, the evaluator will have to determine if selective (or additional targeted) testing is needed. If selective testing is warranted, these tests will have to be purposefully chosen, administered, and then analyzed (e.g., scores considered in relation to other forms of data) to determine whether another round of selective testing is needed to answer the referral question.

The steps for *Assess* include:

- Measuring *core psychological processes*.
- Measuring *core language*.
- Measuring *core achievement*.
- Analyzing norm-referenced data in relation to other data sources and to determine if additional selective testing is required.
- Conduct task demand analysis to further understand the learner.
- Administering additional ***selective*** tests (when required) and to analyze the data to determine if additional diagnostic testing is needed.
- Observing the student in classroom/testing session.
- If questions remain, administer additional diagnostic testing in the area(s) of concern.

During the first phase of *Assess*, data collection continues based on the targeted testing plan. This means that the tests chosen in the previous *Plan* step are administered. Once the selected core tests of cognition, achievement, and language are administered, they must be analyzed. Test publisher's norms and its scoring platform should be used when scoring and analyzing test results. Task demand analysis should equally be conducted to help answer the referral question. Thereafter, this formal data should be integrated with and considered in relation to the other data sources available.

In addition to collecting scores from the instruments administered, evaluators should also collect performance data throughout the *Assess* step. Such data should be collected in two manners. Firstly, the evaluator is expected to collect performance data on the student during the testing session(s). This includes noting the behavior of the student on individual tasks and during the testing process. Secondly, these observations should be considered in conjunction with the documented performance and behaviors acquired from classroom observations in the area(s) of concern. The combination of observation data should likewise be merged and considered in relation to the MSD available.

If the results from core testing and observations during the initial testing phase of *Assess* do not yield the information needed, or if testing introduces new questions, the evaluator will be required to identify and administer selective testing to better understand areas of deficiency. Enough data must be collected to answer the referral question. In instances where the question cannot be answered, the evaluator will have to carefully choose which selective tests to administer following the same criteria outlined for the selection of core tests. After a first round of selective test data has been collected, it too should be analyzed, integrated, and considered in relation to the other data available. Once enough information has been acquired to answer the referral question, make an eligibility decision, and offer recommendations, testing can cease. However, when additional data is required even after core and selective testing has been administered, further select testing may be warranted.

Data Obtained During Formal Testing

Data collected during formal testing can provide valuable information about a student's cognitive processing, academic abilities, problem-solving skills, and their level of communication ability. Additionally, qualitative observational data collected during testing likewise provides rich insight into the just listed characteristics, as well as a student's own perceptions of his/her strengths and weaknesses, and their persistence and motivation when engaging with difficult tasks. As mentioned above, an evaluator should make copious observational notes regarding the student's behavior during test sessions (e.g., how long does it take for the student to respond, how does the student go about solving the problem, what is the student's demeanor during testing, does the student push through difficult tasks or give up easily?). Additionally, the evaluator should note any

errors, self-corrections, repeat requests, and so forth on the test record. All this data provides valuable insight into the student's skills and abilities, especially when specifically linked to each task. As always, this data should be merged with the other MSD to further clarify PSW and inform performance and student needs.

Data Obtained from the Score Report

A score report details a student's performance on core and selective tests. In addition to standard scores and percentile ranks, score reports provide an abundance of information regarding student performance. This data should be analyzed/viewed through multiple lenses including standard scores and percentile ranks, Relative Proficiency Index (RPI), and base rate and intra-individual performance (Schrang, Stephens-Pisecco, & Schultz, 2017). Consult the publisher's literature to accurately score and interpret findings at each level. Once scoring has been conducted and the results interpreted, this Information should be merged and viewed in relation to the other data available.

Standard Scores & Percentile Ranks

Traditionally, standard scores and percentile ranks have been the most widely used scores utilized from NRT. Standard scores and percentile ranks provide information about the individual's relative position among peers and can be useful for determining how common or rare the individual's ability level is within the general population. A standard score of 1 SD below the mean (or a percentile rank below 16) suggests the student performed in the lowest quartile of their national age or grade peer group. Performance at this level often indicates a need for further assessment to determine whether the lower performance can be generalized to a broad area of cognition or achievement, as presented in the cluster score (Schrang, Stephens-Pisecco, & Schultz, 2017).

Many test batteries outline how examiners should analyze and interpret the scores generated by their instruments. Users of KABC-II (Kaufman & Kaufman, 2018), for instance, can analyze the instrument's findings through multiple measures including:

- RPI
- Standard Scores
- Confidence Intervals

- Percentile Ranks
- Normative (Inter) Performance
- Intra-Individual Performance
- Base Rates

Regardless of which combination of score types are leveraged, all the NRT data collected as part of the score report should be carefully analyzed, merged, and viewed in relation to the other available data.

Criterion-Based Interpretations Relative Proficiency Index (RPI)

If evaluators are using the WJ-IV (Schrank, Mather, & McGrew, 2014), a Relative Proficiency Index (RPI) can be generated and used to interpret a student's ability to complete tasks by comparison to other peers in the same age norm. The RPI score is unique to WJ-IV and is a particularly useful measure when evaluating SLD. RPI scores express performance as a ratio scale that compares the individual's relative task proficiency with that of the typical, or median, individual of the same age. See Box 14 for additional information on the RPI.

Box 14. Relative Proficiency Index (RPI)

- Provides a criterion-referenced index of a person's proficiency or functionality.
- Compares the quality of performance on assessed skills and abilities to that of age or grade peers in the norming sample.
- Predicts level of success on similar tasks.
- Shows actual distance from average proficiency.
- Based on the W Difference score.
- Ranges from 0/90 to 100/90.

An RPI score report that flags an individual's test performance as "limited" or lower (or an RPI of about 67/90 or lower) can suggest a need to conduct additional selective testing to yield the related cluster score or provide additional comparative information. Similarly, a proficiency level of "advanced" (or RPI of about 98/90 or higher) might prompt an evaluator to administer additional testing in a domain to establish a broad area of individual strengths should more information in this

area be desired (Schrank, Stephens-Pisecco, & Schultz, 2017). As always, an evaluator should consult the instrument's publisher manuals when scoring and interpreting any test results.

Review of Base Rate and Intra-Individual Performance

The intra-cognitive, intra-oral language, and intra-achievement variation procedures provide an analysis of within-individual variability to derive a profile of relative strengths and weaknesses. Such analysis can be conducted with all major assessment instruments. When using the WJ-IV (Schrank, Mather, & McGrew, 2014), one of the most important scores to consider is the discrepancy percentile rank (discrepancy PR) because it provides information on the base rate of any observed discrepancy. If the difference between an obtained and predicted score is a negative number (e.g., -14 standard score points), then the discrepancy PR directly indicates the percentage of the population who would obtain a difference of that magnitude or greater (the base rate). The difference is sometimes referred to as critical value.

For example, a discrepancy PR of 1 on *COG Test 5: Phonological Processing* indicates that only 1% of the examinee's peer group had the same or larger difference score on this score comparison (the base rate). However, if the difference between the obtained and predicted score is a positive number (e.g., +23 SS points), then the discrepancy PR must be subtracted from 100 to obtain the base rate. For example, a discrepancy PR of 97 indicates that only 3% of the population would obtain a discrepancy between the predicted or obtained score of that magnitude or greater (i.e., $100 - 97 = 3$). In clinical practice, evaluators make note of a discrepancy PR lower than 10 (for negative differences) as an indicator for pursuing additional assessment in the domain area. When the PR is higher than ninety, it indicates areas of strength (Schrank, Stephens-Pisecco, & Schultz, 2017).

Base rate establishes clinical significance in the difference, telling the evaluator if the difference is meaningful. The lower the base rate, the more unusual the difference. Base rates vary by ability level (using FSIQ or GAI score ranges) not by age. Clinical meaningfulness means we can expect the student's performance to manifest in the classroom. Kaufman Assessment Battery for Children (KABC-II; Kaufman & Kaufman, 2018) requires a threshold of < 10% base rate to determine whether a difference is uncommonly large. Sattler (2016), by comparison, suggests differences between scores that occur in less than 15% of the normative sample should be judged as unusual (see the WISC-V

Manual; Wechsler, 2014). Nevertheless, each instrument will have specific instructions which should be followed. Always refer to a publisher's testing manuals when scoring and interpreting scores. See Table 8 for a brief explanation of critical values and base rates.

Table 8. Distinguishing Types of Weaknesses.

Critical Value	Base Rate
Establishes <i>statistical significance</i> of a difference between two scores.	Establishes the <i>clinical significance</i> of a difference between two scores.
Serves as a reference point to identify the likelihood that a difference between two scores occurred by chance.	Identifies the prevalence or frequency of a difference between two scores.
Differences between scores that reach or exceed the critical value are unlikely to have occurred by chance.	The lower the base rate, the more unusual the difference.
The level of statistical significance determines the critical value required to achieve a specified level of confidence that the difference is a true difference and not due to chance. (Typically defaulted to .05 = 95% confidence).	KABC-II requires <10% base rate to determine whether a difference is uncommonly large. Sattler (2016) suggests differences between scores that occur in less than 15% of the normative sample should be judged as unusual (WISC-V Manual).

Analysis of Task Demands

Tests are not pure measures of achievement, cognition, or language, because these measures include an overlapping of various skills and demands associated with input and output. Consequently, when a student performs poorly on a test, the evaluator cannot automatically associate the low performance to the inability to perform an overarching task (e.g., Applied Problems), because the low score might be due to other demands required to complete a given test. As a result, an evaluator must investigate the task demands required to complete each test. Task demands Analysis (TDA) is thus the process of identifying the inputs, processes, overlapping constructs, and output demands of a given test to expose patterns that:

- a) may help explain strengths and weaknesses in performance,
- b) gain a deeper understanding of the learner, and

c) help design individual instruction.

Task demands analysis is not only consistent with the definition of SLD, specifically “listen, think, and speak,” but it allows the examiner to maximize the test data acquired. Examiners can find descriptions of each test included in an instrument in the publisher's test manuals. As relationships between cognition, language, and achievement tests are established, an examiner may be increasingly able to explain “why” a student is behind.

Test publisher manuals often feature tables that offer a clear breakdown of inputs, processes, and outputs associated with each task and the abilities assessed. For instance, if we consider a Reading Comprehension (RC) test, we recognize that it involves written text as input and verbal expression as output. The processes used during the test include decoding words, language comprehension, working memory, and verbal reasoning. If a student receives a low standard score in RC, an examiner can examine the processes to determine what may have contributed to the low score. Below are three selected excerpts from reports that make use of TDA:

“Julie has strong processing speed in tasks like Letter-Pattern Matching, Sentence Reading Fluency, and Sentence Writing Fluency (SS 106-112). However, she struggles with math facts fluency (SS 66, RPI 3/90) indicating processing speed issues specific to math.”

“Tony performs well in tasks that require basic language skills but faces difficulties in tasks that demand advanced language abilities. For instance, he displayed an average performance in Applied Math (SS 98), Calculation (SS-92), and Visual Memory (SS 102), but struggled with Listening Comprehension (SS 78) and Reading Comprehension (SS 74).”

“Sophia struggled on timed tests requiring recalling previously learned information but did well on untimed tasks tests that require recalling previously learned information. Sophia may require extra time in the classroom to recall information as she demonstrated higher performance and reduced anxiety without the pressure of time constraints.”

NRT data is most useful when it is used to generalize about learning in the classroom. It is important to remember that each classroom task also has a demand attached to it. For example, mapping orthography to phonology is a brain process required to spell words correctly regardless of the setting it may occur (e.g., classroom or formal test environment). When comparing the task demands of a test with the task demands of other data (i.e., classroom data, chapter tests, work

samples) these will help provide clear and convincing evidence of the presence of an SLD. When using CSEP™ in this manner, tests help us explain underachievement.

To become proficient at TDA, the examiner must have the following knowledge and skills. They must have the knowledge required for each utilized test in the publisher manuals, specifically in the overview of each test, the theoretical foundation of the test design, as well as their interpretation and validity. Similarly, printing out tables and highlighting low scores is a useful strategy to make connections between tests. In this context, it is important to never “torture” the data but rather ensure that you can support your findings with the manual. The more an evaluator engages in this TDA, the more fluent they will become in making connections and generating useful insight. Overall, conditional or contextual analysis can assist the evaluator in generalizing their findings. In a formal test environment, individual testing conditions are more controlled to elicit optimal performance, while classrooms, by nature, are much more complex, with the latter experiencing multiple variables (social pressure, setting, interference, distractions) that can impact performance.

One way to approach TDA is by considering that SLD are manifested by a specific area of low achievement. These “manifestations” represent a “disorder in psychological processes involved in using language.” For example, the task and subsequent test of reading comprehension requires the input of written language, the processes of decoding, reading fluency, working memory, language comprehension, and simultaneous processing, and the output of either a written response or spoken response. The examiner can investigate a low reading comprehension score, or manifestation, by examining the processes above. A psychological process such as low working memory and poor decoding may explain the underachievement in reading comprehension. If the referral question involves reading comprehension, the examiner can be confident that a student displayed a disorder in the basic psychological process (working memory) involved in using language (reading decoding), which is manifesting itself as an imperfect ability regarding reading comprehension. The analysis is consistent with the definition of SLD and provides the team with specific areas that need to be addressed, namely through memory and decoding strategies.

Combined, the process of investigating TDA ensures evaluators tap into the full capability of the NRT instrument and help understand what the acquired standard scores indicate.

If a significant difference exists between or among the individual test scores with a factor or cluster, report performance on the narrow abilities and, using analysis of the task demands for each test (and other forms of data), attempt to explain the reason or reasons for the difference between scores. Also, consider how this information may alter interpretation or use of the factor/cluster score (Mather & Jaffe, 2016).

Stated differently, through task demands analysis, an evaluator can tease out an area of weakness. The process will likewise provide a better understanding and explanation of splits between scores/non-cohesiveness in data. Hence, evaluators are encouraged to compare task demands on one test to task demands on another to pinpoint strengths and weaknesses. As outlined above, each tests measures multiple constructs, and these are often shared across tests. Similarly, these same constructs are often measured by other data sources, so these alternative data sources can equally be compared to standard scores. Once strengths and weaknesses of each construct have been identified, the evaluator should then consider their implications in the classroom.

In closing this subsection, always attempt to identify the common abilities required on tests on which the student performed well and tests in which the student performed poorly. Examine the types of errors made on test items, determine whether a pattern of errors exists, and note and consider any strategies the examinee used. Through such comparisons, it is usually possible to determine the narrow abilities that appear strong and those that are weak. When making decisions about strengths and weaknesses, be sure to consider both Relative Proficiency Indexes (RPI) and peer comparison scores (Standard Scores). Once identified, consider their implications on classroom performance and which measures can be recommended to improve student outcomes.

Intellectual Development

When using the C-SEP™ strategy to identify Specific Learning Disability (SLD), it is essential to assess intellectual development. This helps to rule out other disabilities, consider if intellectual development can explain underachievement, and gain a deeper understanding of the learner. The term “intellectual development” appears twenty times in the 2006 Code of Federal Regulations in reference to “third method” approaches. For instance, 34 CFR Section 300.309(a)(2)(ii) states:

The child exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, State-approved grade-level standards, or intellectual development that is determined

by the group to be relevant to the identification of a specific learning disability, using appropriate assessments, consistent with §§ 300.304 and 300.305.

Understanding and measuring intellectual development is essential for identifying SLD. During the comment period for the IDEA Regulation, some requested that “cognitive ability” be used to define “intellectual development.” However, the department rejected this proposal and chose to stick with the term “intellectual development” (Schultz, Rutherford, & Cavitt, 2021). As the department did not provide a definition, the term “intellectual development” should be taken to have its ordinary meaning. While cognition is a part of intelligence, one's intellectual development is formed through the interaction between language and academic knowledge. The term "development" implies a process of growth and improvement, which is particularly rapid during childhood and adolescence, facilitated by schools. Therefore, the analogy of the "brain that is under construction" is fitting.

A more nuanced definition of “intellectual development” can be found in Sattler (2016) where he reports most experts in the fields of psychology and education generally agree that the essential elements of “intelligence” include abstract thinking or reasoning, problem-solving ability, capacity to store knowledge (including academic knowledge), memory, environmental adaptation, mental speed, and linguistic competence. These “processes” are interdependent and overlapping (Peterson et al., 2017). Due to its complexity, evaluators must carefully examine student performance at the task demands level to evaluate these aspects of intelligence. This development is influenced by the environmental conditions a child experiences.

Schultz, Rutherford, and Cavitt (2021) provide a practical definition of intellectual development comprised of the most essential elements of intelligence, which includes abstract thinking or reasoning, problem-solving ability, capacity to store knowledge (including academic knowledge), memory, environmental adaptation, mental speed, and linguistic competence. The work of Snyderman and Rothman (1987) and contemporary Cattell-Horn-Carroll (CHC; McGrew, 2009; McGrew, 2023) theory forms the basis of this definition. In simple terms, one’s intellectual development “develops” via the interaction of four variables: innate cognitive ability, language development, acquired academic skills, and interaction with the environment.

The role of language in SLD identification cannot be underestimated. It is a fundamental definition aspect and is closely intertwined with cognition. The definition explicitly links these two

constructs, stating that “psychological processes involved in using language” are central to SLD. Language and thinking are inextricably linked, and language is necessary for thinking. C-SEP™ is unique in its emphasis on linguistic aspects of SLD (e.g., listening, thinking, speaking), making it a more comprehensive model. While other models assess language through comprehension knowledge (Gc) measures, examiners typically do not consider language beyond this score.

When evaluating intellectual development, the most practical, valid, and reliable tools available are individualized norm-referenced tests (NRT). It is important to note that NRT scores should be integrated with other measures of student development when using the C-SEP™ method, while also considering the student’s “instructional response.” This allows the evaluator to determine if the student's development is typical (scores in a similar range) or atypical (significant variance). However, the presence of a variance does not necessarily indicate a disability. Professional judgment should be applied to analyze all available data and determine if there is an impact on learning in combination with these results.

When using the C-SEP™ method to detect Specific Learning Disabilities (SLD), an evaluation of intellectual development provides a more detailed and comprehensive understanding of learning, as opposed to examining discrepancies. This evaluation process is more aligned with the SLD definition and the relevant statute. The most significant aspects of SLD correspond with the concept of intellectual development and the subsequent steps taken to measure it. Psychological tests measure cognitive processes, language tests assess language involvement, while academic constructs like math calculation, basic reading, and written expression are analyzed through achievement tests and other relevant data. It is crucial to consider environmental factors and examine exclusionary or contributory factors to determine their impact. The testing practices of C-SEP™ and other assessment methodologies have been consistent with the SLD definition and statutory regulations.

Norm-Referenced Tests (NRTs)

NRTs have long been used for identifying educational disabilities and are *one tool* among a variety of assessments methods and strategies available to those conducting a full and individual evaluation (FIE). NRTs are usually used in SLD evaluations to explain the meaning behind the scores including variance, low scores, and how these constructs impact learning. This is accomplished by identifying some form of discrepancy (e.g., IQ-achievement discrepancies, dual-

discrepancy/consistency models, concordance- discordance) in SLD evaluations (Decker, Hale, Flanagan, 2013). Without question, NRTs are helpful for identifying and understanding a student's functioning and learning. When assessing using a PSW framework such as C-SEP™, discrepancies and standard scores from norm-reference testing (NRT) data only inform decision-making and professional judgment. NRT scores alone are not determinative.

Unfortunately, in addition to over-emphasizing the value of NRT scores, our experience has shown that evaluators often overlook certain information provided by NRT and norms, and by doing so, they deny themselves the broader and deeper insight many instruments provide. As part of a more in-depth analysis, C-SEP™ recommends evaluators conduct a task demands analysis for each set of NRT scores to exploit the fullest potential of the data. In this manner, it is possible to acquire insight into a student's performance on the constructs being evaluated. At the same time, scores should be considered, calculated, and analyzed according to test publisher guidance.

While NRTs provide insightful data when exploited to their fullest, NRTs equally have their limitations. We have published journal articles about these weaknesses in the context of SLD evaluation (Schrank, Stephens-Pisecco, & Schultz, 2017; Schultz & Stephens, 2015), yet our objective has never been to reject the use of NRTs, but rather to inform evaluators of NRT weaknesses and to get them to rethink how NRTs are utilized. In our training and research, we continuously implore diagnosticians to educate themselves about the strengths and weaknesses of NRT in general, and the test batteries they are using. Without question, no instrument is perfect. Nevertheless, we believe that NRTs are reliable and valid when they are used for their intended purpose, when they are used with their intended target population, when they are administered according to strict protocol, and are scored, analyzed, interpreted, using testing norms, and reported according to the test publishers' guidance.

That said, we would like to take a moment to address five common assumptions concerning NRTs as a means of explaining how tests can responsibly be utilized (Schultz, Rutherford, & Cavitt, 2021). These are assumptions we encourage evaluators to be fully aware of when choosing, administering, and interpreting NRT results.

Assumption #1: *Tests of Cognition (e.g., WISC-V, WJ-IV COG), Tests of Achievement (e.g., KTEA-3, WIAT-4), Tests of Language (e.g., WJ-IV Oral Language, CELF-5) all measure aspects of intellectual development.*

The assumption is accurate since all these tests measure aspects of intellectual development. To illustrate, the Applied Problems of the Woodcock Johnson IV Tests of Achievement (WJ-IV-ACH; Schrank, Mather, & McGrew, 2014) requires the “achievement” area of math calculations and problem solving, construction of mental models via language comprehension, and the cognitive process of quantitative reasoning (McGrew, LaForte, & Schrank, 2014). Similarly, the Understanding Spoken Paragraphs subtest on the Clinical Evaluation of Language Fundamentals (CELF-5; Wiig, Semel, & Secord, 2013), in addition to measuring receptive language, also measures aspects of cognitive processing, specifically verbal working memory, verbal reasoning and the achievement areas of listening comprehension, vocabulary, and prior knowledge (Wiig, Semel, Secord, 2013). One final example provided herein is The Similarities subtest in the Weschler Intelligence Scales for Children–Fifth Edition (WISC-V; Wechsler, 2014), which spans cognition and achievement. Not only does the WISC-V subtest measure cognition, but also requires skills related to achievement such as word knowledge, word meaning (Vocabulary), and the language skill of verbal expression (Weschler, 2014).

Assumption # 2: *Standard Scores are Not Equivalent to Functioning.*

For decades, concerns have been raised about using standard scores for identifying SLD (Katz & Slomka, 2000; Phillips & Clarizio, 1988). There are measurement issues with discrepancy models (Gresham & Vellutino, 2010; Van den Broeck, 2002), and test publishers recognize the limitations of standard score interpretation for students with disabilities. A fundamental and common interpretive error made by evaluators is the interpreting of standard scores as equivalent to functioning or performance. This misunderstanding can lead an evaluator to make improper generalizations and inaccurate educational decisions about the student. For example, a standard score of 90 on a memory test could be misconstrued to mean the student has “average” functioning in memory. However, this score only represents that individual’s relative position or “place” in a line (ordinal placement) in the context of reading (Jaffe, 2009). One manner of averting this error is by conducting a task demands

analysis, whereby an average standard score may be found to be hiding difficulties in certain skills categories.

The Wechsler Individual Achievement Test (WIAT-III; 2009) and WIAT-4 (NCS Pearson, 2020) manual provides useful guidance, stating:

Students with mathematics disabilities typically present a profile of math strengths and weaknesses and may perform considerably below average in some areas and average or better in other areas. For a student with this profile, an overall subtest or composite score may overestimate or underestimate his or her math ability. For this reason, performing a skills analysis is particularly important for evaluating a student's profile of strengths and weaknesses on the Math Problem Solving and Numerical Operations subtest (WIAT-III, 2009).

This limitation is not unique to the WIAT or mathematics, but rather it is often due to item density and standard deviation of the test. Identification models that rely on standard scores may inadvertently misinterpret a construct as being intact when not and vice versa.

To control for this limitation, we must go beyond the score and integrate other data to support a score and be more diagnostically precise. In other words, "to go beyond the score." Evaluators should be cautious when interpreting standard scores. To comprehensively understand a student's strengths and weaknesses, they should consider performance beyond the standard score, utilizing other scores (e.g., RPI), error analysis, task demands, and testing observations. Finally, evaluators should thoroughly read publisher's test manuals to understand what scores mean and how these should be accurately interpreted and reported.

Assumption # 3: *Norm Referenced Tests are Academic Behavior Samples*

This is true. Lewis, Hatton, and Maynard (2017) define behavior as observable, measurable student action. According to this conceptualization, behavior is not limited to social behavior but includes academic behavior as well. For example, total words written during a test is a measurable student action as are the number of correct responses they give. An evaluator can use NRTs to obtain deeper insight into a student on several dimensions when NRTs are viewed as a collection of academic behavior samples, in addition to being a method for making statistical comparisons. This practice is further considered in Assumption #4.

Assumption #4. *Conditional and Task Analysis are as Important as Scores, if Not More.*

Conditional analysis and task demand analysis are accepted educational practices (Stephens, Olvera, & Schultz, 2022) and have been used to understand students with disabilities for decades. For instance, Functional Behavioral Analysis (FBA), recommends examining and considering both environmental and behavioral data. School psychologists often use systematic direct observations to quantify student behavior and identify predictable patterns (Dufrene, Kazmerski, & Labrot, 2017). Within this context, individualized NRTs, by definition, are also systematic as they too require direct observation. When conditional analysis is applied to both the testing session and teaching setting, predictable patterns of academic behavior can be identified and remediated. For instance, scrutiny of the classroom and testing session may demonstrate that a student has difficulty understanding oral instructions. These observations can then be merged with other data collected, including classroom observations, or teacher and parental feedback, to confirm this as a pattern of weakness.

Task demand analysis is similarly valuable for mapping and understanding a student's strengths and weaknesses. To this end, tasks can be analyzed at various levels, including cognitive, linguistic, input (e.g., visual, auditory), and output (e.g., written response, spoken). Once again, classroom and test environments give evaluators the ideal opportunity to observe student performance related to task demands. The overall value of this practice is evident in the following interpretative statement associated with a task demands analysis during a test session:

Victoria can accurately recall verbal information when it is presented in a test session when distractions are reduced, and she is looking directly at the examiner.

This hypothetical observation statement not only identifies the task and the student's challenge, but provides insight into remediation, all of which the evaluator has learned during a test session.

Assumption #5. *Norm-referenced achievement tests are not valid measures of curriculum or grade level standards.*

This is true. Assessment tests (e.g., WIAT-III, WRAT-5, KTEA-2; WJ-IV ACH) are measures of intellectual development and not curriculum or grade level standards and have not been validated for this purpose. Fact is, it is impossible to capture the depth and breadth of a state's curriculum using an NRT due to the limited number of items required to differentiate performance between ages and grades. It would equally be challenging for test publishers to produce instruments that reflect the

innumerable curricula found across the United States. Consequently, the best way to measure achievement by age or grade level standards is with Curriculum-Based Assessments (CBA) or Curriculum-Based Measurements (CBM).

NRTs of achievement, nonetheless, are useful for identifying and understanding students with SLD due to the cognitive complexity of these measures and their relationship to academic domains. Achievement tests are sometimes referred to as a “cognitive process paired with an academic skill.” Test manuals explicitly denote this in their discussion on validity descriptions. We will examine a few statements extracted from popular test manuals to underscore the value of pairing cognitive processes with academic skills:

Wechsler Individual Achievement Test, Fourth Edition (WIAT-4): “For the WIAT-4, validity evidence on response process should provide support that the student engages in the expected cognitive process when responding to subtest items (academic skills)” (NCS Pearson, 2020, p. 40).

Kaufman Test of Educational Achievement, Third Edition (KTEA-3): “The responses should provide support that the examinee engages in the expected cognitive processes when responding to test items (academic skills)” (Kaufman & Kaufman, 2014, p. 50).

Woodcock Johnson IV Tests of Achievement: “The WJ-IV ACH contains test that tap two other identified cognitive abilities: quantitative knowledge (*Gq*) and reading- writing ability (*Grw*). The WJ-IV ACH also includes additional measures of comprehension-knowledge (*Gc*), long-term retrieval (*Gltr*), and auditory processing (*Ga*). Because most achievement tests require the integration of multiple cognitive abilities, information about processing can be obtained by a skilled examiner” (Schrack, Mather, & McGrew, 2014, p. 6).

Wide Range Achievement Test, Fifth Edition (WRAT 5): “The responses should provide support that the examinee engages in the expected cognitive processes when responding to test items” (Wilkison & Robertson, 2017, p. 46).

When “achievement” tests are used in the manner that they have been validated (i.e., academic skill paired with a cognitive process), discrepancy between cognition and achievement can be identified. Moreover, deep analysis and an improved understanding of the student can be obtained if the evaluator considers the cognitive complexity of each test or task, according to the design of each Instrument.

To demonstrate how an achievement test might be leveraged in this manner, let us consider the following impact statement:

Sophia's processing speed can be described as average when completing simple tasks such as discriminating between letter patterns. She struggles with tasks that require processing speed and reading (specifically oral reading), but she demonstrates average ability when using processing speed to quickly solve math problems.

This Impact Statement perfectly depicts how an evaluator used task analysis and the blending of cognitive and academic skills to analyze and explain Sophia's strengths and weaknesses. It clearly explains that Sophia's processing speed is average when solving math, but she simultaneously shows weakness in the same cognitive area when discriminating letter patterns. When these findings are confirmed through additional pieces of data (e.g., teacher feedback, classroom observations), then the evaluator is able to demonstrate the existence of an SLD.

Page Intentionally
Left Blank

Decide

During the *Decide* step, all available data is merged and analyzed in unison to determine whether a PSW exists. Professional judgment is a key throughout the *Decide* stage. During our overview of this stage, this chapter will examine:

- Using integrated data analysis procedures to identify PSWs.
- Conducting systematic data analysis using the chain of evidence technique.
- Using pattern seeking techniques.
- Revisit the law and professional judgment.
- Guidelines for employing professional judgment.
- Recalling the steps for the *Decide* stage.

There are two primary tasks in the *Decide* stage: a) to integrate, merge, organize, and analyze data collected; and b) to apply the data to PSW policy. This step begins after all the required formal data has been collected, and in-depth analysis has occurred. As emphasized above, test publisher's norms and scoring platforms should be used to score and analyze formal test results. To understand scoring differences, a Task Demand Analysis should be used to reconcile these differences. Discrepancy analysis should equally be utilized to show variance and to identify and support patterns that emerge from the data. An analysis of the NRT should then be integrated into other data analysis. NRT results should never be used alone when making an eligibility decision.

Once analyzed independently, all norm-referenced, standard scores should be merged with the multiple forms data collected "prior to and as part of" the evaluation process so that it can be analyzed and triangulated to establish PSW and determine if an SLD exists. In this frame, triangulation is a comparison of data (i.e., informal data, non-standardized testing, formal evaluation data) to determine if it supports the hypothesized academic problem at multiple levels (explained elsewhere). When triangulating data, evaluators should harness their professional judgment, informed by integrated data analysis and pattern seeking techniques, to make eligibility decisions based on what a preponderance of the data collected over time indicates (Schultz & Stephens, 2009). Influenced by the combination of data, the evaluator must then establish an appropriate instructional program that is individualized to the specific needs of the student being evaluated.

Steps for *Decide* are:

- Organize, sort, and make visual representation of data (e.g., Multiple Sources of Data Worksheet).
- Apply data to PSW policy to answer these questions: Does the child exhibit a PSW in performance? In achievement? Both?
- Apply data to PSW policy to determine if the pattern exists.
- Revisit Exclusionary Factors to ensure they are not the primary cause of struggle.
- Report finding and make recommendations to the ARD committee.

Management and Analysis of Data to Identify PSWs

There is power in data! Preliminary PSW often emerges during the *Review* step, therefore, once the evaluator enters the *Decide* step, the additional data collected during test planning and the assessment stages should merge with the patterns earlier identified in the earlier stages. With the referral question in mind, the evaluator should begin to organize the data for analysis. The legal defensibility of your decision is elevated when educational and special education eligibility decisions are supported by MSD, combined with the expertise and professional judgment of the evaluator. Concisely, it is essential that an evaluator's determination of the existence of a PSW be supported by a preponderance of MSD.

Integrated data analysis is therefore crucial for establishing PSW and making sound eligibility decisions. Integrated data analysis involves the pooling of all data and its thorough consideration (e.g., NRT results, RTI data, criterion-referenced test, observations). The practice involves examining the chain-of-evidence as well as applying pattern seeking techniques, determining the trustworthiness (weight, accuracy) of the data collected, organizing it, triangulating it, and conducting a logical cross-validation analysis (Schultz, Simpson, & Lynch, 2012). Below, we provide more information on several aspects of data integration and analysis.

Pattern Seeking Techniques

Numerous operational methods for identifying patterns of strengths and weaknesses have been introduced over the years (Flanagan, Fiorello, & Ortiz, 2010; Floyd, Evans, & McGrew, 2003; Floyd, Meisinger, Gregg, & Keith, 2012; McGill, Styck, Palomares, & Hass, 2006; McGrew & Wendling, 2010; Newton & McGrew, 2010; Schultz, Simpson, & Lynch, 2010; Schultz & Stephens, 2017). These

methods have both unique and shared characteristics. Within the C-SEP™ model, PSW identification is characterized by the following features:

1. The MSD collected spans a period of time.
2. Integrated data analysis using techniques of pattern seeking.
3. Consideration of predictive and treatment validity.
4. The establishment of logical and empirical relationships between data.

When identifying patterns of weaknesses, it is essential to try to determine how long the academic problem has lasted, since, by nature, patterns become predictable. Thus, the evaluator should make every effort to pinpoint when the academic concern(s) began. Once identified, attention can be turned to determining which steps have been taken to address the weakness and to identify the observable outcomes the remedial efforts produced. Such insight will be invaluable to understanding the weakness and will inform testing and evaluator recommendations.

The techniques for pattern seeking described in this manual are adapted from a process outlined by McMillan and Schumacher (2010). We embrace a mixed-methods data analysis process (see Johnson & Onwuegbuzi [2007] for a full description) (Schultz, Simpson, & Lynch, 2012). In this context, when identifying patterns, the evaluator must:

1. Examine the trustworthiness (validity) of the data.
2. Triangulate the data (Leech & Onwuegbuzi, 2007; McMillan and Schumacher, 2010).
 - a. Is the description of the hypothesized academic problem supported by all three levels (links) of evidence?
 - b. Evaluation teams should consider all three levels of data to determine if the student is suspected of having an SLD. This practice is recommended because it provides a more complete picture of the academic problem and is informative when determining more effective intervention strategies (Frijters, et al., 2011).
3. Rule out all exclusionary factors before determining the presence of a learning disability.
4. Analyze the data to identify a PSW consistent with the presence of SLD after all testing and observations have been completed and all extant information has been gathered.
5. Order, *sort*, and review the data.
 - a. Data can be organized in multiple ways to better understand the nature of the problem. It can be organized according to:
 - i. **Construct** (e.g., *Gc*, *Gs*): Organize multiple sources of data according to the construct being assessed (e.g., when assessing *Gc*, the following data should be collected and organized together: vocabulary tests, reading comprehension assignments,

Conditional Analysis: Consider the student's performance in a particular academic area within various instructional environments (e.g., investigate the student's performance in math under testing conditions, classroom conditions, small group instruction).

- Did the student receive adequate grade level instruction from a highly qualified educator?

Since RTI is commonly used to assist students having difficulty, this type of data should be collected because it provides a wealth of information about the student. When considering RTI, we recommend the evaluator determine:

- Were interventions implemented with fidelity?
- Did the interventions match the area(s) of concern?
- What does the progress monitoring indicate?
- Were appropriate tests used with culturally and/or linguistically diverse students?
- Which measures produced results, and which did not?

Similarly, data collected from other stakeholders, such as teachers and parents, must equally be evaluated. When analyzing data collected directly from teachers (e.g., checklists, formal rating scales, interviews), evaluators should take into consideration the following:

- The teacher's familiarity with the student.
- The teacher's familiarity with curriculum/grade level standards.
- The comparison group (who are they comparing the student's performance to).
- Any potential biases the teacher might have towards the student (e.g., he cannot do anything).

For data collected from parents (e.g., checklists, formal rating scales, interviews) evaluators should consider the following:

- The parent's understanding of the referral process.
- The parent's understanding of the reason for referral (what is the problem).
- The parent's ability to provide a comprehensive developmental history.
- The parent's understanding of items on the checklist or formal rating scales.
- The parent's understanding of the importance of an interview.
- Any potential biases the parent might have towards the student or school.

Once again, if the evaluator is suspicious of data collected from a teacher or parent, they should look to other sources. This may require follow-up interviews with the referral source or approaching other teachers and/or family members to balance existing concerns.

If formal/informal measures are available and considered, the evaluator should contemplate questions such as:

- What do the tests measure?
- What do those scores mean/indicate?
- Why has or has the student not performed well on a particular test (e.g., task demand analysis)?

When considering student observations, the evaluator should contemplate questions such as: Did the observation in the classroom and/or testing behaviors further support the reason for referral (e.g., how did the student address the tasks, how did the student approach problem-solving, what was the student's motivation level)?

Inter-Normative & Intra-Relative Measures

Evaluations measure weaknesses either as inter-normative or intra-relative. According to the federal statute 34 CFR § 300.309, both characteristics are necessary and important to establish an SLD. The law states:

(II) exhibits a pattern of strengths and weaknesses in performance achievement, or both relative to age (*normative strengths & weaknesses*), grade level standards, or intellectual ability, as indicated by significant variance (*relative weakness*) among specific areas of cognitive function, such as working memory and verbal comprehension, or between specific areas of cognitive function and academic achievement.

More specifically, “deficit in one or more of the basic psychological processes relative to age” is a normative weakness, while “significant variance” is an intra-related weakness.

When evaluating normative strengths and weaknesses, evaluators should consider a student's performance on tests in relation to age-based/grade-based norms; publisher guidance (qualitative descriptors), RPIs, as well as each score in relation to multiple sources of data. PSW identification models require the assessment of individual differences. Certainly, there must be a difference in a skill (e.g., cognitive, academic, language) to classify the construct (e.g., memory, reading, oral expression, math calculation) as a “strength” or a “weakness.” These individual developmental differences can be described as intra-individual or inter-individual. Intra-individual differences, on the one hand, are differences within an individual (e.g., Kat's math calculation skills are much higher than her basic reading skills) (Sattler, 2016). On the other hand, inter-individual differences are differences between individuals (e.g., Sam's language development is slower when compared to his same age peers) (Sattler, 2016).

Individual differences (PSW) can be measured using NRT data to identify the presence or lack of a disability. However, before we further examine these differences (intra- vs. inter-), we would like to recall core ideas about interpreting standard scores:

- a) Difference does not mean disability.
- b) Scores are used to identify the construct (e.g., reading, long term retrieval, auditory processing) as being a strength or weakness. The scores themselves are not “strengths” or “weaknesses.” They need to be interpreted in the context of other data sources and described in a manner such as, this standard score of 110 may indicate a strength in working memory.
- c) Since standard scores are rank order data and demonstrate a place in line, interpreting them by the score alone may overestimate or underestimate a construct. Scores well above the average range require the student to correctly respond to multiple items, however, due to the nature of norming tests, some low average scores required few correct responses.
- d) A commonly accepted threshold for most psychoeducational tests for a standard score difference is 1 SD. These numerical (statistical) differences are just numbers until the examiner attaches meaning to them. Extracting meaning from the differences requires the use of statistics, professional expertise, and MSD. When analyzing intra- and inter-individual differences in SLD identification and classifying them as a strength or weakness, it is important to understand what scores mean and why they are necessary.
- e) Using a standard bell curve, and interpreting the true scores within the confidence intervals, a Normative Weakness (statistical) would fall below 1 SD below the mean, or eighty-five.

The publisher of each test provides qualitative descriptors, so these should be consulted as different tests do not always describe things in the same manner. For example, some tests consider the average range as SS 90-110 while others consider SS 85-115 as average. It is important to note that whether you call a score of 85 average, low average, or below average, it is still an 85. In this frame, standard score analysis should be considered a necessary first step, albeit superficial interpretation. It is the construct we are describing as a strength or weakness, not scores.

To be identified as SLD using PSW, a student must have scores (constructs) that are both in the average range or above (normative strengths) and scores (constructs) which are in the normative weakness range, or below average. Some weaknesses, however, are not always “meaningful.” For example, a student’s Visual Processing may be below average, but it may not have an impact on his learning. In this instance, it is just a difference. In contrast, if a student’s below average Auditory Processing impacts their acquisition of basic reading skills, this would be an indicator of a disability.

This returns us to our objective of defining inter-individual and intra-individual differences. An inter-individual (normative strength or weaknesses) difference occurs when scores are compared to a population and has both a statistical and meaningful difference. Here, we are observing differences among a group of individuals to identify strengths and weaknesses. In contrast, intra-individual differences (strengths and weaknesses) are also a necessary condition for SLD identification. Support

for this can be found in policy and in the test manuals. Interpretation of intra-individual differences is best interpreted using base rate data and a standard bell curve. An intra-individual difference occurs when scores are compared to the individual and has both a statistical and meaningful difference.

Final Rule-Out of Exclusionary Factors

While some Exclusionary Factors were preliminarily ruled out during the *Review* stage of C-SEP™, evaluators will need to reconsider them at the *Decide* stage. Using the Exclusionary Factors checklist and all the available data, document and rule out each factor independently. Always consider what, if any, effect each contributing factor (e.g., second language, attention, speech/language disorder) has on performance in the classroom or in a test setting? If no impact can be observed, the factor can be ruled out as a *primary* cause of the student's struggle. Rule-out and documentation of the rule-out is necessary to meet both state and federal legal requirements.

Merging the Available Data

C-SEP™ is constantly observing each data point on an individual basis, then merging it with other MSD for further consideration. Merging data requires organizing it in a logical and meaningful manner whereby comparisons can be made, and patterns observed. To this end, data can be organized in diverse ways, and each evaluator will inevitably adapt practices that they find most effective. Regardless of the approach selected, keep in mind that a logical organization ensures that comparisons are accurate, and it makes data triangulation easier. For these reasons, evaluators should:

- Organize the data in a logical manner to better triangulate and analyze the data for interpretation.
- Pull all data together that assesses each cognitive, language, and achievement area (e.g., reading: RTI data [intervention & progress monitoring data], work samples, previous and current grades, teacher information, parent information, state testing, district benchmarks, observation during reading class, testing behavior during formal testing, data from norm-referenced reading assessments).
- Organize data collected for each cognitive area that correlates/contributes to each achievement area (e.g., *Ga*, *Gl*, *Gr*, *Gc*, and *Gs* data should be organized and considered in relation to the impact on the student's performance in each area of reading).
- Do you have an adequate amount of reliable and valid data to answer the referral question? If not, continue collecting additional forms of data to use when making data-based decisions.

Some tools are available to assist data merging, including Dr. Holman's Analysis Tool. Such tools are likewise valuable for documenting resources and for helping to organize written reports.

Systematic Data Analysis Using the Chain of Evidence Technique

Once the quality of the data has been confirmed and it has been merged, the evaluator should conduct a systematic analysis by organizing and analyzing the data using the three-step chain of evidence technique. The chain of evidence used to analyze a PSW in learners suspected of having a SLD functions as follows:

- Link #1: Informal, archival, and extant data, observations, teacher, and parent information.
- Link #2: Information derived from informal, non-standardized assessments such as progress monitoring data and benchmark testing.
- Link #3: Results of standardized testing (i.e., cognitive processing, language, achievement) in all areas of suspected disability.

Each link is examined in closer detail below. Remember, analysis for Link 1 and Link 2 typically occur during the *Review* step of C-SEP™. Consequently, the evaluator is not required to start the pattern seeking process over during the *Decide* step, but rather to integrate and consider the data collected during *Plan* and *Assessment* with any preliminary patterns that emerged during the *Review*.

Link 1:

The first link in the chain of evidence includes analysis of informal assessment data such as attendance records, home language survey, developmental history, school health files, previous test scores, grades and grade history, and records from previous schools, since these sources often provide a wealth of data relevant to the referral concern (Salvia, Ysseldyke, & Bolt, 2010). Other informal sources include classroom observations, writing portfolios, classroom work samples, and parent and teacher interviews which often occur informally. Concerning the latter sources, it is essential for the evaluator to learn how the teacher and parent view the problem since these individuals know the child best and spend a great deal of time with the student on a regular basis.

When examining Link 1, the evaluator should analyze the first link of information (informal data) in terms of how it relates to the referral question. More specifically, consider:

- a. Are the teacher/parent concerns consistent with the referral question?
- b. Does school performance reflect a history consistent with the referral question?

- c. Is there evidence of other explanatory factors that could answer the referral question (attendance rates, previous instruction, linguistic or cultural factors, health issues)? If so, further investigate them by analyzing the data already collected or by collecting more data to clearly document the issue.

After the above analysis, the evaluator should determine which patterns emerge. In particular:

- d. Does the data show a clear pattern of difficulty over time in the area of concern?
- e. Do multiple sources of informal data support the pattern?

Link 2:

The second link in the chain involves examining the results of non-standardized testing, which contains key information about the student's academic functioning. This link includes benchmark testing on grade-level curriculum, curriculum-based testing from end-of-unit tests, progress monitoring data from curriculum-based measures for RTI, classroom running records, reading miscue analyses, and criterion-referenced tests. Concerning the latter resource, criterion-referenced tests provide information on skills that the child has mastered as well as those where the teacher needs to target instruction.

When analyzing these data sources to explore learning behavior and RTI, the evaluator should consider.

- a. What is the rate of improvement based on RTI data?
- b. Is there a discrepancy between end of unit tests and more comprehensive assessments like benchmarks?
- c. How does error analysis contribute to the understanding of the student's performance?

Following this analysis, the evaluator should identify which patterns emerge by considering?

- d. Does the data show a clear pattern of difficulty over time in the area of concern?
- e. Do multiple sources of informal data support the pattern?

Link 3:

The final link in the chain of evidence for identifying a child's academic problem is found in formal evaluations conducted during the Assess stage. This link involves analyzing how the student performed on formal evaluations by comparison to the data from the first two links?

- a. If different, how does it differ and what explains the difference?
- b. How did the student perform in relation to same age/grade level norms and do these findings support pattern(s) that emerged in the first two links?

- c. How did the student perform in relation to relative strengths and weaknesses (e.g., intra-cognitive, intra-achievement, or intra-oral language), and do these findings support pattern(s) that observed in the first two links?
- d. When analyzing the data through different lenses (e.g., standard scores, RPI, CLI, *Gf-Gc*), which pattern(s) emerge and how do they compare to those PSW observed in the first two links?
- e. How does task analysis contribute to what you know about the student?
- f. What patterns emerge from this analysis and how do they fit with the patterns from the previous analysis?

Cross-Validation of the Data

Concisely, cross-validation simply means that you are analyzing all the merged data to determine whether it makes logical sense in its entirety and points in the same direction. Stated differently, when you look at the pool of data, does it all support the same conclusions concerning the student's PSW? Questions to consider include:

- What patterns emerged from the combination of formal and informal data?
- How do the results on formal measures fit into the larger profile of the student's performance?
- Is there more than one data source indicating the same strength and/or weakness?
- How do the data findings align to the referral question?
- Does the student's cognitive and/or language PSW align to his/her achievement PSW in a logical, research-based manner?
- Were exclusionary factors considered and ruled out as the primary cause of academic deficit(s)?
- Did other evaluators come to the same conclusions when analyzing the data?

Naturally, we encourage evaluators to utilize a problem-solving team approach when reviewing cases to validate findings with other evaluators. Engaging in a team review is beneficial for multiple reasons, including that it allows for a fresh pair of eyes on the individual case, it reduces the chances of bias, and it is equally a fantastic way to ensure consistency in determinations within a school or district.

Managing Inconsistencies in the Data

Inconsistencies will inevitably appear in data. When they do surface, they should not be avoided by cherry picking data or buried to hide them. On the contrary, the evaluator and evaluation team should carefully consider why an anomaly occurs. A collective examination of the MSD should be conducted to try to explain inconsistencies based on logical and scientifically supported decision-making.

Otherwise Normal Profile and C-SEP™

We are often asked if there is an “otherwise normal ability profile” expectation for SLD in C-SEP™ like there is in XBA? The short answer is yes, but how C-SEP™ determines it is associated with “unexpected underachievement” and the nature of SLD and not a mathematical formula as in XBA. More specifically, an “otherwise normal ability profile” described by Flanagan and colleagues (2013) refers to students with areas of strengths at or above the average range along with a specific area or areas of processing weakness. This is determined using a *g*-value “which indicates the likelihood that the individual has at least average overall ability to think and reason.” The metric is based on a cognitive standard score contrived by a formula within XBASS (see Essentials of Cross Battery, third edition, Flanagan, Ortiz, & Alfonso, 2013).

By comparison, C-SEP™ comprehensively assesses SLD and goes beyond cognitive explanations and/or mathematical formulas for identifying “unexpected underachievement.” Multiple sources of data are used to determine if the student’s learning difficulties are “specific” and NOT “general” and *unexpected*. When identifying an otherwise normal profile, we encourage readers to take into consideration the following points and examples:

1. We would **EXPECT** a child who was provided appropriate instruction in reading/and or math delivered by qualified personnel **to ACHIEVE**. To **NOT ACHIEVE** is **UNEXPECTED** (appropriate instruction means that most students in that setting (< 80%) are achieving).
2. We would **EXPECT** a child who was provided appropriate intense instruction in reading/and or math via RTI, tutoring, or other pre-referral interventions delivered by qualified personnel **to ACHIEVE**. To **NOT ACHIEVE** is **UNEXPECTED** (appropriate instruction in Tier 2, 3 settings are adequate if the student is showing reasonable progress).
3. We are discriminating between general and specific learning difficulties, and at this point the concept of PSW enters.
 - a. We would **EXPECT** a child who struggles with all academic subjects, has cognitive processes in the same range such as below average, struggles with language, and shows no significant variation in abilities (pervasive weakness). This describes **EXPECTED UNDERACHIEVEMENT** (Not SLD).
 - b. We would **EXPECT** a child who is able to communicate adequately (language) and able to perform certain skills well such as math (showing ability to think and reason quantitatively), should be able to achieve in other areas such as reading. If they can do certain things well, but underachieve in other areas, this describes **UNEXPECTED UNDERACHIEVEMENT** (Possible SLD).
- 4 Other things to consider when determining if there is **unexpected achievement**:

- a. We also need to distinguish SLD from ID. We would **EXPECT** a child with a 50 FSIQ is UNDERACHIEVE = **EXPECTED UNDERACHIEVEMENT**.
- b. We would EXPECT a child who is missed 20% of school (Exclusionary factor) to UNDERACHIEVE = **EXPECTED UNDERACHIEVEMENT**.
- c. A fatal flaw of the old IQ-Achievement discrepancy model was the use of Global scores such as FSIQ and GIA. Due to the nature of SLD (significant variance in scores: PSW), these global scores are not the best estimate of “overall functioning.”
- d. When determining **UNEXPECTED UNDERACHIEVEMENT** using C-SEP™, the evaluator must examine factors other than cognitive standard scores (e.g., instructional/treatment response, language, exclusionary factors, and strengths).

Decision-Making and Professional Judgment

As previously mentioned, throughout the evaluation process, evaluators should employ professional judgment. It will be used when developing the initial referral question, determining the fidelity of general education instruction prior to the referral, considering the fidelity of RTI, and determining if a student is eligible for SLD, among other considerations. Naturally, all decisions made by an evaluator should be based on reliable and valid data collected during the evaluation process. Nevertheless, when making final decisions regarding the eligibility of a student, the evaluator must be assured that a preponderance of the data collected identifies a pattern consistent with the legal definition of SLD. Accurate, pedagogically sound, and legally defensible eligibility decisions are best made using a multi-dimensional assessment process (Dehn, 2006), using multiple sources of data, and through the exercise of professional judgment (Schultz & Stephens, 2009).

Schultz and Stephens (2009) argue that professional judgment is an essential component of an evaluator’s understanding of learning, assessment, and eligibility. Since professional judgment can ensure the accuracy of decision-making in complex situations involving eligibility, classification, and evidence-based practices (Schalock & Luckasson, 2005), it is vital for the evaluator to be confident in his/her ability to converge knowledge, theory, experience, contextual factors, and critical thinking skills when determining SLD eligibility. Some states, such as Texas, New Mexico, and Georgia, have established guidelines regarding the use of professional judgment.

Schalock and Luckasson (2005) recommend six strategies to ensure competent decision-making when considering eligibility. The six strategies are:

1. Conduct a thorough social history.

2. Align data and its collection to the critical questions at hand.
3. Apply broad-based assessment strategies.
4. Implement intervention best practices.
5. Plan, implement, and evaluate support.
6. Reflect on cultural competence and diversity.

Building upon these six strategies, Schalock and Luckasson (2005) further provide guidelines for meeting the requirements of each strategy. For the sake of space, we use Table 9 below to summarize their theory.

Table 9. Guidelines for Professional Judgment (Adapted from Schalock & Lucasson, 2005)

Strategies and Guidelines	Characteristics of Strategy and Guidelines
Conduct a thorough social history.	• Be familiar with current legal findings. • Investigate, collect, and organize all relevant data covering student's life stages. • Obtain information from the individual, teacher, parents, and other pertinent sources. • Identify educational experiences or lack thereof. • Identify differences in data findings and investigate and address possibilities for differences.
Align data and its collection to the critical question(s) at hand.	• Develop a focused referral question (FRQ) and align data with FRQ. • Converge data and analyze findings for patterns of strengths and weaknesses. • Use evidence-based knowledge to incorporate current measures and strategies. • Assess and increase competence in professional judgment (e.g., additional training).
Applying broad-based assessment strategies.	• Incorporate a broad evaluative strategy. • Use multiple sources of data as part of the evaluation process. • Use a variety of formal, informal, and formative assessments. • Consider treatment validity when assessing. • Link assessment to instruction or intervention.
Implementing intervention best practices.	• Integrate the results of the analysis into the selection of clearly stated intervention strategies.

Planning, implementing, and evaluating supports.	• Provide appropriate training and integrity checks of intervention implementation. • Utilizing a problem-solving model with progress monitoring to evaluate programming. • Consider relevant individual, family, and school factors.
Reflecting cultural competence and diversity.	• Follow all legal and regulatory requirements for nonbiased assessments. • Be sensitive to culturally diverse populations when selecting assessment instruments. • Understand language proficiency and the implications on learning.

Prior to closing this chapter, it is essential to recall the importance of adhering to the guidelines and standards which govern our profession. From adhering to federal and state law to reading and obeying test publisher guidance, evaluations are managed and overseen by numerous restrictions and guidelines. These standards are further buttressed by best-practices, like those just outlined concerning professional judgment. It is the responsibility of the evaluator to educate themselves and observe these standards.

Page Intentionally
Left Blank

Report Writing

As evaluators, we are charged with conducting a thorough evaluation and this thoroughness should be demonstrated in the report. The Full Individual Evaluation (FIE) is a psychoeducational report that concisely summarizes the assessment process and findings. In this sense, the FIE is written documentation that the evaluator and other team members have fulfilled their obligation and conducted a comprehensive and individualized evaluation of all areas of suspected disability. In its broadest sense, this document should present the reason for referral, indicate the areas of suspected disability, state the Focused Referral Question (FRQ), and discuss the assessment procedures utilized. It simultaneously should determine whether a disability condition exists and explain how the disability impacts the student's access and progress in the general education curriculum. It will also provide recommendations for adapting the content, methodology, or delivery of instruction to meet the individual needs of the student being evaluated in addition to other services and accommodations required.

This chapter lays out some suggestions for how report writing can be approached. Although there are no specific guidelines on how an FIE should be organized, some schools or districts may provide general templates or specific guidance on the FIE, and we encourage evaluators to leverage these resources. Common sense tells us that its layout should be logical, easy to read, and its content should be simple to understand.

Nonetheless, federal law provides guidance for what should be contained in the report, so will begin by quoting this at length. It reads:

34 CFR § 300.311 Specific documentation for the eligibility determination

(a) For a child suspected of having a specific learning disability, the documentation of the determination of eligibility, as required in §300.306(a)(2), must contain a statement of—

- (1) Whether the child has a specific learning disability;*
- (2) The basis for making the determination, including an assurance that the determination has been made in accordance with §300.306(c)(1);*
- (3) The relevant behavior, if any, noted during the observation of the child and the relationship of that behavior to the child's academic functioning;*
- (4) The educationally relevant medical findings, if any;*
- (5) Whether—*

(i) The child does not achieve adequately for the child's age or to meet State-approved grade-level standards consistent with §300.309(a)(1); and
(ii)

(A) The child does not make sufficient progress to meet age or State-approved grade-level standards consistent with §300.309(a)(2)(i); or

(B) The child exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, State-approved grade level standards or intellectual development consistent with §300.309(a)(2)(ii);

(6) The determination of the group concerning the effects of a visual, hearing, motor disability, or an intellectual disability; emotional disturbance; cultural factors; environmental or economic disadvantage; or limited English proficiency on the child's achievement level; and

(7) If the child has participated in a process that assesses the child's response to scientific, research-based intervention—

(i) The instructional strategies used and the student-centered data collected; and

(ii) The documentation that the child's parents were notified about—

(A) The State's policies regarding the amount and nature of student performance data that would be collected and the general education services that would be provided;

(B) Strategies for increasing the child's rate of learning; and

(C) The parents' right to request an evaluation.

(b) Each group member must certify in writing whether the report reflects the member's conclusion. If it does not reflect the member's conclusion, the group member must submit a separate statement presenting the member's conclusions.

Regardless of how the above information is incorporated into the document, these legal standards should create the backbone of your FIE. Adherence to this federal framework, both in terms of the assessment process and its documentation, should ensure that the evaluator has complied with federal law. Evaluators should equally consult any guidance provided by their state, although this varies nationwide.

In total, the FIE generated should meet the standards and requirements set by federal and state law, as well as any guidelines established by your district or school. We also make additional recommendations which evaluators may find useful in the following sections.

Dos and Don'ts

There are several dos and don'ts of report writing. In this section, we will briefly address the following topics: timelines, legal defensibility, legalese, accessibility, and shortcuts.

Timelines

It is essential that established timelines for assessments and report writing be maintained. Keep in mind that timelines sometimes vary according to state or district.

Strengthening Legal Defensibility

When writing a FIE, it should be written as if it were destined to be presented before a court of law. While only a fraction of assessments ever become embroiled in due process, we recommend conducting your assessment and drafting your report as if litigation is certain. In this manner, an evaluator will adhere to all applicable laws and best practices. Unfortunately, if an assessment is poorly conducted and the FIE poorly written (lacking data/support), it is increasingly unlikely to provide a legally defensible determination or the appropriate framework for addressing a student's individual needs. This requires putting in the necessary time and energy.

The FIE serves as the legal foundation for the IEP. It provides the IEP/ARD team members with the information they need to make appropriate eligibility and instructional decisions about the student. For this reason, it is fundamental that evaluators create a clear connection between the data contained in the FIE with the educational needs, goals, services, and placement. These are the aspects a hearing officer will examine should an FIE be brought to due process.

FIE reports should present evidence that legal mandates were adhered to. This does not mean that the report should *state* that it met the law, but rather should clearly demonstrate how the law was met (Hass & Carriere, 2014). Solid and well-explained data acquired from multiple sources should be outlined as well as the use of multiple and varied assessment practices and data sources (34 CFR § 300.304(b)(1)). Moreover, as mentioned above, it is the decision of eligibility, needs, goals, services, and placement that will be the points of contention should litigation be pursued. When well designed, these characteristics of the FIE will be reflected in the Present Levels of Academic Achievement and Functional Performance (PLAAFP). The PLAAFP drives the goals, goals drive services, and services drive placement. Be certain to link your assessment data to the PLAAFPs, the PLAAFPs to educational needs, educational needs to the goals and accommodations, and the combination of these to placement.

Unfortunately, many FIEs do not fully integrate and synthesize their data into meaningful content for readers. Instead, most only present the test scores, thereby leaving it up to the reader to interpret the data shown. Fact is, most readers of an FIE, namely parents and teachers, are unlikely to be qualified to interpret hard data. It is therefore the evaluator's job to explain to them what the assessment included, what its findings were, and which recommendations are made, and this should be done in a manner that the average individual understands (Hass & Carriere, 2014).

By way of summary, always remember that the purpose of the assessment is to recommend instructional programming that links to the data acquired from the assessment process. All students do not require the same interventions and accommodation, so carefully select these to assure they reflect the *individual* needs of the student being evaluated.

Avoid Legalese

Although the law must be followed, evaluators should limit the use of legal jargon, law quoting, or explicit references to the law when drafting reports (Hass & Carriere, 2014). These practices are disparaged for several reasons. Foremost, the general audience of an FIE are not legal scholars, attorneys, or law clerks, so it makes little sense to reference federal or state legislation. Its inclusion only increases the risk of creating confusion. Second, just because someone explicitly quotes the law, it does not prove that they obeyed it. To illustrate, rather than quote 34 CFR 300.304(c)(1)(i) while insisting that the assessment instruments used in the evaluation were "selected and administered so as not to be discriminatory on a racial or cultural basis;" explaining how the instrument is not biased because it was designed to be used with the demographic group the student represents and denote that precautions x, y, and z were taken to manage bias during test administration.

We provide an example of legalese below to demonstrate why these are not suited for a general audience. A commonly referenced statute, 34 CFR § 300.304(c)(1)(11), allegedly proves an assessment met the linguistic needs of a student. It is usually stated that the assessments utilized:

...were provided and administered in the child's native language or other mode of communication and in the form most likely to yield accurate information on what the child knows and can do academically, developmentally, and functionally, unless it is clearly not feasible to so provide or administer."

Interestingly, this quote is generic. It does not provide any valuable information to the reader about how the law was adhered to. In fact, by reading this sentence, it is not even clear in which language a child was assessed. Equally confusing, what are these other modes of communication, and were they used?

An alternative approach is to provide information in a manner that is more accessible. To adhere to the elimination of legalese to explain the law was adhered to, a more appropriate paragraph might state:

Maria's first language is Spanish, and she is an English Learner (EL). Her academic instruction has been in English since Kindergarten. Maria scored in the Intermediate level on the 2017 California English Language Development Test (CELDT), which indicates that she has a good command of English that is equal to her average peers. Her mother reported that she speaks Spanish with her parents and English with her siblings. Maria reports that she speaks English with her classmates and neighborhood friends. Both Maria and her mother think her English skills are stronger than her Spanish skills. For these reasons, Maria was interviewed in English and the tests used during this assessment were administered in English. Her cognitive abilities were assessed by using tests that minimized receptive and expressive language.

The above example provides a comprehensive yet clear explanation of which language Maria was assessed in, and how and why that language was chosen for the evaluation. By comparison to the "legalese" example, this statement has far more substance. Equally important, should someone be called upon to read the para statement through the eyes of the law, it is clear 34 CFR § 300.304(c)(1)(11) was adhered to.

Regardless of our recommendation to avoid directly referencing the law, please keep in mind that some schools or districts may demand these statements be incorporated into SLD evaluation documents. In this case, evaluators have an obligation to include them. However, we remind evaluators not to simply insert the required legal references (verbatim or otherwise) and stop there. Instead, they should also contextualize and demonstrate how the assessment and its decisions meet the legal requirements cited. Once again, these explanations should be done in a manner that can be digested by the average consumer.

Ensure Accessibility

In addition to satisfying the legal requirements governing the assessment process, evaluators have an ethical obligation to produce a report that is useful and accessible to relevant stakeholders, as noted above. The individuals that have the greatest stake in the assessment process are the child being assessed, their parents or guardians, and the teachers that are charged with that student's education. With this in mind, we suggest that these individuals be your target audience, and that you write in a manner that they understand.

Table 10. Ethical Standards of Report Writing

NASP (2010) Principles for Professional Ethics	Educational Diagnostician Standards
Report • Present findings in language that is understandable to the reader. • Emphasize interpretation and recommendations. • Support recipients in their work or interactions with the child. Interpretation • Written in simple language. • Based on a convergent and comprehensive assessment source.	• Use appropriate communication skills to report and interpret assessment and evaluation results. • Effectively communicate to parents/guardians and professionals the purposes, methods, findings, and implications of assessment. • Interpret and use assessment and evaluation data for targeted instruction and ongoing review. • Use assessment and evaluation information to plan individualized programs, establish measurable annual goals and objectives, and make instructional decisions that result in appropriate services for individuals with disabilities, including those from culturally and/or linguistically diverse backgrounds.

Holman, 2021

Stylistically, there are numerous ways an evaluator can make their reports readily accessible to their audience. For instance, evaluators can select to use shorter words as opposed to longer ones. In addition, it is recommended to use active, everyday English, as opposed to the passive voice that is further complicated by scientific vocabulary and professional jargon. If jargon must be introduced,

such as making references to Standard Scores or RPI, these should be clearly defined in simple language. We also suggest providing clear examples to illustrate what these terms mean.

Once the document has been drafted, it is likewise recommended that an evaluator use the Flesch Kincaid readability measure to determine the readability of your report. According to this scale, it is suggested that your report should not exceed ninth grade readability. In most cases, this should be sufficient to clearly present your assessment and its findings while keeping the information accessible to a wide audience. The combination adheres to the professional standards established by recognized organizations in our field, as observed in Table 10.

Avoid Shortcuts

Some evaluators may be tempted to pursue shortcuts, such as copying and pasting reports, to hasten report writing. While this practice is tempting, it is littered with pitfalls that range from being outright embarrassing to downright risky. On the one hand, it would be quite easy to accidentally transform the gender of a student, for example, when an evaluator uses Robert's report as a template for Suzie's. While colleagues might be taken aback or confused by such a mistake, a parent is likely to be offended. On the other hand, it would be extremely easy to make clerical errors while modifying a copied report. For instance, certain tests used in the copied version may not have been used with the student being assessed. Similarly, assuming that the two students were given the exact same tests, an evaluator might forget to change a score, meaning that Suzie's scores were Robert's. Such errors are easily overlooked, and should an evaluator miss them, then the FIE will be presenting tests, scores, and interpretations that do not apply to the student whose name is on the report. The potential legal implications of this hypothetical should be sufficient to discourage the use of cutting and pasting into reports.

Structure and Content

When writing an FIE, there are specific recommendations that we would like to make as it pertains to the structure and content of an FIE.

Structuring a Report

The structure of a report has a direct impact on its readability and accessibility. If templates or software are used to produce a text, there is little to no flexibility in design. Moreover, flexibility and style can be impacted by having multiple individuals contributing to its content—a practice which is sometimes inevitable in FIEs. For these reasons, the general components and layout of an FIE should be agreed upon by all contributing authors.

Nonetheless, there are three commonly used structures for FIE reports: domain-based; test-based; and referral question-based. The domain-based approach organizes the report according to the domains of concern. One of its limitations includes its tendency to concentrate substantial amounts of data and their interpretation into one or two sections. If data is not properly explained, sections of the report can become overwhelming and confusing to readers. By comparison, a test-based structure arranges the text around the testing instruments that were used during the assessment. One downside to this approach is that it is test-centric, and if written incorrectly, could give the impression that NPTs carried more weight in decision-making than other sources of data.

The last common report structure is the Referral-Based Consultative Assessment Report Model (Batsche & McCoy, 1983) which recommends a referral question-based arrangement. This approach walks the reader through the reason for the referral and the referral question, to the assessment plan, the data collection tools and methods used to answer the referral question, then the results in the context of the questions, and finally provides recommendations based upon the referral question. Hass and Carriere (2014) advocate a referral or question-based approach. Evaluators will have to select the method that they prefer and feel most comfortable writing.

Regardless of how an FIE is structured, it should contain five elements, and these should be clearly identifiable to the reader:

1. A review of instructional interventions.
2. A clear statement of the student's present levels of academic achievement and functional performance (PLAAFP).
3. An explanation of why the student is not making adequate academic progress.
4. A determination if the student meets eligibility criteria for a specific learning disability and/or other educational disabilities.
5. An appropriate and effective plan that is individualized to the educational needs of the student.

Document Assessment Procedures

When drafting reports, be certain to demonstrate the use of the variety of instruments and measures used throughout the assessment process. Nevertheless, the instruments and measures utilized should not only be denoted in passing but linked to the referral question(s) and clearly connected to the relevant and functional information they provided. However, keep in mind that a table is insufficient. Resources must also be contextualized into the text of the FIE, where they should be individually introduced, the data acquired from them explained, and how that data links to eligibility decision-making and recommendations. Table 11 illustrates RIOT, one method that can be used to document the measures and instruments used to collect data during an SLD evaluation.

Table 11. Documenting Procedures and Sources using RIOT

Procedure	Source	Date
Records Review	Cumulative school report	10/18/2023
	RTI progress monitoring	10/18/2023, 11/6/2023
	Classroom work samples	10/18/2023, 11/6/2023
Interviews	James Wilson, Student	10/18/2023, 11/6/2023
	Mr. Lewis, 5th Grade Math Teacher	10/29/2023
Observations	Small group reading	10/31/2023
	Whole class direct math instruction	11/1/2023
Testing	Woodcock-Johnson IV Tests of Achievement	11/1/2023
	Woodcock-Johnson IV Tests of Cognition	11/2/2023

Integrate and Synthesize Data

Be certain to integrate and synthesize the data presented in the FIE. One method of doing this is to use an evaluation matrix that contains the research question(s) and RIOT components. Using the RIOT matrix, it is possible to identify which assessment procedures and resources were used to answer each question. In fact, the arrangement provides an ideal framework for presenting information into the FIE, where resources and instruments can be described, synthesized, and contextualized with the area of concern, the research questions, as well as assessment findings and recommendations.

Another recommendation we make is when displaying scores and score reporting, select one presentation format and maintain it throughout the report. For instance, an evaluator may select to show scores by domain or test type, while simultaneously distinguishing between core and selective testing. Regardless of the arrangement chosen, be certain to use it throughout the document.

Our final advice in this context is to never just present data. Instead, interpret the data you present in a clear and understandable manner. For example, synthesize and contextualize scores within the FIE, explaining both their meaning and significance in terms of the student's strengths and weaknesses and the services or accommodations they require.

Include Impacts and Needs Statements

Reports must present and describe a student's academic and behavioral performance. It must likewise address the potential impacts and needs of the disability identified. For these reasons, Impact and Needs Statements are essential components of an FIE and IEP. Among their functions, these statements inform the provision of services required to assist the student achieve within the general education curriculum. An Impact Statement, for instance, describes how a disability impacts the student's access and progress in the general education curriculum. The Impact Statement is constructed on the weaknesses and/or behavioral problems caused by the disability and are identified in the FIE, and should likewise address the functional, developmental, and academic concerns. The Impact Statement is naturally associated with the Needs Statement, the latter of which should:

- 1) Determine whether a student has a disability and
- 2) Determine whether the student requires specifically designed instruction or related services.

The Needs Statement points to which supports are required within the framework of the student's identified strengths and weaknesses. Every impact identified should be accompanied by a needs statement and recommended programs and/or services to address the individual need.

Both Impact and Needs Statements inform the development of the present levels of academic achievement and functional performance (PLAAFP), as well as goals and services. The PLAAFP establishes individualized goals, objectives, and programming recommendations which can be leveraged to make targeted recommendations for specialized instruction. The PLAAFP should perfectly correspond with the needs identified in the Needs Statements (1 for 1). This means that an

evaluator must establish clear and specific methods and objectives for each of the needs presented to ensure that individualized support and monitoring measures are put in place to individually address the child's unique circumstances. In this frame, task analysis is valuable for understanding what may facilitate or inhibit learning (e.g., visuals). After identifying the disability and its impact on accessing and progressing in the curriculum, attention can turn to which measures are recommended to improve student outcomes.

Offer Meaningful Recommendations

A PLAAFP statement leads to the development of a needs statement and individualized recommendations based on the student's specific needs. Effective needs statements will use the evaluation data to identify the disability condition and consider its impact on learning whereby specific recommendations for services and support can be made. Needs Statements should provide specific information about the types of services and supports a student requires to access and progress in the general curriculum. They thus operationalize data and findings from the individualized evaluation, explaining what may affect the student's access and/or progress in general education.

Nevertheless, regardless of whether a student qualifies for special education services or not, an FIE must include recommendations. The recommendations provided should be based on assessment data and be specifically targeted to the individual needs of the student. Use the existing data to develop an individualized educational program and identify recommendations and necessary accommodations to support the student. However, we caution evaluators not to use generic checklists or templates when making suggestions but instead to tailor their recommendations to the individual strengths and weaknesses of the student being assessed. Accepted best practice is to provide recommendations that are specific enough that a teacher or parent could implement them. The law, unfortunately, does not dictate how many recommendations should be included, so professional judgement is essential in this context. In general, the principle of quality over quantity should apply.

Page Intentionally
Left Blank

Special Populations

C-SEP™ for Dyslexia Evaluations

C-SEP™ with Evaluations for Culturally and Linguistically
Diverse (CLD) Students

Page Intentionally
Left Blank

C-SEP™ for Dyslexia Assessments

Students with dyslexia struggle with reading and spelling, while demonstrating the ability to learn in the absence of print. These characteristics are mentioned in common descriptions of the disorder. The International Dyslexia Association (IDA) (2002) defines dyslexia as follows:

Dyslexia is a specific learning disability that is neurological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge. (International Dyslexia Association, 2002)

From these definitions, we can observe the *primary characteristics* of dyslexia are:

- difficulty reading words in isolation,
- difficulty accurately decoding unfamiliar words,
- difficulty spelling, and
- having slow, inaccurate, or labored oral reading.

Demonstration of one or more of these listed traits are most often associated with weaknesses in the *cognitive processes* of:

- Phonetic Coding (phonological awareness),
- Naming Facility (rapid automatic naming/speed of lexical access), and
- Memory Span (phonological memory), as well as difficulty learning the names of letters and their associated sounds.

Secondary manifestations of dyslexia can likewise include:

- difficulty with reading comprehension,
- difficulty with written language, and
- limited vocabulary development.

Documenting and assessing these attributes are essential for both identifying dyslexia and for providing options for remediation.

The *Review, Plan, Assess, and Decide* steps of the C-SEP™ framework can be applied to create a targeted and efficient evaluation of students suspected of having dyslexia. As usual, pre-referral data should inform the area(s) of suspected disability and the selection of standardized measures.

Multiple sources of data should be collected and integrated to produce a comprehensive picture of the student's abilities at every step of the evaluation.

Table 12 depicts a generic dyslexia evaluation. This framework is not intended to be a checklist, but rather an illustration of how data collection, integration, and interpretation might be pursued when evaluating a student suspected of having dyslexia. However, since each student is unique, their strengths, weaknesses, and individual needs will vary, and these differences will naturally impact the manner and instruments required during an evaluation. The table below depicts a dyslexia assessment as a part of a Full Individual Evaluation.

Table 12. Dyslexia Assessments

From Referral to Decision (Review, Plan, Assess, Decide)	
Review: <i>The purpose of this data collection stage is to a) clarify referral question, b) conduct preliminary assessment of exclusionary factors, c) assess instruction and instructional response, and d) identify patterns of academic strengths and academic weaknesses.</i>	
Actions	Tools and Strategies
1. Review Referral Concerns.	Clarify referral question: - Does the referral question suggest that the reading concerns are phonological in nature? - Does the referral question suggest that the reading difficulties are unexpected?
2. Review Educational Records.	Document findings related to referral question including the student's exposure to explicit and systematic instruction in essential components of reading: phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension strategies. - Educational records particularly pertinent to a dyslexia referral include measures of reading such as: - ISIP- examine and compare individual areas of performance (e.g., word analysis, vocabulary, text fluency, phonological awareness, spelling, comprehension). - DRA- examine and compare accuracy percentage, words per minute, comprehension points. - TPRI- examine and compare phonemic awareness, graphophonemic knowledge,

	word reading, accuracy/fluency, reading/listening comprehension. 5. CBM- examine and compare individual areas of performance (phonemic awareness, phonics, reading fluency, word reading, reading comprehension, vocabulary). 6. Running Record- conduct an error analysis to determine strengths and weaknesses and determine instructional reading level.
3. Review response to scientific, research-based intervention in the area of suspected disability.	- Do the interventions implemented match the areas of deficit revealed from the reading measures administered? - What is the pattern of response?
4. Collect evidence to support that the student's difficulties are not due to lack of appropriate instruction in reading, irregular attendance, lack of experiential background, or language difference.	- Attendance. - School history including special programming (504, Special Education, BIL/ESL). - Language proficiency measures/LPAC documentation. - Parent information (does a parent or family member have a history of reading difficulty or diagnosis of dyslexia?). - Teacher information.
<i>Plan:</i> The planning stage involves creating an assessment plan based on a hypothesis generated from previously collected data, the referral question, what is known about the construct and the individual student.	
1. Organize and analyze informal data collected.	Document emergence of patterns if evident from Review stage.
2. Develop a working hypothesis of the referral concern.	- Where is the breakdown in reading for the student (decoding, word retrieval, comprehension, fluency)? - How does the student's performance in reading compare to other areas? - Do results from the Review stage suggest the student is demonstrating characteristics of dyslexia?
3. Determine what additional data is needed to answer the referral question.	"Based on the student's academic difficulties and characteristics and/or language acquisition, areas related to vocabulary, listening comprehension, oral language proficiency, written expression, and other cognitive abilities may need to be assessed" (TEA, 2014, Dyslexia Handbook, p. 21).
<i>Assess:</i> Data collection continues based on the assessment plan.	

1. Measure Language (if applicable):	Administer norm-referenced tests of oral language proficiency in native (if available) and second language.
2. Measure Cognitive Processes:	Administer norm-referenced tests of phonological/phonemic awareness and rapid naming of symbols or objects.
3. Measure Achievement:	Administer norm-referenced test of achievement to assess the following skills: reading words in isolation, decoding unfamiliar words accurately, reading fluency (rate & accuracy), reading comprehension, and spelling. Administer a measure of letter knowledge (name & sound).
4. Analyze Norm-Referenced Data:	Examine data in relation to task demands. Document qualitative observations of behavior during testing. Consider influence of oral language proficiency in relation to academic skills.
5. Administer Additional Tests (if necessary):	Potential areas for additional assessment include vocabulary, listening comprehension, verbal expression, written expression, handwriting, orthographic processing, math calculation / reasoning, phonological memory, and processing speed.
<i>Decide:</i> This stage is conducted after all data is collected and has two main objectives: a) integrate and analyze data and b) apply data to definition of dyslexia	
1. Analyze the data to determine if the student's "difficulties in the areas of reading and spelling reflect a pattern of evidence for the primary characteristics of dyslexia with unexpectedly low performance for the student's age and educational level" (TEA, 2014, Dyslexia Handbook, p. 22).	1) Does the student exhibit one or more of the primary characteristics of dyslexia: • Reading words in isolation • Decoding unfamiliar words • Reading fluency (rate & accuracy) • Spelling (an isolated deficit in spelling is not sufficient to identify dyslexia) **Students being taught to read/write a transparent language may be able to decode and spell "real and nonwords adequately but demonstrate difficulties in reading rate with concurrent deficiencies in phonological awareness and rapid automatized naming" (TEA, 2014, Dyslexia Handbook, p. 22). Reading difficulties are more often noticed in <u>reading speed</u> and <u>reading comprehension</u>.

	2) Does the student exhibit deficits in the cognitive processes typically associated with dyslexia? “Average phonological awareness scores alone do not rule out dyslexia. Ongoing phonological processing deficits can be exhibited in word reading and/or spelling” (TEA, 2014, Dyslexia Handbook, p. 22). 3) Are the reading/spelling/phonological processing deficits unexpected? Does the student demonstrate cognitive ability to support academic learning? Potential sources of supporting evidence include: • GIA/FSIQ • Reasoning & Knowledge: Gf-Gc • Oral Language: Listening Comprehension, Oral Expression, Vocabulary • Math Calculation/Reasoning Skills • Academic Knowledge: Science, Social Studies, Humanities • General Information 4) Are there secondary characteristics evident in reading comprehension/written expression?
2. Report findings and make recommendations:	Are there additional complicating factors that warrant further investigation/intervention and how do they impact the student’s ability to benefit from the Gen. Ed. dyslexia program?

Table Reference: Holman, S., Stephens-Pisecco, T., Schultz, E., & Moon, G. (2018).

Retrieved from Stephens-Pisecco & Schultz (2018). *C-SEP™ Handbook*.

Page Intentionally
Left Blank

C-SEP™ and CLD/EL Evaluations

Assessment of culturally and linguistically diverse (CLD) students requires special consideration of cultural and linguistic variables. These variables are to “understand the referral, determine the appropriate language and form of the testing, interpret the test results in light of other information, and participate effectively in the decision-making process for the student” (Alvarado, 2011). Consideration of these variables not only apply to students identified as Limited English Proficient/English Learners (LEP/EL) but also to students for whom another language has had considerable influence on their English language skills development that may *not* be identified as LEP/EL. The reasons to make special considerations are quite evident, but a brief review is provided below.

Among the many considerations, it is important that evaluators distinguish students struggling to learn a language from those who have a disability. We propose this because LEP/ELs often struggle with the semantical and syntactical aspects of learning a second language (Mitkovska, 2016; Talosa & Maguddayao, 2018), while students with SLD often struggle with several aspects of language. Within this frame, evaluators must determine if a student’s struggles are due to normal second language acquisition challenges or if it is disability related. Error analysis is beneficial in this frame can provide more specific information about an examinee’s performance than can be derived from a composite of test scores. Error analysis can likewise help identify common sources of difficulty underlying diverse skills areas including language, and it is equally valuable for planning intervention (Breaux, 2020).

Similarly, TDA can help identify patterns related to input, output, and language demands of an individual subtest. These input and output demands can be related to sensory, motor, cultural, or linguistic factors (Breaux, 2020). Analyzing processing demands of a subtest can thus provide important insight into the cognitive and linguistic demands of a test. For example, the *Passage Comprehension* test in the Woodcock-Johnson IV Tests of Achievement requires the cognitive and linguistic process of integrating syntactic and semantic properties of printed words and sentences into a representation of the whole passage. If a student does poorly on this task, an analysis of this process may help to determine if the issue is due to a disability or language acquisition challenges (McGrew, LaForte, & Schrank, 2014).

Naturally, evaluators are expected to take great care when selecting testing instruments for these students. There are several legal requirements pertinent to an assessment of CLD students.

Most notably, IDEA (2004) stipulates that assessments and other evaluation materials be:

- 1) Selected and administered so as not to be discriminatory on a racial or cultural basis.
- 2) Provided and administered in the child's native language, or other mode of communication, and in the form most likely to yield accurate information on what the child knows and can do academically, developmentally, and functionally, unless it is clearly not feasible to so provide or administer.

Hence, language mastery and acculturation should be considered when selecting and administering assessments. Acculturation is defined as the process of adopting the behaviors and traits of the dominant culture (Croucher & Kramer, 2017).

While first-generation ELs are usually exposed to their immigrant parent's culture in the home, they adopt the dominant culture through their social and educational experience. However, it is important to understand that acculturation is a process, and each student may acculturate at varying levels and diverging speeds. Factors to consider include, but are not limited to, visiting their native country, the reason for immigration, family attitudes towards acculturation, English fluency, and community socialization. When provided with educationally relevant and accommodative environments, post-secondary outcomes can be promising. Studies have demonstrated that first-generation immigrants tend to have better educational outcomes when provided with bilingual instruction, culturally resonant caring pedagogy, supportive teachers, and high post-secondary expectations (Hopkins et al., 2013).

When interpreting assessment results, particularly as it relates to first and second-generation students, evaluators must carefully consider the student's home and learning environment. For example, access to English language development support, newcomer programs, and their integration with non-EL peers will help to indicate the potential and degree of acculturalization. In addition, evaluators should be cognizant of test items that may reflect the values and experiences of the dominant culture. Missed items or low scores should be analyzed by the examiner to determine if these low scores are reflective of low acculturation, a lack of exposure, or bias, rather than the manifestation of a disorder.

The *Review, Plan, Assess, and Decide* framework of C-SEP™ can be applied to assessments of culturally and linguistically diverse students. Pre-referral data should inform the selection of standardized measures and guide the evaluation process in a manner that accommodates linguistic and cultural diversity. Multiple sources of data collected using varied and targeted methods should generate data that can be integrated to yield a comprehensive picture of a student’s abilities.

Table 13 outlines a generic evaluation of an EL following the C-SEP™ model. This table should not be used as an assessment checklist, but rather is designed to illustrate how the process might function with culturally and linguistically diverse students that are suspected of having a specific learning disability. Portions of the information contained in the Tools and Strategies section of the framework was adapted from *Best Practices in the Special Education Evaluation of Students Who Are Culturally and Linguistically Diverse* by Dr. Criselda Guajardo Alvarado (2011).

Table 13. CLD/EL Evaluations

From Referral to Decision (<i>Review, Plan, Assess, Decide</i>)	
<u>Review:</u> <i>The purpose of this data collection stage is to a) clarify referral question, b) conduct preliminary assessment of exclusionary factors, c) assess instruction and instructional response, and d) identify patterns of academic strengths and academic weaknesses.</i>	
Actions	Tools and Strategies
1. Review Referral Concerns:	Clarify referral question: • Are the documented learning difficulties consistent with typical second language acquisition? • Are the documented learning difficulties the result of incongruous home/school expectations and experiences? • If the referral concern is vague or unclear, clarify by interviewing source or requesting additional data.
2. Review Educational Records:	Document findings related to the referral question including the student’s native language, length of exposure to native language/English, quality of exposure to native language/English, history of special language programming, cultural/background information. Educational records particularly pertinent to a SLD referral for a CLD student include: 1. Home Language Survey 2. LPAC Language Proficiency Tests (over time)

	3. Functional Language Ability Survey 4. Home Language Questionnaires (Parent's & Student's perspective) 5. TELPAS (over time) 6. School History (in/out of U.S.) 7. State-wide/Other Assessments (over time)
3. Review response to scientific, research-based intervention in the area of suspected disability.	• Do the interventions implemented match the areas of deficit indicated in the referral question? • How were the student's language needs addressed in the RTI process? ○ Was the intervention "meaningfully accessible"? • What is the pattern of response?
4. Collect evidence to support that the student's difficulties are not due to lack of appropriate instruction in reading or math and are not primarily the result of a visual, hearing, or motor disability, intellectual disability, emotional disturbance, cultural factors, environmental or economic disadvantage, or Limited English proficiency.	Parent information Teacher information Attendance School history including special programming (e.g., 504, Special Education, BIL/ESL). Language proficiency measures/LPAC documentation This information should be interpreted through a lens of: • Typical second language acquisition process. • Typical academic benefit of special language programming based on type of service (see Collier & Thomas, 2001).
5. Establish failure to meet age- or grade-level state standards in one of eight areas when provided appropriate instruction.	Multiple data sources require review and consideration in relation to the referral question and the individual student. Data sources to analyze: a) RTI data b) In-class tests c) Grade average over time d) Norm* or criterion referenced tests e) Statewide assessments f) Teacher input g) Work samples h) Observations i) Cognitive Academic Language Proficiency (CALP) j) Any other useful data *Norm referenced data to be collected during formal testing sessions.
6. Review research related to referral concerns and review data collected.	Review and Assessment sheets for each area of SLD eligibility.

<i>Plan:</i> The planning stage involves creating an assessment plan based on a hypothesis generated from previously collected data, the referral question, what is known about the construct and the individual student.	
1. Organize and analyze informal data collected.	Document emergence of patterns if evident from Review stage.
2. Develop a working hypothesis of the referral concern.	• What area(s) of weakness emerged from the review stage? • What are the potential causes of/explanations for those weaknesses?
3. Determine what additional data is needed to answer referral question(s).	a) Select parallel measures of oral language proficiency if available in both languages. ➤ Tests such as the WMLS III or WJ-IV Oral Language can be used to determine oral language proficiency and dominance in English and Spanish. ➤ If tests such as the WMLS III or WJ-IV Oral Language tests have been administered previously, results should be reviewed and compared with multiple sources of data (e.g., RTI, work samples, benchmarks, statewide assessment) and used in the planning of a comprehensive evaluation. b) Select norm-referenced tests of cognition and achievement. ➤ Tests of cognition should be in the student's dominant language, if available. If dominant language is unclear even after task analysis, administer core set of norm-referenced tests of cognition in the student's native language/language and form most likely to yield accurate results. ➤ English academic testing should occur if the student has received English instruction for a year or more. ➤ Native language academic testing (formal or informal) should occur if the student has received instruction in that language for a year or more. ➤ For a referral in Reading, Writing, Listening, or Speaking, tests such as WMLS III, can be used as the core of the evaluation, with selective testing from the WJ-IV ACH or Bateria IV. ➤ Review CALP scores c) Observation tools (language demands, classroom functioning).

<u>Assess:</u> <i>Data collection continues based on the assessment plan.</i>	
1. Measure Language Proficiency:	Administer norm-referenced tests of oral language proficiency in native (if available) and second language. • Interpret results: take into consideration second language acquisition process and quantity and quality of language exposure. • Determine dominance: compare oral language total and conduct task analysis of language complexity if necessary (basic vs. applied). • Use informal measures to determine oral language proficiency in native language if formal tests are unavailable.
2. Measure Psychological Processes:	Administer the selected core set of norm-referenced tests of cognition.
3. Measure Achievement:	Administer norm-referenced test(s) of achievement based on assessment plan.
4. Analyze Norm-Referenced Data:	✓ Organize data by construct and publisher's quantitative and descriptive analysis. ✓ Examine data in relation to task demands. Document qualitative observations of behavior during testing. ✓ Interpretation of <u>cognitive testing</u> should include consideration of: • Consider the degree of linguistic demand and cultural loading of tests. ✓ Interpretation of <u>academic testing</u> should include consideration of: • Amount and quality of academic exposure in the language. • Effects of different instructional programming on academic performance. • Impact of socioeconomic factors on language learning. • Native language loss phenomena.
5. Administer Additional Tests (if necessary):	Administer additional tests in areas that show deficiencies and/or are needed to explore referral question.
6. Observe in Classroom:	1. Observe the student in the general education classroom in the subject(s) related to the area(s) of concern. a) Describe how the academic area(s) of concern impacts the student's performance in the classroom.

	b) Note specific behaviors related to the academic area(s) of concern including the student's reactions to learning and feedback from the general education teacher. c) Complete language demands assessment.
7. Organize Existing Informal Assessment Data and/or Administer Informal Assessments (if needed).	Useful informal assessments can be: a) Curriculum-based measures (e.g., nonsense word fluency, oral reading fluency) b) Tests of limits c) Language sample
<u>Decide:</u> This stage is conducted after all data is collected and has two main objectives: a) integrate and analyze data and b) apply data to PSW policy.	
1. Organize, sort, and make visual representation of data	Use integrated data analysis and pattern seeking techniques. ➤ Educational history, sociological / cultural / economic factors, and language background must be considered in the interpretation.
2. Apply data to PSW policy to answer these questions:	The child exhibits a pattern of strengths and weaknesses in: ▪ Performance: yes or no? ▪ Achievement: yes or no? ▪ Both: yes or no? The pattern is relative to: ▪ Age; (how does the student compare to same age peers-both in the classroom and on norm-referenced measures?), ▪ State-approved grade-level standards; (how far from grade standards is the student's performance?) or ▪ Intellectual development (Unexpected under-achievement? Intra-individual strengths / weaknesses?)

3. Apply Data to PSW policy to answer these questions:	The pattern is evident as indicated by significant variance: ▪ Among specific areas of cognitive function such as working memory and verbal comprehension; or (yes or no and which ones?) ▪ Between specific areas of cognitive function and academic achievement; and (yes or no and which ones?) The pattern is relevant to the identification of an SLD using appropriate assessments (see guidance below).
From Guidance Document:	In evaluating specific areas of cognitive functioning to determine a pattern of strengths and weaknesses, schools should take into consideration the federal definition of SLD as “a disorder in one or more of the basic psychological processes involved in understanding or in using language” (CFR §300.8(c)(10)). An identified pattern of strengths and weaknesses should be linked to the failure to achieve adequately as described above in the identification process. Students whose classroom achievement indicates a pervasive weakness that does not constitute a pattern of strengths and weaknesses should not be determined to have a SLD. Students who meet the criteria as having intellectual disability should not be determined to have a SLD.
4. Report findings and make recommendations to ARD committee.	ARD committee determines whether the child is a child with a disability and the educational needs of the child. ➤ To determine language of instruction, consider: • Native language • Dominant language • Proficiency levels in each language • Years of exposure to each language • Type of exposure to each language • Academic skills in each language • Age • Current language of instruction

FAQs

Many questions about C-SEP™ has been posed during conversations with colleagues and peers over the past decade. Many of the same queries have repeatedly surfaced. For this reason, we want to take a moment to address several in the form of FAQs.

- 1. *What is the role of RTI in C-SEP™?*** RTI plays a critical role in C-SEP™ as it serves numerous purposes in the overall SLD identification process. RTI:
 - a.** Helps establish underachievement and documents that the student received appropriate instruction.
 - b.** Helps differentiate instructional causalities vs students with SLD as well as students with general low achievement. A student with SLD is unable to learn despite intensive instruction/intervention or makes minimal progress when analyzing rate of improvement (ROI) over time (e.g., flat trend line). A student with general low achievement makes progress when provided with intensive intervention and instruction over time (e.g., trend line going up steadily when considering rates of improvement). A student with this profile is learning “abled” (NO SLD) but at a slower rate. A student who is considered as having an “instructional casualty” makes considerable progress when considering ROI.
 - c.** Provides additional cognitive, language, and achievement data when used as “part of” the evaluation and supports norm-referenced testing. For example, oral reading fluency measures when examined on the task demand level involve processing speed and decoding. The MAZE task requires using context clues (fluid reasoning) to fill in blanks. Nonsense words fluency requires processing speed and is a measure of auditory processing.
 - d.** Assesses exclusionary factors and motivation.
- 2. *I am not a neuropsychologist; I do not feel adequately trained to implement some of the Cog-PSW models.*** C-SEP™ is an educational classification model that adheres to educational policy. The approach uses a single core battery of tests which publishers have organized in a manner that has construct validity and are psychometrically sound. Each instrument contains specific instructions on administration and interpretation, with major cognitive batteries that are informed by the field of neuropsychology. By using the test publisher’s organizational structure

and applying it to the referral concerns, tests are already organized. In other words, neuropsychology has been done for you. This not only aids in your decision-making capabilities but meets the letter and spirit of the procedural guidelines for evaluations.

3. ***Sometimes the core set of tests needs to be supplemented. Is this the “selective” part of C-SEP™?*** A feature of the C-SEP™ approach is that “selective” testing is used throughout the process. The first “selection” determines which core set of tests are to be used. C-SEP™ was initially conceptualized using the WJ-IV as this set of tests emphasized a strong core set of tests for each battery (COG, ACH, and OL) and includes a “selective” testing table. Other revisions of tests also share this feature (having a “core”) such as the CELF-5 and KTEA. An examiner may select any cognitive battery of tests as the “core” (i.e., WISC-V, SB-5, KABC) that best answers the referral question and is most appropriate for the student’s background. The examiner would also select a language test that measures oral expression and listening comprehension. After the core is administered and interpreted, referral questions can be further explored by using “selective” tests which will provide additional information.
4. ***Doesn’t the use of professional judgment allow too much subjectivity and allow examiners too many grey areas?*** All methods of SLD require the use of professional judgment throughout the process from test selection, interpretation, to deciding if the results are valid. However, good data is equated with good decisions. C-SEP™ procedures help examiners collect diagnostically precise data by moving from a broad testing perspective (core) to an increasingly narrow scope (selective) of testing. When using C-SEP™, data is analyzed using integrated data analysis to obtain a complete picture of the student. Moreover, it recommends strictly adhering to local and state policy, so eligibility decisions reduce grey areas and “false positives.”
5. ***We are used to administering many more tests than C-SEP™ recommends, don’t we have to administer at least two tests for every ability?*** No. C-SEP™ is a PSW model and considers norm-referenced testing as part of the overall data collection. It does not require a minimum number of tests for a formula (outside the battery) to work. Reliability is ensured through integrated data analysis. Abilities can be assessed and evaluated using a “variety of assessment tools and strategies” and not just norm-reference testing. In addition, tests have been updated and improved to make them “cognitively complex” which allows for more than one ability to be assessed. Some models require a minimum number of tests for the formula to work, but this results in the administration of an excessive number of tests without providing additional data.

Furthermore, administering too many tests increases the chances of error, misinterpretation, and an over-reliance on standard scores. C-SEP™ balances comprehensiveness with efficiency.

6. ***What distinguishes a child with a specific learning disability from low achievers?*** This question has perplexed the field for years and influences discrepancy and other mathematical models which have standard score cut-offs. C-SEP™ uses standard scores when making these distinctions, however, scores are not the sole determinant since other data is considered. To distinguish these populations using C-SEP™, consider the following factors outlined in the table (Tables Adapted from the Ventura County SELPA: PSW Model of SLD Eligibility Procedural Manual, <http://www.vcselpa.org/Resources-for-Teachers-and-Staff/Pattern-of-Strengths-and-Weaknesses>).
7. ***Which Skills/Knowledge/Training is Required to Use C-SEP™:*** We offer the following recommendations to enhance professional development and your ability to use C-SEP™ appropriately. This is not intended to be an exhaustive list.
- a) Attending graduate training in test administration, scoring, and interpretation.
 - b) Understanding of tests and measurements including the limitations of norm-referenced tests.
 - c) Understanding of the Cattell-Horn-Carroll (CHC) Theory of Cognitive Abilities along with a working knowledge of Luria (PASS theory), processing theories, executive functions, working memory, language, and development of academic abilities.
 - d) Understanding of the relationships between cognition, language, and achievement.
 - e) Being able to evaluate each instrument for validity, reliability, and construct development by using the examiner and technical manual. In addition, be able to understand the scores derived from the publisher.
 - f) It is likewise important to understand the reciprocal relationship between teaching and learning and classroom implications. Likewise, obtain an understanding of informal assessments used in teaching and learning to supplement existing limitation of this data.

Page Intentionally
Left Blank

C-SEP™ Case Illustrations

The following excerpts are case illustrations using the C-SEP™ methodology. While the case illustrations below contain scores and extracts from reports, these examples are not intended to be, and should not be viewed as, full case studies. On the contrary, they are designed to show how data can be documented and explained.

Page Intentionally
Left Blank

Case Study: Lori (Full)**Name:** Lori Redacted**Date of Birth:** 05/28/2015**Age:** 6 years, 9 months**Sex:** Female**School:****Teacher:** SCHULTZ (OL)**Grade:****ID:****Date of** 03/04/2022 (OL)**Testing:**

03/03/2022 (COG)

03/03/2022 (ACHA)

Examiners: Edward Schultz (ACHA)**Tests Administered***Woodcock-Johnson IV Tests of Cognitive Abilities* (Norms based on age 6-9)*Woodcock-Johnson IV Tests of Oral Language* (Norms based on age 6-9)*Woodcock-Johnson IV Tests of Achievement Form A and Extended* (Norms based on age 6-9)**Table of Scores***Woodcock-Johnson IV Tests of Cognitive Abilities* (Norms based on age 6-9)

CLUSTER/Test	AE	RPI	SS (95% Band)	SS Classification	PR
GEN INTELLECTUAL ABIL	6-0	68/90	86 (81-92)	Low Average	18
AUDITORY PROCESS (<i>Ga</i>)	5-10	76/90	88 (79-97)	Low Average	22
VOCABULARY	6-9	90/90	100 (91-109)	Average	51
Oral Vocabulary	6-3	83/90	94 (84-103)	Average	34
Number Series	6-4	68/90	92 (84-100)	Average	29
Verbal Attention	6-7	87/90	98 (88-108)	Average	44
Letter-Pattern Matching	5-6	7/90	77 (62-92)	Low	6
Phonological Processing	5-3	42/90	75 (62-87)	Low	4
Story Recall	7-8	94/90	108 (96-121)	Average	70
Visualization	6-10	91/90	101 (89-113)	Average	52
Nonword Repetition	7-5	93/90	104 (95-113)	Average	61

Woodcock-Johnson IV Tests of Oral Language (Norms based on age 6-9)

CLUSTER/Test	AE	RPI	SS (95% Band)	SS Classification	PR
ORAL LANGUAGE	7-5	94/90	107 (97-117)	Average	69
PHONETIC CODING	6-8	89/90	99 (93-105)	Average	48
VOCABULARY	6-9	90/90	100 (91-109)	Average	51
Picture Vocabulary	7-7	95/90	107 (95-120)	Average	68
Oral Comprehension	7-3	94/90	106 (94-118)	Average	65
Segmentation	6-10	91/90	100 (93-108)	Average	51
Rapid Picture Naming	8-4	98/90	110 (100-120)	Average	76
Sound Blending	6-4	87/90	98 (89-107)	Average	44

Case Illustrations

Woodcock-Johnson IV Tests of Achievement Form A and Extended (Norms based on age 6-9)

CLUSTER/Test	AE	RPI	SS (95% Band)	SS Classification	PR
READING	6-3	50/90	91 (87-94)	Average	27
BASIC READING SKILLS	6-6	79/90	96 (92-100)	Average	39
MATHEMATICS	6-7	87/90	98 (92-104)	Average	45

CLUSTER/Test	AE	RPI	SS (95% Band)	SS Classification	PR
WRITTEN LANGUAGE	6-0	38/90	88 (81-94)	Low Average	20
ACADEMIC SKILLS	6-3	60/90	91 (87-95)	Average	27
ACADEMIC APPLICATIONS	6-3	62/90	91 (86-96)	Average	28
BRIEF ACHIEVEMENT	6-2	56/90	88 (84-93)	Low Average	22

Letter-Word Identification	6-4	57/90	92 (88-97)	Average	30
Applied Problems	6-5	83/90	96 (85-106)	Average	38
Spelling	5-7	24/90	79 (70-88)	Low	8
Passage Comprehension	6-1	43/90	89 (84-95)	Low Average	23
Calculation	6-9	89/90	100 (94-105)	Average	49
Writing Samples	6-3	55/90	93 (84-101)	Average	31
Word Attack	6-10	92/90	102 (93-111)	Average	55
Oral Reading	6-6	76/90	95 (90-100)	Average	36
Sentence Reading Fluency	5-7	46/90	78 (72-84)	Low	8

Cognitive-Academic Language Proficiency (CALP) Scores and Interpretation

CLUSTER	W	WDiff	RPI	CALP	Interpretation
ORAL LANGUAGE	483	6	94/90	4	Fluent
READING	424	-20	50/90	3	Limited
BASIC READING SKILLS	443	-8	79/90	3.5	Limited to Fluent
WRITTEN LANGUAGE	428	-24	38/90	3	Limited
ACADEMIC SKILLS	425	-16	60/90	3	Limited
ACADEMIC APPLICATIONS	439	-15	62/90	3	Limited
BRIEF ACHIEVEMENT	428	-18	56/90	3	Limited

STANDARD SCORES	DISCREPANCY				Interpretation at	
<u>VARIATIONS</u>	<u>Actual</u>	<u>Predicted</u>	<u>Difference</u>	<u>PR</u>	<u>SD</u>	<u>+ or - 1.00 SD (SEE)</u>
<i>Intra-Cognitive [Extended] Variations</i>						
AUDITORY PROCESS (<i>Ga</i>)	88	95	-7	28	-0.58	--
VOCABULARY	100	92	8	77	+0.73	--
ORAL LANGUAGE	107	92	15	89	+1.21	Strength
PHONETIC CODING	99	95	4	62	+0.30	--
Oral Vocabulary	94	92	2	57	+0.19	--
Number Series	92	92	0	47	-0.07	--
Verbal Attention	98	93	5	66	+0.41	--
Letter-Pattern Matching	77	96	-19	7	-1.51	Weakness
Phonological Processing	75	95	-20	3	-1.88	Weakness
Story Recall	108	92	16	90	+1.28	Strength
Visualization	101	93	8	72	+0.59	--
Nonword Repetition	104	96	8	73	+0.62	--
Picture Vocabulary	107	93	14	85	+1.02	Strength
Oral Comprehension	106	93	13	86	+1.06	Strength
Segmentation	100	96	4	65	+0.37	--
Rapid Picture Naming	110	97	13	83	+0.96	--
Sound Blending	98	96	2	54	+0.10	--

Case Illustrations

	STANDARD SCORES			DISCREPANCY		Interpretation at
<u>VARIATIONS</u>	<u>Actual</u>	<u>Predicted</u>	<u>Difference</u>	<u>PR</u>	<u>SD</u>	<u>+ or - 1.00 SD (SEE)</u>
<i>Intra-Oral Language [Extended] Variations</i>						
PHONETIC CODING	99	104	-5	36	-0.36	--
VOCABULARY	100	105	-5	34	-0.41	--
AUDITORY PROCESS (<i>Ga</i>)	88	105	-17	11	-1.24	Weakness
Picture Vocabulary	107	104	3	59	+0.23	--
Oral Comprehension	106	106	0	51	+0.02	--
Segmentation	100	103	-3	44	-0.15	--
Rapid Picture Naming	110	102	8	73	+0.62	--
Sound Blending	98	104	-6	33	-0.43	--
Oral Vocabulary	94	105	-11	18	-0.92	--
Phonological Processing	75	104	-29	1	-2.34	Weakness
Nonword Repetition	104	104	0	51	+0.03	--

	STANDARD SCORES			DISCREPANCY		Interpretation at
<u>VARIATIONS</u>	<u>Actual</u>	<u>Predicted</u>	<u>Difference</u>	<u>PR</u>	<u>SD</u>	<u>+ or - 1.00 SD (SEE)</u>
<i>Intra-Achievement [Extended] Variations</i>						
BASIC READING SKILLS	96	91	5	73	+0.61	--
Letter-Word Identification	92	91	1	55	+0.13	--
Applied Problems	96	91	5	65	+0.38	--
Spelling	79	94	-15	4	-1.78	Weakness
Passage Comprehension	89	92	-3	37	-0.34	--
Calculation	100	91	9	84	+1.00	Strength
Writing Samples	93	93	0	47	-0.08	--
Word Attack	102	93	9	84	+0.98	--
Oral Reading	95	93	2	57	+0.18	--

	STANDARD SCORES			DISCREPANCY		Significant at
<u>COMPARISONS</u>	<u>Actual</u>	<u>Predicted</u>	<u>Difference</u>	<u>PR</u>	<u>SD</u>	<u>+ or - 1.00 SD</u> <u>(SEE)</u>
<i>GIA/Achievement Discrepancy Procedure</i>						
BRIEF ACHIEVEMENT	88	88	0	52	+0.05	No
B						
READING	91	88	3	59	+0.22	No
BASIC READING SKILLS	96	89	7	76	+0.69	No
MATHEMATICS	98	88	10	84	+1.01	Yes (+)
WRITTEN LANGUAGE	88	89	-1	45	-0.12	No
ACADEMIC SKILLS	91	88	3	61	+0.28	No
ACADEMIC APPLICATIONS	91	88	3	67	+0.43	No
ORAL LANGUAGE	107	91	16	91	+1.37	Yes (+)

Woodcock-Johnson IV Tests of Cognitive Abilities Test Session Observations

Level of conversational proficiency: Typical for age/grade

Level of cooperation: Uncooperative at times

Level of activity: Appeared fidgety or restless at times.

Attention and concentration: Distracted often.

Self-confidence: Appeared at ease and comfortable (typical for age/grade).

Woodcock-Johnson IV Tests of Cognitive Abilities Test Session Observations

Care in responding: Prompt but careful in responding (typical for age/grade) Response to difficult tasks: Would not try difficult tasks at all.

Summary of findings

Lori's intellectual development was assessed using the Woodcock-Johnson IV set of tests that measure cognition, language, and her academic skills. These results were analyzed and compared to other data collected during the evaluation. One indicator of a specific learning disability is significant variability between the scores, indicating a pattern of strengths and weaknesses. This data is used as one part of the preponderance of data required for SLD identification. Lori's results related to her intellectual development are consistent with a student with a specific learning disability, specifically in written expression with the condition of Dysgraphia and Basic Reading with the condition of Dyslexia.

Lori obtained a General Intellectual Ability (GIA) score of 86 (low average), but this score may slightly underestimate her overall ability due to some variability in the score, specifically low scores on Letter Pattern Matching (SS 77) and Phonological Processing (SS 75). The Letter Pattern Matching test requires visual-motor integration under timed conditions, which requires discriminating between orthographic patterns (letter patterns). Her test behavior also impacted her scores. She did not appear to give her best effort to the Phonological Processing tests as she continually looked at the clock and kept asking if she was done. She was friendly and polite, but she was comfortable saying "I don't know" to items as soon as they became difficult. She was encouraged to do some "difficult" items and would get some of them correct, but then she would resort to "I don't know." She was fidgety and restless at times and was easily distracted, which is consistent with parent and teacher information regarding attention. Despite this, she could do well on most tasks presented to her.

She had strengths in language-based tasks. Her verbal/oral language is more developed than her written language. Tests such as Oral Vocabulary (SS 94), Picture Vocabulary (SS 107), and Oral Comprehension (SS 106) all required semantics and were well in the average range. She did obtain average standard scores in Passage Comprehension (SS 89) and Writing Samples (SS 93), but her Relative Proficient Index (RPI) scores for these tests were in the frustration level (43/90; 55/90). This means despite having average scores; she may struggle with similar tasks. These low RPI scores are due to limited items on

the test for her age range, meaning she did not have to get too many items correct to obtain an “average” standard score.

Lori’s scores measuring the phonological aspects of language and underlying skills most related to Dyslexia and Dysgraphia were inconsistent. She scored average on the Segmentation test (SS 100) because she could segment words and syllables but could not segment on the phoneme (sound) level. Rapid Picture Naming (SS 110), Word Attack (SS 102), Sound Blending (SS 98), and the Nonword Repetition tests (104) were all in the average range. All these tasks are Dyslexia indicators and, except for Rapid Picture Naming, require the phonological aspects of language to be intact. Her low Phonological Processing (SS 75) test scores can best be attributed to her attention and word retrieval difficulties. Her low scores in Spelling (SS 79) and Letter Word Identification (SS 92, RPI 57/90) require decoding words.

She also struggled on Reading Fluency Tasks (78 RP, RPI 46/90), which require decoding ability. She performed well on other times tests, indicating her struggle is specific to reading and not processing speed itself. Her progress monitoring data indicated that she was significantly below her peers and not reaching her goal. Performance inconsistency is common among students with attention problems. These results were consistent with the Dyslexia screener administered in late January 2022, which was also inconsistent across the subskills assessed. All of this data is consistent with the concerns expressed by her teacher and parent and corresponds with classroom data. Due to this, she may sometimes struggle with some basic reading and writing tasks. Her Reading Comprehension and Applied Written Expression difficulties are more likely related to this skill deficit than a language deficit.

Students with Dyslexia struggle with accurate and/or fluent word reading, poor spelling, and decoding skills, typically resulting from a deficit in the phonological component of language. These difficulties continue despite conventional instruction and are unexpected due to the student’s other abilities. The current data collected, integrated with the historical data set, and current intervention data are all consistent with a child with Dyslexia.

Regarding dysgraphia, when integrated with other informal data, this data suggests the presence of dysgraphia with Lori. According to the Dyslexia Handbook (2021), Dysgraphia is defined as a neurodevelopmental disorder manifested by illegible and/or inefficient handwriting due to difficulty with letter formation. Lori’s writing samples provided by her teacher and writing samples obtained during testing indicate most of the characteristics of dysgraphia as outlined in the handbook: variably and poorly formed letters, poor spacing between letters and words, slow and illegible writing, low volume of written output, and difficulty with unedited spelling. (See Achievement section for actual achievement data related to dysgraphia)

The act of writing by nature requires visual-motor integration, and with all the data obtained (formal testing, writing samples, parent and teacher information), it is evident that she struggles.

Additionally, she struggled with the Letter-Pattern Matching test in the WJ IV cognitive battery. To be successful on this test, a student must be able to recognize orthographic patterns, and use visual perception and motor output to process in concert. Dysgraphia often occurs with deficits in the phonological aspects of language, and Lori has deficits in certain aspects that contribute to her writing and attentional difficulties.

Lori's data indicated a disorder in the psychological processes of auditory processing (Ga), specifically the phonological aspects of visual-spatial (Gv) processing. These weaknesses, combined with her history of struggles with attention, will impact her basic skill development. These processes are related and essential (relevant) to basic Reading (Dyslexia), Reading fluency (Dyslexia), and written expression (dysgraphia).

Impact Statement:

Academic

Lori's specific learning disability in written expression (w/condition of Dysgraphia) impedes her ability to meet the first-grade standards, specifically in handwriting development, which involves printing words and sentences and answering questions legibly. The data indicated that she struggles with all twenty-six letters in the alphabet. These difficulties also impact her spelling, and she could only spell some two-letter words; she should be developing into being able to spell VC, CVC, and CCVC words. Lilly can read words with one syllable and identify syllables but has difficulties with multisyllabic words and connected text due to her disability in basic reading.

Functional

Due to her academic struggles in written language, she will rush through work to "get it done" and not put forth her best effort. Lori also tends to give up easily when tasks get difficult and needs to be redirected and encouraged to complete tasks.

Needs and Recommendations

1. Lori will need additional direct instruction in the foundational components of written language (to help both reading and spelling). This should be in addition to her regular classroom instruction delivered in a small group (4-8 students). Her small group instruction should address the foundational reading and spelling skills since she is only in first grade.
2. The IEP team should consider an evidence-based Dyslexia program.
3. Lori will need direct instruction in spelling in addition to her regular classroom instruction in small groups. The sequence of skills in the state standards is appropriate due to her age.
4. To address Dysgraphia, she will need: a) Activities to support writing (strengthening hand muscles, copying letters, tracing letters, and so forth; b) Activities to help develop automatic writing, and c) explicit practices to address fluency.
5. To improve motivation, activities need to be varied and engaging, and success is critical at each skill.

Case Study: Olivia (DNQ)

Name: Olivia

DNQ: Intellectual Ability

Actual Achievement Data

Background

Olivia is a third grader and has never been retained. She has always been at grade level or above, so she has never participated in any intervention. She decodes fluently and self-corrects while reading orally, according to her teacher. Her spelling is average in isolation and similar in accuracy when writing in context. Her report cards, MAP data, and classroom data indicate that she meets grade-level TEKs with typical support from her teacher. A parent requested this referral along with her advocate.

Scores

GORT-5:		WISC-V:	
Rate	9	Similarities	15
Accuracy	10	Vocabulary	15
Fluency	9	Block Design	9
Comprehension	11	Matrix Reasoning	12
ORI	100	Figure Weights	14
		Digit Span	10
		Coding	6
		Symbol Search	11
		FSIQ	111

WIAT-4:		CTOPP	
Total Achievement	111	Phon. memory	88
Word Reading	96	Rapid naming	85
Reading Comprehension	126		
Spelling	96	Low score(s) in:	
Sentence Composition	123	Nonword Repetition	5
Math Problem Solving	106	Rapid Letter Naming	7
Numerical Operations	107		
Pseudoword Decoding	93		
Listening Comprehension	114		
Oral Expression	116		

Findings

To understand your child's learning pattern and to help with the identification process, we assessed your child's intellectual development by administering norm-referenced tests that measure cognition, language, and academic skills (WISC-V, WIAT-4, GORT- 5, CTOPP-2). These test results were analyzed and compared to other data. This data is considered part of the preponderance of data when determining if your student meets the criteria for a specific learning disability. Our examination of your child's test scores and other data show that Olivia's scores are not consistent with a student with a specific learning disability in reading and the condition of Dyslexia.

According to the data, Olivia is a skilled reader, defined as having fluent execution and coordination of word recognition and text comprehension. This report section will describe the relationship between her intellectual development and reading.

Basic reading skills include decoding abilities, sight words, and phonological processing. Olivia's test scores (Word Reading, SS 96; Pseudoword Decoding, SS 93) in these areas were average. Her ability to read fluently depends on decoding and identifying words. This data indicates she possesses the ability to perform basic reading skills. Additionally, decoding is required to be successful in fluency tasks. Her Oral Reading Index Score (SS 100) on the GORT-5 is in the average range, as is her spelling (SS 96, Spelling); these abilities depend on decoding and processing phonological information.

An underlying ability required to help facilitate fluency is processing speed ability. Several tests of processing speed were administered. Her lowest processing speed scores were in Rapid Naming (SS 85) and Coding (SS 6). This is in the low average range but does not impact her ability to read fluently, as evidenced by her Oral Reading fluency score (GORT Fluency SS 95). Other tests of processing speed (Symbol Search, SS 11, Fluency, SS 9) were in the average range. Grades, classroom data, and MAP scores are consistent with these findings, as she is making adequate progress with the general education curriculum in basic reading.

Reading Comprehension depends on basic reading, reading fluency, and higher-order language skills. Reading comprehension and oral language ability are significant strengths for Olivia. The data indicated very advanced higher-order language skills. Verbal Reasoning (Gc/Gf) test scores (Similarities SS 15; Listening Comprehension SS 114) were above average. Other abilities that require higher level language skills were also well developed. On expressive tasks of language (Oral Expression, SS 116) she could perform above average. Grades, classroom data, and MAP scores are consistent with these findings, as she is making adequate progress in the general education curriculum in reading comprehension.

According to the IDEA, a child with an SLD is one who is not adequately achieving state standards and has a pattern of strengths and weaknesses in achievement, performance, or both. Olivia is achieving adequately in all areas, including reading, according to several sources of data, including test results, MAP data, grades, and progress reports. Her overall intellectual development and profile of scores is absent of any weaknesses and can therefore be described as average-to-high average. She had two

scores in the low-to-low average range, but other data do not support them, and it does not prevent her from achieving at grade level and making progress in the general education curriculum.

Olivia's data was analyzed to consider the presence of dyslexia. According to the Dyslexia Handbook (TEA, 2014), the list of questions in Figure 3.7 (p. 31) *must* be considered when determining dyslexia.

Do the data show the following characteristics of dyslexia?

1. Difficulty with accurate and/or fluent word reading? The preponderance of evidence indicates Olivia can read words accurately and fluently as indicated by her Word Reading (SS 96), Oral reading Index score of 100 on GORT 5, and Pseudoword Decoding (SS 93). These scores are also supported by classroom and MAPs data. Additionally, she can attach meanings to words as evidenced by her advanced reading comprehension skills.
2. Poor spelling skills? Formal measures (WIAT-4, Spelling 96) indicate average spelling ability. Does classroom data support this?
3. Poor decoding ability? Olivia can decode words, as evidenced by her ability to fluently read words accurately, as described above. The Pseudoword Decoding (WIAT-4 SS 93) test also supports this finding. In addition, spelling requires decoding skills.
4. Do the difficulties result from a deficit in the phonological component of language? Olivia does not have a deficit in the phonological component of language. The preponderance of evidence collected that specifically measures aspects of phonological component (Spelling, SS 96, Pseudoword Decoding, SS 93, Word Reading SS 96, Phonological Memory, SS 88) are all in the average range. Additionally, skills that require the phonological component of language in the classroom (Spelling, Reading) are average.
5. Are these difficulties unexpected for the student's age in relation to the students' other abilities and provision of effective classroom instruction? The responses to questions 1-4, and classroom performance show that Olivia can progress in the general education curriculum.

Page Intentionally
Left Blank

Case Study: Timothy (Achievement)

Name: Timothy

Actual Achievement data

A records review and parent information were used to obtain Timothy's educational and developmental history. According to the parent information, Timothy has a history of hearing problems and has had four sets of ear tubes (Tympanostomy) to help with hearing. He was also provided early intervention from ECI for communication and cognition. Timothy was referred for an evaluation by his parents and ECI in 2019 when he was ~2 years and 10 months old for cognition and communication concerns. At that time, he did not exhibit a communication disorder, and his cognitive development was age appropriate.

In Kindergarten, Timothy was provided intervention in the school's MTSS program. By the end of the year, it was reported that he made noteworthy progress toward his goals (letter naming, letter sounds, and sight words), but he was still below grade level. He scored 7 (Letter Sounds) and 8 (Blending) on the Dyslexia screener (at-risk is considered 5 or below), so he did not show at-risk on the screener. Spring MAPS data showed he was in the average range. Despite this, according to the teacher's information, he struggled in the classroom with these skills.

This school year, 22/23, Timothy began the year in the MTSS program and received pull-out reading instruction with the reading specialist and classroom accommodations (visual aids, extra time, and preferential seating). According to the data from the reading specialist, he is making some progress in some skills but still struggles with remembering words from page to page, confusing and reversing letter sounds, and blending. Timothy was assessed with MAPS at the beginning of the school year (BOY, 9/15/22) and mid-year (MOY, 12/6/22). He had an RIT score of 151 in the BOY (fall) administration which indicates low growth/low achievement. The MOY (Winter) administration yielded an RIT score of 145, which also shows low growth/low achievement. His score decreased by 6 points, which is in the 6th percentile.

An RIT score of 166 is the average for students his age and grade. He scored at the average grade level on the MAPS assessment measuring math. According to his classroom teacher, he is not meeting grade-level standards and is a "non-reader." She also reports some inattentive behavior and distractibility. It should be noted that according to information obtained from the mother, he has ADHD.

Timothy has received appropriate instruction in general education as indicated by the data, and according to the data-based documentation of repeated assessments collected in the MTSS program, he is not meeting state standards for his grade level.

Page Intentionally
Left Blank

Case Study: Virginia (Intellectual Development & Achievement)

Name: Virginia

Intellectual Development & Achievement

Virginia's intellectual development was assessed using the WJ IV tests of Cognition and Language. Her test behavior can be described as friendly and cooperative. She was not eager to test and asked several times, "How much longer do we have to do this?" She was distracted often and appeared fidgety and restless at times. The overall results are a fair estimate of her abilities tested. On the WJ IV Tests of Cognition, her overall general intellectual ability (GIA) score of 83 is considered low average; however, her score profile contained significant variance as standard scores ranged from 63-106. Her score profile is consistent with a child with a specific learning disability (SLD) as she demonstrates a pattern of strengths and weakness in performance relative to her overall intellectual development.

Virginia had strengths in tasks that required quantitative/fluid reasoning (Number Series; SS 104). Fluid reasoning is considered a higher-order thinking ability. She also demonstrated an average ability to spatially configure, as evidenced by the Visualization test with a standard score of 103. In addition to strengths in these tasks, she can segment words (Segmentation, SS 106) and process orthographic patterns (Letter-Patterns, SS 95) quickly, indicating average processing speed. This is consistent with the scores on measures related to this (Letter Word Identification, SS 89, and Word Attack, SS 103). This suggests that her reading problems are not related to the phonological component of reading.

Virginia demonstrated several areas of weakness, especially in tasks that required written and spoken language. She struggled in tasks that required listening to and recalling meaningful information, as evidenced by low scores on both the Story Recall Test (SS 77) and Oral Comprehension Test (SS 74). This finding is consistent with her history of oral language problem problems and previous eligibility areas. These skills are also needed to comprehend what she reads. Underlying skills directly involved in these tasks are working memory tasks. Virginia struggled with tasks on tests that measure working memory and scored in the low range in Verbal Attention (SS 87, RPI 66/90) and Sentence Repetition (SS 84, RPI 46/90). These memory tasks and scores are influenced by attention. Her inattentiveness contributes to struggles in this area. The Phonological Processing Tests (SS 63) score is low and can be attributed to her difficulty retrieving words that start with a particular sound. This process is related to both memory stores and word access. This deficit contributes to her reading difficulties, word production when writing, and spelling struggles.

These language deficits will have implications across all subjects (math, science, social studies). For example, the Applied Problems test yielded a standard score of 74, which is in the low range. However, her calculation standard score of 95, and her Number Series (Quantitative Reasoning) standard score of 104, being in the high range, suggest her math problem-solving abilities (word problems) are resulting

from poor oral and reading comprehension. Her language deficits are also evident in her writing abilities.

This set of formal data is consistent with her eligibility data from 2020 (Reading Comprehension, Oral Expression, Listening Comprehension) and other data sources (IEP 2/28/22). Her most recent MAPs testing (12/7/22) indicates that with support, she is making progress. Virginia has a disorder in the basic psychological processes of auditory memory and verbal reasoning. These deficits will impact her language abilities, both written and spoken. The preponderance of evidence indicates Virginia meets the criteria for a child with a specific learning disability in the areas of listening comprehension, reading comprehension, written expression, and oral expression.

Actual Achievement Data

Virginia has received special education services in 2020 since she was in the first grade. She is identified as having an SLD in reading comprehension, oral expression, and listening comprehension. She was identified with a speech/language disorder in articulation at that time as well but was dismissed from speech in the Spring of 2022. MAP data for reading indicates significant growth this last year. In Fall 2022, her Reading -MAP score of 178 was below average compared to her Winter Reading MAP score of 201. This places her in the high range for Growth (93rd Percentile) and average Achievement (46th percentile).

In Fall 2022, her Math-MAPs score of 184 was in the below-average range; her score of 199 in the Winter administration is still in the below-average range (34th percentile) but indicative of high growth (94th percentile). These MAP scores indicate Virginia is benefitting from her IEP as she continues to progress as a result of her program.

She is receiving individualized instruction with reading, and according to her most recent progress report (see IEP and Progress Report), she is making progress toward her reading goals. She receives in-class support, accommodations (extra time, preferential seating, oral administration of tests when appropriate), and content and language & vocabulary support.

Her writing was assessed for this evaluation using curriculum-based measures (CBM). She was asked to respond to a writing prompt for 3 minutes. On writing prompt one, she could produce 33 Total words (TWW), 14 correct writing sequences (CWS), and spell 75% of the words she chose correctly. On the second writing prompt, she produced 18 total words, with 4 correct writing sequences and 88% of the words spelled correctly. Her performance is in the at-risk range. A typical 4th grader should be able to write 46 total words, 44 correct words, and 42 correct writing sequences. A qualitative analysis indicates in addition to spelling struggles, she struggles with capitalization and punctuation.

Her current grades are Science (80), Social Studies (84), Math (80), and ELAR (82); her PE is 97 and Art 100. Norm-referenced testing (see intellectual for scores) of academic skills were all in the

average range (SS 85-103) except for Applied Problems (math) where her standard score is in the low range. Her calculation standard score of 95 and her Number Series (Quantitative Reasoning) standard score of 104, being in the high range, suggest her math problem-solving abilities (word problems) result from poor oral and reading comprehension.

Overall, she is making progress toward her individualized goals. The IEP committee must consider the response to removing support when determining continuing eligibility. At this time, Virginia is benefiting from her support and would most like regress if these supports are removed.

Page Intentionally
Left Blank

References

- Alvarado, C. G. (2011). *Best Practices in the Special Education Evaluation of Students Who Are Culturally and Linguistically Diverse*. Pearland, TX: Education & Evaluation Consultants.
- Barrera, M. (2006). Roles of definitional and assessment models in the identification of new or second language learners of English for special education. *Journal of Learning Disabilities, 39*(5), 142-156. DOI: 10.1177/00222194060390020301
- Barrett, C., & Matthews, M. S. (2021). Ortiz Picture Vocabulary Acquisition Test. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The Twenty-First Mental Measurements Yearbook*. University of Nebraska Press.
- Batsche, G., & McCoy, G. (1983). Pupil Personnel Services Recommended Practices and Procedures Manual. Illinois State Board of Education.
- Braden, J. P., & Matthews, M. S. (2021). Wechsler Intelligence Scale for Children--Fifth Edition, Spanish. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The Twenty-First Mental Measurements Yearbook*. University of Nebraska Press.
- Breaux, K. C. (2020). *Wechsler Individual Achievement Test* (4th Edition): Technical & interpretive manual. NCS Pearson.
- Cho, E., Fuchs, L. S., Seethaler, P. M., Fuchs, D., & Compton, D. L. (2020). Dynamic assessment for identifying Spanish-speaking English learners' risk for mathematics disabilities: Does language of administration matter? *Journal of Learning Disabilities, 53*(5), 380-398. DOI:10.1177/0022219419898887
- Collier, V. P., & Thomas, W. P. (2001). Educating linguistically and culturally diverse students in correctional settings. *The Journal of Correctional Education, 52*(2), 68-73.
- Comerchero, V., & Krach, S. K. (2014). Test of Auditory Processing Skills 3: Spanish-Bilingual Edition. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The Twenty-First Mental Measurements Yearbook*. University of Nebraska Press.
- Croucher, S. M., & Kramer, E. (2017). Cultural fusion theory: An alternative to acculturation. *Journal of International and Intercultural Communication, 10*(2), 97-114. DOI: 10.1080/17513057.2016.1229498
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121-129.
- 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*(3), 300-313. <https://doi.org/10.1002/pits.21675>
- Dehn, M. J. (2014). *Essentials of Processing Assessment*. Second Edition. New York: Wiley.

- Dehn, M. J. (2006). *Essentials to processing assessment*. Hoboken, NJ: John Wiley & Sons.
- Doll, B. J., & Plata-Potter, S. I. (2021). Woodcock-Muñoz Language Survey III. The twenty-first mental measurements yearbook.
- Dufrene, B. A., Kazmerski, J. S., & Labrot, Z. (2017). The Current Status of Indirect Functional Assessment Instruments. *Psychology in the Schools*, 54(4), 331-250.
<https://doi.org/10.1002/pits.22006>
- Evans, J., Floyd, R., McGrew, K., & Leforgee, M. (2001). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescents. *School Psychology Review*, 31, 246-262.
- Every Student Succeeds Act (ESSA). (2017). S-1177 114th Congress: Retrieved from
<https://www.govtrack.us/congress/bills/114/s1177>
- Facione, P., Facione, N., & Giancarlo, C. (1997). *Professional judgment and the disposition toward critical thinking*. The California Academic Press.
- Farrall, M. L. (2012). *Reading Assessment: Linking Language, Literacy, and Cognition*. Wiley.
- 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(8), 835-853.
- Flanagan, D., Fiorello, C., & Ortiz, S. (2010). Enhancing practice through the application of Cattell-Horn-Carroll theory and research: A “third method” approach to specific learning disability identification. *Psychology in the Schools*, 47(1), 739-760.
- Flanagan, D., Ortiz, S., & Alfonso, V. (2013). *Essentials of cross-battery assessment (3rd ed.)*. Hoboken, NJ: John Wiley & Sons.
- Floyd, R., Evans, J., & McGrew, K. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools*, 40, 155-171.
- Floyd, R., Meisinger, E. Gregg, N., & Keith, T. (2012). An explanation of reading comprehension across development using models of Cattell-Horn-Carroll theory: Support for integrative models of reading. *Psychology in the Schools*, 49(8), 725-743.
- Frijters, J., Lovett, M., Steinbach, K., Wolf, M., Sevcik, R., & Morris, R. (2011). Neurocognitive predictors of reading outcomes for children with reading disabilities. *Journal of Learning Disabilities*, 44(2), 150-166.
- Gatlin-Nash, B., & Dockterman, D. (2020). Supporting Early Language Development for Diverse Learners. *The Right Start for Early Childhood*, 16(6).
https://scholar.harvard.edu/files/ddockterman/files/supporting_early_language_development_for_diverse_learners.pdf
- Gonzalez, V., Castellano, J. A., Bauerle, P., & Duran, R. (2013). Attitudes and behaviors toward testing-the limits when assessing LEP students: results of a NABE-sponsored national

- survey. *The Bilingual Research Journal*, 20(3&4), 433-463.
<http://doi.org/10.1080/15235882.1996.10668638>
- Gregory, T., Dal Grande, E., Brushe, M. *et al.* (2021). Associations between School Readiness and Student Wellbeing: A Six-Year Follow Up Study. *Child Indicators Research*, 14, 369–390.
<https://doi.org/10.1007/s12187-020-09760-6>
- Gresham, F. M., & Vellutino, F. R. (2010). What is the role of intelligence in the identification of specific learning disabilities? Issues and clarifications. *Learning Disabilities Research & Practice*, 25(4), 194-206.
- Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experiences of Young American Children*. Baltimore, MD: Brookes Publishing.
- Hass, M. R., & Carriere, J. A. (2014). *Writing Useful, Accessible, and Legally Defensible Psychoeducational Reports*. New York: Wiley.
- Hass, M.R., & Leung, B.P. (2021). When you can't R.I.O.T., R.I.O: Tele-assessment for school psychologists. *Contemporary School Psychology*, 25, 33-39.
<https://doi.org/10.1007/s40688-020-00326-5>
- Holman, S. B., & Stephens, T. (2021). Conducting a comprehensive bilingual evaluation using C-SEP. *Beyond the Score Webinar Series*.
- Holman, S., Stephens-Pisecco, T., Schultz, E., & Moon, G. (2018). *C-SEP Dyslexia Assessment: From Referral to Decision*.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing. An Interdisciplinary Journal*, 2(2), 127–160. <https://doi.org/10.1007/BF00401799>
- Hopkins, M., Martinez-Wenzl, M., Aldana, U.S., & Gandara, P. (2013). Cultivating capital: Latino newcomer young men in a U.S. urban high school. *Anthropology and Education Quarterly*, 44(3), 286-303. <https://doi.org/10.1111/aeq.12026>
- Individuals with Disabilities Education Act (IDEA) regulations. (2006). 34 C.F.R. Part 300.
- Individuals with Disabilities Education Act (IDEA) regulations. (2004). 34 C.F.R. Part 300.
- Individuals with Disabilities Education Improvement Act of 2004 (IDEA). (2004). Pub. L. No. 102-446, 118 Stat. 2647.
- International Dyslexia Association (IDA). (2002). Definition of Dyslexia. Available at <https://dyslexiaida.org/definition-of-dyslexia/>
- Jaffe, A. (Ed.). (2009). *Stance: Sociolinguistic perspectives*. Oxford University Press.
- Johnson, R., & Onwuegbuzie, A. (2007). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Kaufman, A. S., & Kaufman, N. L. (2018). *Kaufman Assessment Battery for Children—Second Edition (KABC-II)*. Administration Manual. NCS Pearson.
- Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement, Third Edition (KTEA-3)*. Administration Manual. NCS Pearson.
- Katz, L. J., & Slomka, G. T. (2000). Achievement testing. *Handbook of psychological assessment*,

- 149-182.
- Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and instruction*, 25, 95-103.
- Kwiatek, R., & Schultz, E. K. (2014). Using informal assessment data to support the diagnosis of a specific learning disability. *The DiaLog*, 43(1), 12-15.
- Leech, N., & Onwuegbuzie, A. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, 22(4), 557-584.
- Leung, B. (1993). Back to basics: Assessment is a R.I.O.T.!. *NASP Communique*, 22(3), 1-6.
- Lewis, T. J., Hatton, H. L., & Maynard, D. (2017). What Beginning Special Educators Need to Know About Conducting Functional Behavioral Assessments. *Teaching Exceptional Children*, 49(4), 231-238. <https://doi.org/10.1177/0040059917690885>
- Mather, N., & Jaffe, L. E. (2016). Woodcock-Johnson IV: Reports, recommendations, and strategies. John Wiley & Sons, Inc.
- Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, 20(3), 142-155.
- McGill, R.J., Styck, K.M, Palomares, R., & Hass, M. R. (2016). Critical issues in specific learning disability identification: What we need to know about the PSW model. *Learning Disability Quarterly*, 39(3), 159-170.
- McGrew, K. S. (2023). Carroll's Three-Stratum (3S) Cognitive Ability Theory at 30 Years: Impact, 3S-CHC Theory Clarification, Structural Replication, and Cognitive–Achievement Psychometric Network Analysis Extension. *Journal of Intelligence*, 11(2), 32.
- 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), 1-10.
- McGrew, K., & Wendling, B. (2010). CHC cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, 47(7), 651-675.
- McGrew, K., LaForte, E., & Schrank, F. (2014). Technical Manual. *Woodcock-Johnson IV*. Rolling Meadows, IL: Riverside Publishing.
- McMillian, J., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry*. Upper Saddle River, NJ: Pearson.
- Mitkovska, L. (2016). Error analysis in the English language classroom. *Vizione*, 26, 25-31.
- 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: Wiley.
- Naglieri, J. A. (1999). *Essentials of CAS Assessment*. Hoboken: John Wiley & Sons, Inc.
- National Association of School Psychologists (NASP). (2010). *Principles for Professional Ethics*.
- National Institutes of Health (NIH). (2021). *Speech and Language Development Milestones*. <https://www.nidcd.nih.gov/health/speech-and-language>
- NCS Pearson. (2020). *Wechsler individual achievement test* (4th ed.) (WIAT-IV). Manual.

- Pearson.
- Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell-Horn-Carroll-based assessment. *Psychology in the Schools*, 47(7), 621 – 634.
- Ortiz, S. O. (2019). On the measurement of cognitive abilities in English learners. *Contemporary School Psychology*, 23(1), 68–86. <https://doi.org/10.1007/s40688-018-0208-8>
- Peterson, R. L., Boada, R., McGrath, L. M., Willcutt, E. G., Olson, R. K., & Pennington, B. F. (2017). Cognitive prediction of reading, math, and attention: Shared and unique influences. *Journal of learning disabilities*, 50(4), 408-421.
- Phillips, S. E., & Clarizio, H. F. (1988). Limitations of standard scores in individual achievement testing. *Educational Measurement: Issues and Practice*, 7(1), 8-15.
- Portland Public Schools. (2015). Identification of Specific Learning Disabilities via Patterns of Strengths and Weaknesses. PSW Manual. <https://www.pps.net/Page/3002>
- Raiford, S.E., Drozdick, L., Zhang, O., & Zhou, X. (2015). *WISC-V Technical Report #1: Expanded Index Scores*. NCS, Pearson. Retrieved from http://downloads.pearsonclinical.com/images/assets/wisc-v/wisc-vtechreport1_fnl_v2.pdf
- Random House. (2011). Pattern. Webster's Unabridged Dictionary.
- Salvia, J., Ysseldyke, J., & Bolt, S. (2010). *Assessment in special and inclusive education*. Belmont, CA: Wadsworth/Cenage.
- Sattler, J M. (2016). *Assessment of Children: WISC-V and WPPSI-IV*. La Mesa: Sattler, Publisher, Inc.
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy*. New York: Guilford Press.
- Schalock, R. & Luckasson R. (2005). *Clinical judgment*. American Association of Mental Retardation: Washington, DC.
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-144). New York, NY: Guilford Press.
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Test of Achievement*. Rolling Meadows, IL: Riverside.
- Schrank, F. A., Stephens-Pisecco, T., & Schultz, E. (2017). *The WJ IV Core-Selective evaluation process applied to identification of a specific learning disability*. (Woodcock-Johnson IV Assessment Services Bulletin No. 8). Itasca, IL: Houghton Mifflin Harcourt.
- Schultz, E., & Stephens, T. L. (2017). Using the core-selective evaluation process (C-SEP) to identify a pattern of strengths and weaknesses. *The Journal of Texas Educational Diagnostician Association*, 46(1), 9-15.
- Schultz, E., & Stephens, T. L. (2015). Core-selective evaluation process: An efficient and

- comprehensive approach to identify SLD using the WJ IV. *The Journal of Texas Educational Diagnostician Association*, 44(2), 5-12.
- Schultz, E., & Stephens, T. L. (2009). Utilizing professional judgment within the SLD eligibility determination process: Guidelines for educational diagnosticians and ARD committee members. *The DiaLog*, 38, 3-6.
- Schultz, E., & Stephens, T. L. (2008). SLD identification: An analysis of state policies 2008-2009. Manuscript submitted for publication. Texas Education Agency (2008). Special education rules and regulations. <https://www.tea.state.tx.us/special.ed/rules/sbs.html>
- Schultz, E., Rutherford, E., & Cavitt, D. (2021). *Intellectual development and specific learning disability: The role of norm-referenced tests*. Special Education Research, Policy & Practice. Hofstra University.
- Schultz, E., Simpson, C., & Lynch, S. (2012). Specific learning disability: What constitutes a pattern of strengths and weaknesses? *Learning Disabilities: A Multidisciplinary Journal*, 18(2), 87-97.
- Schrank, F.A., McGrew, K.S., & Mather, N. (2014). *Woodcock-Johnson IV*. Rolling Meadows, IL: Riverside Publishing.
- Schrank, F., 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. Itasca, IL: Houghton Mifflin Harcourt.
- Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, 42(2), 137.
- Stephens, T. L., & Moon, G. (2020). Exclusionary factors checklist. Adapted for COVID-19.
- Stephens, T. L., & Moon, G. (2014). Exclusionary factors checklist. Adapted from Stephens et al. (2013).
- Stephens, T. L., Dykes, F., Proctor, C., Moon, G., Gardner, R., & Pethick, L. (2013). Ruling out exclusionary factors through the utilization of a response-to-intervention (RTI) model. *The DiaLog*, 42(1), 5-14.
- Stephens, T. L., Olvera, P. & Schultz, E.K. (2022). Conducting a Targeted SLD Assessment Using the Core-selective Evaluation Process (C-SEP): Implications for Assessing English Learners. *Contemporary School Psychology*. <https://doi.org/10.1007/s40688-022-00443-3>
- Stephens, T. L., Proctor, C., Moon, G., Gardner, R., & Pethick, L. (2013). Ruling out exclusionary factors through the utilization of a response-to-intervention (RTI) model. *DiaLog*, 42(1), 5-14.
- Stephens, T. L., Schultz, E. K., Holman, S. B., & Moon, G. (2018). The Core-Selective Evaluation Process: Overview and Procedures. Dragonfly Publishing, TX: Houston.
- Stephens-Pisecco, T. L., & Schultz, E. (2018). *C-SEP Handbook*. Dragonfly.
- Talosa, A. D., & Maguidayao, R. N. (2018). Evaluation of second language learners' syntactic

- errors in ESL writing. *TESOL International Journal*, 13(4), 172-181.
<https://files.eric.ed.gov/fulltext/EJ1244109.pdf>
- U.S. Department of Education, Office of English Language Acquisition. (2017). English learner tool kit (2nd Rev. ed.). Washington, DC: Author.
- van Aken, L., van der Heijden, P. T., Oomens, W., Kessels, R. P. C., & Egger, J. I. M. (2019). Predictive Value of Traditional Measures of Executive Function on Broad Abilities of the Cattell-Horn-Carroll Theory of Cognitive Abilities. *Assessments*, 26(7), 1375-1385.
- Van den Broeck, W. (2002). The misconception of the regression-based discrepancy operationalization in the definition and research of learning disabilities. *Journal of Learning Disabilities*, 35(3), 194-204.
- Wasowicz, J. (2021). A Speech-to-Print Approach to Teaching Reading. *LDA Bulletin*, 53(2), 10-18.
- Watkins, M. W. (2009). Errors in Diagnostic Decision Making and Clinical Judgment. In T. B. Gutkin & C. R. Reynolds (Eds.), *Handbook of School Psychology*, Fourth Edition, (pp. 210-229). Hoboken: Wiley.
- Wechsler Individual Achievement Test, 3rd Edition (WIAT-III). (2009). Technical Manual. San Antonio: Pearson.
- Wechsler, D. (2014). Wechsler Intelligence Scale for Children (5th ed.). Bloomington: PsychCorp.
- Wiig, E. H., Semel, E., & Secord., W. A. (2013). Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5). Bloomington: NCS Pearson.
- Wilkison, G.S., & Robertson, G. J. (2017). Wide Range Achievement Test, 5th Edition (WRAT-5). Pearson.
- Woodcock, R., Alvarado, C. Ruef, M., & Schrank, F. (2017). *Woodcock-Munoz Language Survey*. Rolling Meadows, IL: Riverside Publishing.
- Wright, J. (2010). The RIOT/ICEL Matrix: Organizing Data to Answer Questions About Student Academic Performance and Behavior. How RTI Works. Intervention Central.

Page Intentionally
Left Blank

C-SEP™ Legal Alignment

34 Code of Federal Regulations	C-SEP™ Framework/Guidance
The child must be <u>assessed</u> in all areas of suspected disability. (4) The child is assessed in all areas related to the suspected disability, including, if appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities.	The federal law is explicit in that <i>all areas of suspected disability should be assessed</i>. C-SEP™ is comprehensive, as it requires the integration of formal assessment results with multiple other data sources to include all areas related to suspected disability. ** Districts can decide whether to mandate testing in all areas or just in the areas of concern.
SLD is a disorder in one or more of the basic <i>psychological processes</i> involved in <i>understanding or in using language</i> that is <i>spoken or written</i>, that may manifest itself in the <i>imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations...</i> (34 CFR § 300.8(c)(10).	C-SEP™ is comprehensive and all areas of SLD are assessed/considered using multiple sources of data: • Language, • Cognition, and • Achievement C-SEP™ recognizes the importance language development plays in Cognition and Achievement.
<i>Exclusionary Factors</i> must be ruled out as the primary cause of academic struggle (§ 300.8).	C-SEP™ requires that all <i>exclusionary factors</i> be considered and ruled out as the primary cause of SLD. Preliminary rule-out occurs during the Review Stage when multiple sources of data are organized and considered. Documentation is necessary to support rule out.
<i>“Prior to and part of the [SLD] evaluation....”</i> (b) To ensure that <u>underachievement</u> in a child suspected of having a specific learning disability is <i>not due to lack of appropriate instruction in reading or math</i>, the group must consider, as part of the evaluation described in §300.304 through § 300.306— (1) <i>Data</i> that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel; and	• Multiple sources of data collected prior to and as part of the evaluation are considered and used to establish underachievement. • Quality of instruction is considered (e.g., did a certified teacher deliver instruction, what curriculum was used, was the child homeschooled, did the child have excessive moves). • Was instruction conducted in the student’s dominant language? • Attendance is considered. Were absences or tardies excessive?

(2) <i>Data-based documentation</i> of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction, which was provided to the child’s parents (§ 300.309).	- RTI data and progress monitoring data are considered and incorporated into data analysis. - Consider the student’s performance in relation to peers (e.g., is this an individual problem or a class-wide problem).
34 Code of Federal Regulations § 300.309	C-SEP™ Framework/Guidance
(a) For a child suspected of having a specific learning disability, <i>the documentation of the determination of eligibility</i>, as required in §300.306(a)(2), must contain a statement of— (7) If the child has participated in a process that assesses the child’s response to scientific, research-based intervention— (i) The instructional strategies used and the student-centered data collected.	C-SEP™ is an assessment model that is comprehensive and <i>merges data collected</i> from the RTI process. List and description of interventions implemented (duration & frequency). Progress monitoring data (ROI).
The public agency must ensure that the child is <i>observed in the child’s learning environment</i> (including the regular classroom setting) to document the child’s academic performance and behavior in the areas of difficulty.	<i>C-SEP™ integrates various forms of observation data with other data points</i> when interpreting evaluation results.
May <i>not use any single measure or assessment as the sole criterion</i> for determining whether a child is a child with a disability and for determining an appropriate educational program for the child.	C-SEP™ uses multiple measures and integrated data analysis to ensure that SLD identification is based on multiple criteria. <i>Standard scores are NOT used as the sole criterion for SLD identification.</i>
Use a <i>variety of assessment tools and strategies</i> to gather relevant functional, developmental, and academic information about the child. Including information provided by the parent, which may assist in determining whether a child has a disability; and use it for individualized educational planning.	C-SEP™ requires assessment results to be used for program planning. Adequate data is collected to assist in establishing current academic functioning and developing IEP goals and objectives.
Assessments and other evaluation materials used to assess a child under this part— (i) Are selected and administered so as not to be discriminatory on a racial or cultural basis;	Characteristics of race, culture, socioeconomic status, and language acquisition are considered when choosing assessment instruments and when conducting evaluations.

34 Code of Federal Regulations § 300.309	C-SEP TM Framework/Guidance
Use <i>technically sound instruments</i> that may assess the relative contribution of cognitive and behavioral factors, in addition to physical or developmental factors. The child is assessed in all areas of suspected disability.	C-SEP TM procedures require adherence to the publisher's/author's administration and interpretive guidelines. In addition, publisher software and/or norm tables are used to score tests.
Tests are <i>administered in accordance</i> with any <i>instructions provided by the producer</i> of the assessments.	C-SEP TM procedures rely on the technical specifications of the test publisher/author for norm-referenced tests to ensure reliability and validity.
The child <i>exhibits a pattern of strengths and weaknesses in performance, achievement</i> , or both, relative to age, State-approved grade-level standards, or intellectual development, which is determined by the group to be relevant to the identification of a specific learning disability, using appropriate assessments, consistent with §300.304 and §300.305.	C-SEP TM uses integrated data analysis to determine a pattern of strengths and weaknesses. C-SEP TM utilizes pattern seeking strategies consider the student's performance across multiple data sources over an extended period of time.

Page Intentionally
Left Blank

Resource Links

Resource	A digital copy is available at:
Multiple Sources of Data Worksheet (Holman, 2017)	www.csep.online/documents/Multiple_Sources_Data.pdf
Exclusionary Factors Checklist (Stephens and Moon, 2022)	www.csep.online/documents/Exclusionary_Factors_Checklist.pdf
Portland Public Schools (Portland Public Schools, 2015)	www.csep.online/documents/PPS_Manual.pdf
Language Demands Worksheet	www.csep.online/documents/Language_Demands_Checklist.pdf
For a comprehensive collection of useful resources, consider: Sarah Holman's <i>C-SEP™ Supplemental Resources</i>, available on Amazon.	

Page Intentionally
Left Blank

C-SEP™ Quick Start (Testing Procedures)

Responsible use of norm-referenced tests underlines C-SEP™ testing procedures. When identifying specific learning disabilities, we must remember that norm-referenced tests are informative, not determinative. Test selection (both core and elective) is informed by the publisher's test manuals and the referral question. Keep in mind the purpose of a norm-referenced tests in C-SEP™:

1. We are using norm-referenced testing to assess **"intellectual development"** which is comprised of: our innate cognitive ability, language, academic skills and environment.
2. We assess intellectual development to help us understand and explain the student's underachievement.

Step 1: Select and administer a Core Set of Cognitive tests from the manual.

WJ Cog	WISC-V (age 6-16)	KABC (age 7-18)
1. Oral Vocabulary 2. Number Series 3. Verbal Attention 4. Letter-Pattern Matching 5. Phonological Processing 6. Story Recall 7. Visualization	1. Block Design 2. Similarities 3. Matrix Reasoning 4. Digit Span 5. Coding 6. Vocabulary 7. Figure Weights *You may give these as core or use these to follow-up on a low score. 8. Visual Puzzles 9. Picture Span 10. Symbol Search	1. Number Recall 2. Word Order 3. Story Completion 4. Pattern Reasoning 5. Atlantis 6. Rebus 7. Rover 8. Triangles for age 7-12 or Block Counting for age 13-18. 9. Verbal Knowledge 10. Riddles

Step 2. Select and administer a core set of language tests.

WJ Oral Language	WIAT 3	KTEA -3
1. Picture Vocabulary 2. Oral Comprehension 3. Segmentation 4. Rapid picture Naming	1. Listening Comprehension (Receptive Vocabulary, Oral Discourse) 2. Oral Expression (Expressive Vocabulary, Oral Word Fluency, Sentence Repetition)	1. Listening Comprehension (test 16) 2. Oral Expression (test 18)

Step 3. Select and administer a core set of “achievement tests” (cognitive process paired with an academic skill, and equally acts as a language measure). **Only test in suspected areas.**

Reading Core	Math Core	Writing Core
1. Basic Reading 2. Reading Fluency 3. Comprehension	1. Calculation test 2. Applied Problems	1. Spelling 2. Written Expression.

Step 4. Score the core set of tests and analyze any global score to either rule out or consider disabilities (e.g., Intellectual disabilities, language disorders).

Step 5. If you have scores that show significant variance, this is ONE piece of evidence that “this child does not have typical intellectual development” and has a pattern of strengths and weakness. Scores such as these are consistent with a student with SLD, but alone do not indicate a student has SLD.

Step 6. Interpret the scores.

- a) Average or high scores can be interpreted to mean the student has met the cognitive, linguistic, and skills demands of that set of tests. You can **generally** infer that a student would be able to complete a similar task with less cognitive demands.
- b) Low or below average scores mean that an area of “intellectual development” may be weak (“disorder in one or more of the basic psychological processes involved in using language”). **Low scores must be explained and followed up.** Look within your test data using the task demands to corroborate a low score. Examine the informal data set and this is where the disability “manifests itself.” Selective testing procedures may also be used to corroborate a low score, but make sure to look within your test data and informal data prior to using selective testing procedures to avoid over-testing.

Step 7. Cluster the low scores (all of them) and examine the task demands. Answer these questions.

- 1. What do they have in common? Do they struggle on fluency tasks, memory tasks, listening tasks, auditory memory tasks, comprehension tasks, etc.?
- 2. Is there a similarity or difference in “inputs” (auditory, visual) or “outputs” (motor, oral)?
- 3. Are the tasks they struggle with like the cognitive, linguistic and skill demands of the referral question (reading, math, writing, listening)?
- 4. Are there any low scores that correlate to some average or high scores (e.g., visual memory vs auditory memory, oral comprehension vs reading comprehension, spelling vs reading decoding, quantitative reasoning vs math problem solving)? **Explain the split using task demands analysis.**
- 5. Consider the low scores in context of all the other data. Evaluators sometimes “over-value” low scores so it is important to put these scores in context with the other data (formal and informal) before labeling it a weakness. Remember, all tasks require multiple abilities and strengths in other areas may offset or reduce the effects of the weakness.

Step 8. Integrate the formal data with the informal data set. Some questions you want to ask when looking at this combination of data is:

- a. Is the informal data consistent with the norm-referenced data?
- b. Are these low areas of testing manifesting themselves in the classroom?
- c. Does this formal data help explain why a student is not meeting state standards?

Step 9. If after steps 1-8, you cannot draw a conclusion or need additional information (for example, the data showed many indicators of dyslexia, or you want to further explore a low score), use selective testing procedures and give additional tests to improve your understanding of the child. When select testing, keep in mind:

1. The core set of tests you administered gave you the most “bang for your buck” as they adequately cover the construct of “intellectual development” and are the most cognitively complex.
2. We do not want to over-rely on norm-referenced testing.
3. You have an informal data set, which includes “instructional response” data as it represents “teaching and learning.” The data you have combined with a poor instructional response may be sufficient.
4. When selecting additional selective tests, have a clear rationale for why you are selecting tests. Here are some examples:
 - a. I want to see how his visual memory is compared to his auditory memory. This is referred to as a “planned comparison.”
 - b. This student profile had several indicators of dyslexia during testing, so I am going to administer “word attack,” “nonsense word decoding,” and “spelling of sounds.”
 - c. This student scored low on “coding,” but I am unsure he understood the directions; I will give “symbol search.”
 - d. I want to include more non-verbal measures due to this student's language development.
5. You have plenty of data from the core to determine whether the student's intellectual development is “typical” or “atypical.”
6. Selective testing should be used to “inform educational decisions” by understanding the learner.

Step 10. If you have followed these procedures, you have completed the “testing” aspects of “Intellectual development” and C-SEP™. Your next step is to combine these findings with the preponderance of data required to determine SLD.