

8. Gathering Data for Comprehensive Evaluation

Contents of this Section

Chapter Overview	2
Regulations and Rules	2
Quality Practices Determining Service and Education Requirements	5
Quality Practices in Using a Problem-Solving Protocol to Design the Comprehensive Evaluation	7
Major Sources of Data	14
Comprehensive Achievement Batteries	22
Quality Practice in Collecting Information Processing Data	24
Collecting Data on Cognitive or Intellectual Functioning	31
Designing Comprehensive Evaluations for Young Children	33
Designing Comprehensive Evaluations for Learners with Cultural and Language Differences	36
Next Steps	43
Appendix	44
References	48

Chapter Overview

Comprehensive evaluation involves intensive and comprehensive problem-solving that leads to a special education eligibility determination. Evaluations should be grounded in theory, driven by specific hypotheses, and tailored to each student. Data from discrepancy scores or scientific research-based interventions may be considered in the determination evaluation, but should not be used as the lone determinant.

Beginning with a comprehensive presentation of all laws pertaining to data gathering, this chapter discusses sources of data, provides guidance on determining service and education requirements, and provides sections relating to young children and English Language Learners. It contains various tools, such as FAQs, and suggested achievement and cognitive measures to help teams complete this step.



Regulations and Rules

Note: Regulations, statutes, and rules form the basis for legal compliance and are provided here to help readers understand what the law requires. The following regulations, rules, and statutes govern practices of data collection and the design of the comprehensive evaluation.

Comprehensive Evaluation

Full requirements for comprehensive evaluation are covered in 34 C.F.R. section 300.301 through 300.306. The most relevant requirements to the content in this chapter have been included.

- 34 C.F.R. section 300.304(b) The public agency must:
 - Use a variety of assessment tools and strategies to gather relevant functional, developmental, and academic information about the child, including information provided by the parent that may assist in determining:
 - Whether the child is a child with a disability.
 - The content of the child's IEP, including information related to enabling the child to be involved in and progress in the general education curriculum.
 - Not use any single measure or assessment as the sole criterion for determining whether a child is a child with a disability and for determining an appropriate educational program for the child.
 - Use technically sound instruments that may assess the relative contribution of cognitive and behavioral factors in addition to physical or developmental factors.
 - Ensure that the evaluation is sufficiently comprehensive to identify all of the child's special education and related services needs, whether or not

commonly linked to the disability category in which the child has been classified.

- Minnesota Rule 3525.2710 subpart b (2): In conducting the evaluation, a district shall not use any single procedure as the sole criterion for determining whether a child is a pupil with a disability or determining an appropriate education program for the pupil.

Evaluation Materials and Procedures

- 34 C.F.R. section 300.304(c)(1): Each public agency must ensure that:
 - Assessments and other evaluation materials used to assess a child under this part are:
 - Selected and administered as to not be discriminatory on a racial or cultural basis.
 - 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.
 - Used for the purpose for which the assessments or measures are valid and reliable.
 - Administered by trained and knowledgeable personnel.
 - Administered in accordance with any instructions provided by the producer of the assessments.
 - Assessments and other evaluation materials used include those tailored to assess specific areas of educational need and not merely those that are designed to provide a single general intelligence quotient (34 C.F.R. § 300.304(c)(2)).
 - Assessments are selected and administered so as best to ensure that, if an assessment is administered to a child with impaired sensory, manual, or speaking skills, the assessment results accurately reflect the child's aptitude or achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual, or speaking skills, unless those skills are the factors that the test purports to measure (34 C.F.R. § 300.304(c)(3)).
 - 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 C.F.R. § 300.304 (c)(4)).
 - Assessments tools and strategies that provide relevant information that directly assists persons in determining the educational needs of the child are provided (34 C.F.R. § 300.304 (c)(7)).

- Additional procedures defined in Minnesota Rule 3525.2710 c (2): Each district shall ensure that materials and procedures used to evaluate a child with limited English proficiency are selected and administered to ensure that they measure the extent to which the child has a disability and needs special education and related services, rather than measure the child's English language skills.

Variance from Standard Evaluation Conditions

Minnesota Rule 3525.2710, subp. 3(c)(6): If an evaluation is not conducted under standard conditions, a description of the extent to which it varied from standard conditions must be included in the evaluation report.

Review of Existing Evaluation Data

- 34 C.F.R. section 300.305(a) As part of an initial evaluation, if appropriate, and as part of any reevaluation under this part, the IEP Team and other qualified professionals, as appropriate, must:
 - Review existing evaluation data on the child including:
 - Evaluations and information provided by the parents of the child.
 - Current classroom-based local or state assessments and classroom-based observations.
 - Observations by teachers and related service providers.
 - On the basis of the review, and input from the pupil's parents, identify what additional data, if any, are needed to determine:
 - Whether the pupil has a particular category of disability, as described in as defined in section 300.8, and the educational needs of the child OR in the case of a reevaluation of a child, whether the child continues to have such a disability and educational needs of the child.
 - Whether the child needs special education and related services, OR in the case of a reevaluation of a pupil, whether the child continues to need special education and related services.
 - Whether any additions or modifications to the special education and related services are needed to enable the child to meet the measurable annual goals set out in the individualized education program of the child and to participate, as appropriate, in the general curriculum.
- 34 C.F.R. section 300.305(c) Sources of data: The public agency must administer such assessments and other evaluation measures as may be needed to produce the data identified under subpart a.
- 34 C.F.R. section 300.305(d):
 - Requirements of additional data are not needed if the IEP team and other qualified professionals, as appropriate, determine that no additional data are

needed to determine whether the child continues to be a child with a disability, and to determine the child's educational needs, the public agency must notify the child's parents:

- Of that determination and the reasons for the determination.
- Of the right of the parents to request an assessment to determine whether the child continues to be a child with a disability, and to determine the child's educational needs.
- That the public agency is not required to conduct the assessments previously described in unless requested to do so by the child's parents.

Secondary Transition Needs

Minnesota Rule 3525.2900, subp. 4(A): For each pupil, the district shall conduct an evaluation of secondary transition needs and plan appropriate services to meet the pupil's transition needs. The areas of evaluation and planning must be relevant to the pupil's needs and may include work, recreation and leisure, home living, community participation, and postsecondary training and learning opportunities. To appropriately evaluate and plan for a pupil's secondary transition, additional IEP team members may be necessary and may include vocational education staff members and other community agency representatives, as appropriate.

Use of Assessments Transferred from Other Public Schools

34 C.F.R. section 300.304(c)(5): Assessments of children with disabilities who transfer from one public agency to another public agency in the same academic year are coordinated with those children's prior and subsequent schools, as necessary and as expeditiously as possible, consistent with section 300.301(d)(2) and (e), to ensure prompt completion of full evaluations.



Quality Practices Determining Service and Education Requirements

In order for teams to conclude that a student is eligible for special education due to an Specific Learning Disability, the disability must meet eligibility criteria under 34 C.F.R. section 300.309. During the required comprehensive evaluation, teams must also determine the educational and/or related service needs of the student. Finally, teams use the data to determine:

- The student's continuing educational needs and the instruction that will address the student's needs.
- Any factors that contribute to poor performance (e.g., mobility, untreated vision problems, English language acquisition).
- If more than one disability is indicated, identify the primary and co-existing disability(ies).

OR

- Any educational needs that must be met through accommodations or modifications, and special education services.

- The next steps to meet the student's instructional needs if the student is not determined to have a disability and require specially designed instruction under IDEA 2004 or does not have a disability and require modifications under 504.

Criteria and Sources of Data Used In Decision Making

Use of the discrepancy formula or data from research-based interventions alone is insufficient to accurately identify a student as having an SLD. A discrepancy score disconnected from an understanding of how a student functions in a classroom and responds to quality instruction is insufficient to address the questions put forth in the eligibility determination. Data from interventions are important for extracting information about many of the exclusionary variables that can affect learning in the classroom, notably poor or inappropriate instruction, cultural bias, issues of language acquisition, etc. However, data illustrating a child's response to interventions is insufficient to generate comprehensive evaluation of a child's achievement and a hypothesis for the learning difficulty.

Teams will find that data indicating response to intervention, observation data, interviews, and record reviews provides ecological validity to test data gathered during comprehensive evaluation. The process of finding convergence among various sources of data as well as teasing out explanations from divergent data increases accuracy of identification and informs the design of special education services.

The figure below illustrates the two evaluation criteria options and the corresponding types of data required.

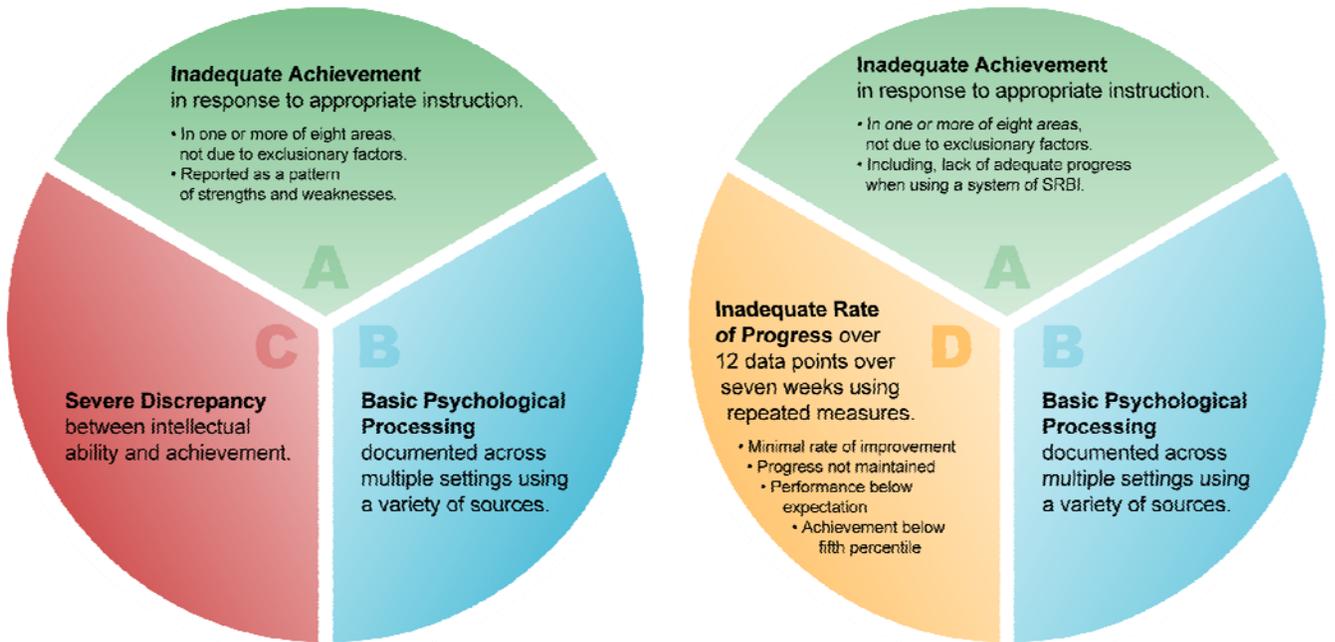


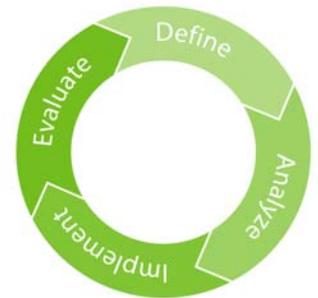
Figure 8-1. Determination Criteria.

Note: See Minnesota SLD Rule Summary in Chapter 1 for more information.



Quality Practices in Using a Problem Solving Protocol to Design the Comprehensive Evaluation

As discussed in Chapter 4, for the system of SRBI process, the determination process includes four iterative steps:



Step 1: Define the Problem. Define the problem and why it is happening.

Step 2: Analyze the Problem. Validate the problem, identify the variables that contribute to the problem and develop a plan.

Step 3: Implement the Plan. Carry out the intervention as intended.

Step 4: Evaluate the Plan. Determine whether the data indicate the plan is working (for more information see chapter 5 for further discussion of monitoring progress).

During comprehensive evaluation, teams should follow the same process, but use different tools such as formal tests and measures of achievement and cognitive abilities. The assessment plan should be informed by data gathered prior to the evaluation planning

meeting. The more that teams are able to integrate existing data, the more efficient and individualized the comprehensive evaluation process.

At the point of designing the comprehensive evaluation, teams should thoroughly review the results of attempts to address gaps in achievement, language development, social-emotional, behavioral challenges, physical limitations, and suspected weaknesses in basic psychological processes. Teams will need to redefine the problem, re-examine why the problem persists despite high-quality instruction and intervention as well as reassess what further data needs to be gathered. During the evaluation process, teams must be prepared to integrate the data gathered from formal tests and measures with existing data and analyze the salient findings and relationships between achievement and basic psychological processes.

Figure 8-2 provides a basis for informed decision-making and shows how data gathered from each phase in the eligibility process informs the next step in data collection and decision-making. This framework for problem solving provides one means of systematically analyzing student needs. Districts are encouraged to specify and train staff in their own protocols and tools.

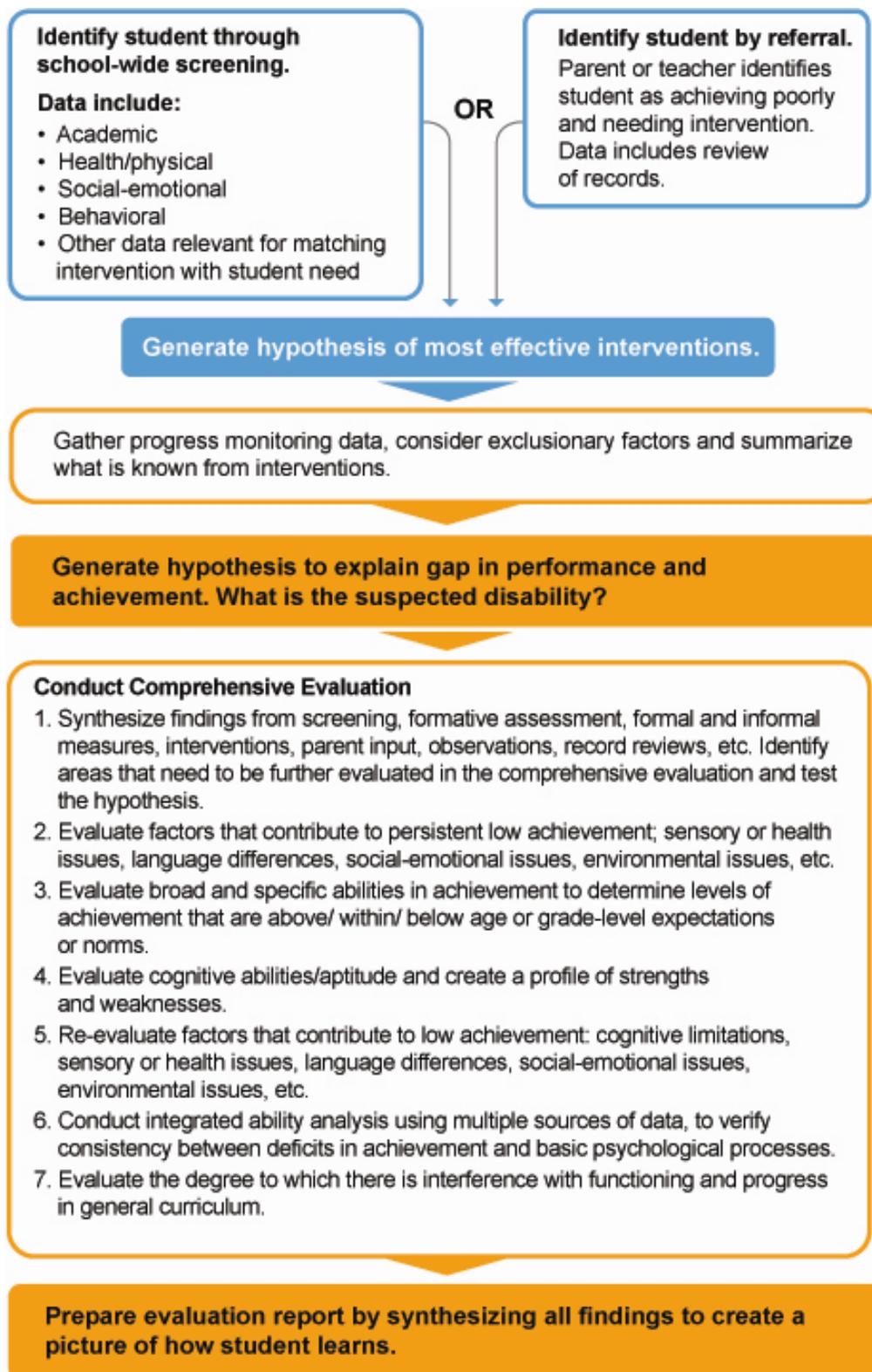


Figure 8-2. Assessment Process.

--Adapted from Operational Definition of Learning Disabilities by Flanagan, D. et al., (2006).

Note: See Chapters 3 and 5 for more information on screening and progress monitoring noted in the figure above.

Whether a team uses the proposed model or another research-based model for organizing the comprehensive evaluation, all those involved in making the eligibility determination must have adequate information to address eligibility, instructional needs, and next steps. Districts may want to investigate other research-based models for organizing an SLD evaluation. Other research-supported models include Concordance-Discordance Model of SLD Determination by Hale and Fiorello’ (2004) and Discrepancy/Consistency Model based on the Planning, Attention, Simultaneous, Successive (PASS) processing theory by Naglieri (1999).



Quality Practices

Questions Guiding the Design of Comprehensive Evaluation and Collection of Data with Corresponding Regulatory Citations

The design of the comprehensive evaluation should be grounded in theory, guided by specific questions and research-informed practices in assessment. Teams will notice that the guiding questions at the end of each chapter are duplicates, but are organized to address the statutory requirements that come with determining eligibility and necessary specially designed instruction and related services. To the extent that existing data has been integrated and used to inform the next step, the data that remains to be gathered may be different for each student. Teams should focus on collecting data that address the persistent and complex educational needs of the student and not be driven by a standardized template or testing kit. The table below provides guidance regarding these issues.

Table 8-1
Questions to Guide Comprehensive Evaluation Design and Collection of Data

Guiding Questions	Core Instruction	Supplemental Intervention	Specialized Instruction (IEP)
Access to high-quality scientific research-based instruction	How has the team determined the student has had sufficient access to high-quality instruction and opportunity to perform within grade-level standards? Minn. R 3525.1341, subp. 1 B	What supplemental efforts, aligned with grade-level standards, were implemented to accelerate the student’s rate of learning and level of performance? Minn. Stat. 125A.56 Minn. R 3525.1341, subp. 2 D	What has and has not worked to increase access and participation in the regular classroom environment? What additional supports, accommodations or modifications are necessary to provide access to grade-level standards?

Guiding Questions	Core Instruction	Supplemental Intervention	Specialized Instruction (IEP)
<p>Limitations in adequate achievement or performance</p> <p>(ELL, lack of appropriate instruction in reading or math)</p>	<p>What areas of educational performance/ achievement continue to be below grade-level expectations?</p> <p>34 C.F.R. § 300.8(a)(2)</p> <p>34 C.F.R. § 300.304(b)(3)</p> <p>34 C.F.R. § 300.304(c)(2)</p> <p>Minn. R 3525.1341, subp. 2 A</p>	<p>What factors limit performance? What supplemental efforts have been successful in mediating the impact?</p> <p>34 C.F.R. § 300.304(c)(1)(ii)</p> <p>What about the student's profile leads the team to suspect a disability and the need for special education and related service supports?</p> <p>34 C.F.R. § 300.306(b)</p>	<p>What special education supports would be sufficiently rigorous to accelerate performance towards grade-level achievement standards?</p> <p>OR</p> <p>Given previous efforts, what additional supports are required to help the student gain control over academic, non-academic, and transition goals?</p>
<p>Impairment/Disability</p> <p>(Sensory, cognitive delay, emotional or behavioral)</p>	<p>How is the student functionally limited from making progress towards grade-level standards?</p> <p>34 C.F.R. § 300.304-306</p> <p>How is the student limited from participating in the five areas of transition: namely, work, recreation and leisure, home living, community participation, postsecondary training and learning opportunities.</p> <p>Minn. R 3525.2900, subp. 4(A)</p>	<p>What evidence is there that indicates the student needs protections afforded through Reauthorized Federal IDEA 2004 for <i>specific learning disability</i> to make progress towards grade-level standards?</p> <p>34 C.F.R. § 300.8(a)(1)</p> <p>34 C.F.R. § 300.304(c)(2)-(7)</p> <p>34 C.F.R. § 300.8(b)</p>	<p>What are all the needs that must be addressed and the evidence-based instruction that will accelerate achievement towards grade-level standards?</p> <p>34 C.F.R. § 300.305(a)(2)</p> <p>34 C.F.R. § 300.304(b)(1)</p> <p>Minn. R 3525.2710, subp. 4 (D)-(E)</p>

The questions are organized from least restrictive environment starting in the top left corner, moving right and down as documentation is gathered to identify the appropriate specially designed instruction. Teams maximizing the guiding questions will have documentation sufficient to meet eligibility requirements, design special education services and develop an individualized education program as well as document the need for participation in the modified assessment.

Table 8-2

Data Collection Best Practice for Culturally and Linguistically Diverse Student Learners

Data Collection Best Practices for Culturally and Linguistically Diverse Learners

Ortiz (2008) outlines revised and refined structural guidelines, which provide a comprehensive framework for engaging in fair and equitable assessment of diverse individuals. A practical framework to guide assessment is found in *Best Practices in School Psychology V* and consists of 10 essential components. The framework below is adopted with permission.

1. Assess for the Purpose of Intervention. An intervention-driven process can be one discriminatory aspect of the assessment and can bias all subsequent activities. The intervention(s) need to provide ways to accelerate acquisition of skills and learning rather than identifying the underlying cause of observed problems.

2. Assess Initially with Authentic and Alternative Assessment Procedures. Intervention-based assessments have value in reducing some of the discriminatory aspects of evaluation as well as improving academic achievement. Interventions and documentation of intervention fidelity assist in assuring progress in skill development and reflect what the student has been taught. Implementation of a proper response to intervention framework that is culturally and linguistically appropriate is can be a rigorous approach using authentic methods.

3. Assess and Evaluate the Learning Ecology. An exploration of extrinsic causes that might be related to learning difficulties should occur prior to exploration of intrinsic factors like ability. Assessment of culturally and linguistically diverse students is often related to experiential factors. Acculturation and differences in language are equally important to consider. Additional differences impacting all students might include health, family situations, socioeconomic issues, teacher biases, and access to effective instruction, to name a few.

4. Assess and Evaluate Language Proficiency. For dual language learners, assessment of language proficiency in both languages for Basic Interpersonal Communication (BICS) and Cognitive Academic Language Proficiency (CALP) must be current (within 6 months) and is crucial for development of appropriate linguistically interventions. This information addresses questions such as opportunity to learn, expected level of functioning relative to English language development, etc.

5. Assess and Evaluate Opportunity for Learning. The educational system, including the curriculum, personnel policies, instructional setting, etc., must be carefully evaluated to determine whether the student has been provided with adequate opportunity to learn. Some of the factors to consider are parent interview, regularity of school attendance, match between native language in instruction and parents' ability to support language instruction, culturally appropriate instruction and curriculum, etc.

6. Assess and Evaluate Educationally Relevant Cultural and Linguistic Factors. Many factors outside of school should be assessed because of their possible influence on student learning and language development. The effects of small amounts of exposure to two or more languages or cultures during early childhood development may create circumstances that impact school performance. Observations across multiple environments and observers, interviews, and review of records are a few of the multiple methods and sources of information that should be accessed.

Data Collection Best Practices for Culturally and Linguistically Diverse Learners

7. Evaluate, Revise, and Retest Hypotheses. The convergence of data and multiple information sources should be thoroughly evaluated. Systematic interventions need to be carefully analyzed. Sometimes external factors may be present but not directly contributing to learning difficulties. When there are no plausible or demonstrable external factors that can account for the learning difficulties, then consideration of intrinsic factors is warranted.

8. Determine the Need for Language(s) of Assessment. IDEA 2004 mandates that assessors consider the student's primary language ability (in addition to English ability) in the development of the assessment plan. Factors that influence test selection are based on information collected from steps 1-7 above as well as other relevant outside data. Although each case is individual, basic guidelines are that students **who are not** proficient in English should be assessed in the primary language in addition to any English testing that may be appropriate, and students who are proficient in English may be assessed in their primary language in addition to any English testing that may be appropriate. All students, whether proficient in English or not, whose histories and backgrounds are not comparable to U.S. mainstream, should be evaluated by an assessor who possesses knowledge of the factors relevant to the student's unique experiences and how they may effect learning.

9. Reduce Bias in Traditional Testing Practices. The process of nondiscriminatory assessment using tests is represented in two distinct options: (a) administer test(s) in a standardized way and attempt to evaluate the results in a nondiscriminatory manner, or (b) modify the testing process in a way that is less discriminatory initially. Rationale for each is summarized below.

(a) Maintaining standardization allows application of systematic methods to reduce bias.

- Use locally developed, pluralistic norms.
- Provide a foundation for nondiscriminatory assessment based on research/empirical evidence. For example, the Culture Language Interpretive Matrix (included in the appendix).
- Use a knowledge of test properties relative to cultural loading and linguistic demands as the basis for test selection.

(b) Modification and adaption of tests to reduce the effect of acculturation or linguistic bias violates standardization and negates the validity and interpretability of results for quantitative data. Therefore, protocols should not be scored and no quantitative data reported. Tests may provide qualitative information, but should remain guided by efforts to intervene and not to diagnosis.

10. Support Conclusions Via Data Convergence and Multiple Indicators. A convergence of data from multiple sources, including the student's unique experience and background, should be integrated and used as the appropriate context from which to evaluate the data. The data collected should come together in a cohesive and convincing manner that supports the plausibility of the final conclusion. A convergence of evidence is sufficient to provide validity to conclusions, but care should be taken not to assign unwarranted significance to any single piece of evidence. In the final analysis, equivocal data should be interpreted as the learning problem is not intrinsic to the learner, but that functioning is within normal limits. Any observed difficulties are the result of factors other than those related to a disability.

Ortiz (2008) found in *Best Practices in School Psychology V*. National Association of School Psychology

Major Sources of Data

This section discusses the major sources of data that may be collected to meet each of the criteria for SLD determination, namely inadequate achievement (including intervention data), information processing, and IQ (for discrepancy). Teams should consider how data will be gathered so that any area of concern identified through the evaluation has multiple sources of data confirming/validating the deficit. Ideally, teams will have three independent pieces of data confirming the area of deficit.

See Chapter 7 for more information on areas of inadequate achievement and academic functioning relevant for SLD determination.

Reminder: *Academic functioning below age or grade-level standards is required for eligibility under SLD.*

Collecting Achievement Data

In order to document achievement for the eligibility determination and to develop instruction after the eligibility decision is made, the team should collect data on the following:

- Listening comprehension.
- Oral expression.
- Basic reading skills.
- Reading comprehension.
- Reading fluency.
- Written expression.
- Math calculation.
- Mathematical problem solving.

Teams should note the differences in how the achievement data should be documented for the choice in criteria being used. In cases where the discrepancy (criteria ABC) is being used, the achievement must be reported as a pattern of strengths and weaknesses. In cases where the lack of response to instruction (criteria ABD) is to be used, the data indicating lack of response must be documented.

Currently, there is no legal definition of inadequate achievement or pattern of strengths and weaknesses. Teams are obliged to document all areas of educational need 34 C.F.R. § 300.304(c)(1). Educational need may be conceptualized as any areas of achievement that require continued support to make progress towards grade-level standards. Minnesota Rules 3525.1341 also require that documentation must be representative of the child's curriculum and useful for developing instructional goals and objectives.

Sources of data teams may use in their analysis include, but are not limited, to:

- Repeated measures of achievement.
- Cumulative record review.
- Class work samples.
- Teacher records.
- State or district assessments.
- Formal and informal tests.

- Curriculum-based evaluation results.
- Results from targeted support programs.

Teams should use the sources of data available to construct a holistic picture of the student and how the student is performing relative to age and grade-level standards. Integrating multiple sources of achievement data provides a picture of how well a student is meeting grade-level standards. It also reveals which conditions that improve a student's skill acquisition and those conditions that constrain performance. Teams may find *Figure 9-1: Likely Patterns of Performance for SLD Identification* helpful in constructing a holistic picture of academic performance.

Note: Achievement is compared to age and grade-level standards, but information processing deficits are normative deficits.

The rest of the segments in this section cover sources of data that help develop a complete picture of the student performance and learning preferences.

Classroom Data and Professional Judgment

Classroom data and professional judgment are required by teams to determine the extent to which the instruction and environment have been changed to improve student learning. Information that should be used from observations, student work, record reviews, etc., includes:

- Potential positive influences to achievement (successful means of differentiation).
- Whether core instruction and interventions were nondiscriminatory and delivered with quality.
- Whether response to faithfully implemented interventions was sufficient.
- The extent to which additional supports or interventions are likely to improve achievement.
- Nuances in performance noted across teachers, classroom, non-classroom, tutorial environments.
- Rigor of instructional goals and objectives and performance across time.

Data used to make professional decisions may include, but is not limited to:

- Observations of the student in the regular classroom setting that document the student's academic performance and behavior in the area of difficulty.
- General education teacher's assessment of the student's daily work revealing a relative lack of quality and depth on a consistent basis (work samples may be representative of interventions, targeted support program such as Title 1, or daily work from regular curriculum).

- Pre and post measures indicating a lack of achievement over time (informal inventories, screening assessments, formative/summative assessments).
- Records and reporting systems showing a pattern of poor achievement.
- Family members' concerns that the student is not achieving to potential.
- Student reports indicating frustration with achievement, comprehension, following directions, completing assignments, building and maintaining friendships, etc.
- Classroom work that is below expectations in depth, breadth, or complexity for this student when compared to his or her peers.
- Classroom work that demonstrates a breakdown in a specific stage of learning: lack of skill acquisition, proficiency, maintenance, generalization or adaptation.
- Teacher records, e.g., results of conferences, anecdotal reflections of student learning.
- Behaviors or approach to tasks, or thinking strategies observed during assessment administration.

Examples of notable behaviors include, but are not limited to:

- Attitude and interests toward testing or any changes, before, during, after testing.
- Degree of comprehension and compliance with assessment directions.
- Response to visual, auditory or motor demands.
- Receptive and expressive language characteristics.
- Recognition of errors and attempts to change or solve a problem.
- Repetition of mistakes with or without level of self-awareness or monitoring of responses.
- Management of frustration.
- Verbalizations or thinking aloud before, during, after tasks.
- Task approach (impulsive, thoughtful, gives up easily, persists, revisions of answers, etc.).
- Response to success, failure, reinforcers (verbal and physical).

Observation Data

Data from observations made during instruction should be integrated into judgments about the quantitative results. Observation data can be gathered to provide context for standardized assessment data may include, but is not limited to:

- Curricular influences on achievement:
 - Whether core instruction and interventions were nondiscriminatory and delivered with fidelity and quality.
 - Rigor of instruction as compared with grade-level expectations and student performance across time.

- Instructional influences on achievement:
 - Whether interventions were implemented with fidelity.
 - Nuances in performance noted across teachers, classroom, non-classroom and tutorial environments.
 - Response to directions, success, failure, use of reinforcers (verbal and physical), etc.
 - Response to instruction with different size groups, delivery methods or materials.
 - Demonstration of a breakdown in stage of learning: lack of skill acquisition, proficiency, maintenance, generalization or adaptation.
 - Instructional adjustments, modifications, or additional supports within intervention that would likely strengthen response and rate of skill acquisition.

- Learner centered influences on achievement:
 - Frequency, duration, or latency of behaviors.
 - Approach to task and management of frustration (impulsive, thoughtful, gives up easily, persists, revisions of answers, etc.).
 - Verbalizations or thinking aloud before, during and after tasks.
 - Use of strategies, cues or problem solving to regulate attention, emotion, or behavior.

Informal Assessment Procedures

Informal measures and procedures provide assessment teams with the ability to test limits, determine instructional levels, verify mastery of competency or curriculum, identify factors that contribute to skills, and test assumptions given differences between performance on open-ended and close-ended tasks. Supplement standardized measures with informal and other assessment procedures, such as:

- Criterion-referenced tests that indicate whether a student has met a pre-determined standard of performance.

- Work samples collected under varying conditions that show the breadth of skills under different learning conditions and environmental contexts.

- Informal writing, math or reading inventories that consist of graded prompts indicating student's instructional, independent or frustration level.
 - Examples include:
 - Jerry John's Informal Reading Inventory.
 - Informal Phonics Survey.
 - Qualitative Reading Inventory.
 - Qualitative Spelling Inventory.
- Checklists and rubrics developed from research or qualitative analysis. Examples include:
 - Multidimensional Fluency Scale.
 - National Assessment of Educational Progress Integrated Reading Performance Record.
 - Teacher-made formative and summative assessments linked with curriculum and state standards.

Repeated Measures of Achievement (Progress Monitoring Data)

Repeated measures of achievement or progress monitoring data may be the strongest indicator of a student's degree of impairment or limits to participation in general education when provided with high-quality instruction. Use progress monitoring data whenever data are determined to be a valid and reliable measure of the student's achievement. Progress monitoring data should:

- Indicate baseline performance.
- Indicate changes or shifts in intervention/instructional strategies via marked graphs.
- Indicate that regular measurements were taken.
- Contain a minimum of 12 data points gathered over the course of intervention(s) consistently implemented over at least seven weeks.
- Reflect level of performance expected across time when given the full intensity of intervention.
- Reflect the trend or slope of the student's growth rate when given the full intensity and duration of intervention.
- Reflect the trend of correct—and when appropriate incorrect—responses.
- May also reflect loss and recoupment time over breaks in instruction.

Data should reflect that interventions were modified or changed according to pre-determined decision rules with any extenuating circumstances noted. Judgments of the data should include consideration of the intensity, frequency, length, duration and fidelity of intervention received by the student, number of and quality of probes used to gather data and consistency in scoring and interpreting data.

When data from progress monitoring measures do not meet requirements for technical adequacy, use other standardized measures to document inadequate achievement in the area of academic concern. Teams may choose to include in the comprehensive evaluation progress-monitoring data from independent tutoring or instruction provided outside the school day that meets the criteria stated above.

Standardized Measures of Achievement

Standardized, norm-referenced measures of achievement help teams determine how well a student is performing relative to a peer group. It is important to note that group-administered achievement tests, including Minnesota Basic Skills Tests and Statewide Testing, do not have the sensitivity and are not intended to be adequate either for specific eligibility criteria or for writing IEP goals and objectives.

Important: The following lists of assessments have been selected for the skills they measure and are not equal in their ability to address referral questions. The list is not intended to be exhaustive. Teams may choose other assessment tools or versions with updated norms as long as they are adequate measures of the abilities being tested (see each test manual for intended uses, strengths, limitations and interpretive guidance).

Professionals have an obligation to stay updated on appropriate assessment tools for the purposes for which they are being used. Findings from tests should be useful for developing the student's instructional programming.

Phonological Skills

- Comprehensive Test of Phonological Processing (CTOPP).
- SCAN-C Test for Auditory Processing Disorders in Children Revised (SCAN-C-R).
- Test of Auditory Comprehension of Language 3rd Ed (TALC-3).
- Test of Phonological Awareness 2nd Ed PLUS (TOPA-2+).
- Test of Phonological Skills (TOPAS) and Test of Phonological Awareness in Spanish (TPAS).

Reading

- Measures of Academic Progress (NWEA-MAP), for formative assessment only.
- Gray Diagnostic Reading Tests 2nd Edition (GDRT-2).

- Gray Oral Reading Test 4th ED (GORT-4).
- Early Reading Diagnostic Assessment (ERDA).
- Gray Silent Reading Test (GSRT).
- Slosson Oral Reading Test-Revised (SORT-R3).
- Standardized Reading Inventory-2nd ED (SRI-2).
- Test of Early Reading Ability 3rd ED (TERA-3).
- Test of Irregular Word Reading Efficiency.
- Test of Reading Comprehension (informs professional judgment and instruction, but unless norms are updated, do not use scores to establish a pattern of inadequate achievement).
- Test of Silent Word Reading Fluency (TOSWRF).
- Test of Silent Contextual Reading.
- Test of Word Reading Efficiency, Test of Silent Word Reading Fluency, Test of Irregular Word Reading Efficiency.
- Woodcock-Johnson III Diagnostic Reading Battery (WJ III DRB).
- Woodcock Reading Mastery Tests-Revised/Normative Update (WRMT-R/NU).
- Word Identification and Spelling Test (WIST).

To determine the appropriate uses of tests listed in this manual, read the administration manual and independent reviews in the latest editions of:

- Mental Measurements Yearbook.
- The Achievement Test Desk Reference: A Guide to Learning Disability Identification 2nd Ed (Flanagan, Ortiz, Alfonso & Mascolo, 2006). Boston: Allyn & Bacon.
- Assessment of Children's Cognitive Applications (Sattler 2008).
- Handbook of Psychological and Educational Assessment of Children, 2nd Ed: Intelligence, Aptitude, and Achievement (Reynolds, Kamphaus, & Hendry 2003).

How to determine if a test is appropriate for use in the discrepancy calculation:

1. The technical manual states that the assessment is valid and reliable for discriminating between individuals with SLD and other groups. Validity of .9 or .7-8 with corroborating evidence from other sources of data.
2. The normative sample is less than 10 years old and represents the student being assessed.
3. Cluster scores, composite scores, or scores derived from multiple subtests are available.
4. Standard scores can be calculated.
5. The test items are developmentally appropriate and are sufficient to represent the skills requiring assessment.

Do not use scores from a test with out-of-date norms for the calculation of discrepancy or as indicators of student performance. Teams may use performance, observed behaviors during administration, and error analysis to inform professional judgment, verify a hypothesis or design instruction.

Math

- Measures of Academic Progress (NWEA-MAP), for formative assessment only.
- Comprehensive Mathematical Abilities Test (CMAT).
- Key Math III.
- Test of Early Math Abilities (TEMA).

Written Language

- Oral and Written Language Scales: Written Expression (OWLS:WE).
- Test of Early Written Language 2nd Ed (TEWL-2).
- Test of Written Language 3rd Ed (TOWL-3).

Language Tests (to use score in the discrepancy calculation, the selected test must measure achievement)

- Clinical Evaluation of Language Fundamentals 4th Ed (CELF-4) (this is not a measure of achievement but has been useful in designing instructional programming).
- Comprehensive Assessment of Spoken Language (CASL) (informs professional judgment and instruction, but do not use scores unless norms are updated).
- Comprehensive Receptive and Expressive Vocabulary Test 2nd Ed (CREVT-2).
- Comprehensive Receptive and Expressive Vocabulary Test-Adult (CREVT-A).
- Expressive One-Word Picture Vocabulary Test (EO-WPVT).
- Illinois Test of Psycholinguistic Abilities 3rd Ed (ITPA-3).
- Oral and Written Language Scales (OWLS) (Inform professional judgment and instruction, but do not use in calculating discrepancy scores until norms are updated).
- Peabody Picture Vocabulary Test III (PPVT-III).
- Receptive One-Word Picture Vocabulary Test (RO-WPVT).
- Test of Adolescent and Adult Language.
- Test of Early Language Development 3rd Ed (TELD-3).
- Test of Language Development-Intermediate: 3rd Ed (TOLD-I:3).
- Test of Language Development-Primary: 3rd Ed (TOLD-P:3).
- The WORD Test 2nd Ed (WORD-2).

- Test of Expressive Language.
- Woodcock Language Proficiency Battery—Revised.

Comprehensive Achievement Batteries

Comprehensive achievement batteries provide the broadest picture of a student's achievement. Data gathered from other sources, such as interventions in one academic area, do not provide a comprehensive understanding of the student's academic needs. Although teams may find that items on the assessments do not adequately correlate to state standards, it is important to understand that test items have been selected for their ability to differentiate learners. Comprehensive assessment batteries are appropriate measures to use when identifying comprehensive patterns of achievement. Users need to be sure the subtests in the tests they select to use:

- Are developmentally appropriate.
- Have adequate specificity and sensitivity to identify areas of strength and weakness in students of similar age.
- Closely align with curricular expectations.
- Measures used in calculating discrepancy must provide standard scores (mean standard score of 100, standard deviation of ± 15).

Important: The following lists of assessments are not equal in their ability to address referral questions. It is the obligation of the professionals selecting and administering the tests to use the most appropriate test for each student and referral concern. The list is not intended to be exhaustive. Teams may choose other assessment tools or versions with updated norms as long as they are adequate measures of the abilities being tested (see each test manual for intended uses, strengths, limitations and interpretive guidance). Professionals have an obligation to be trained and knowledgeable about the tests they are administering.

Batteries to document inadequate achievement and calculation of the discrepancy score:

- Diagnostic Achievement Battery (DAB-3).
- Kaufman Test of Educational Achievement 2nd Ed (KTEA-II).
- Peabody Individual Achievement Test Revised/Normative Update (PIAE-R/NU).
- Wechsler Individual Achievement Test 2nd Ed (WIAT—II).
- Woodcock-Johnson III Test of Achievement (WJIII)/NU.

Test Selection for Eligibility Decision

Use the following suggestions when selecting technically adequate assessment tools for eligibility decisions:

- Use tests with age-based norms that are no more than 10 years old.
- Use tests designed specifically for, or considered an appropriate and robust measure of, one of the eight areas of academic functioning specific to Minnesota State Rule. See rule language at beginning of chapter 10.
- Use tests with adequate norming sample. The norming should have been conducted using a sample of people from the United States with adequate samples of students at the age of the student being tested.
- Use tests selected and administered in a manner as to not be discriminatory on a racial or cultural basis.
- Ensure that the test's technical manual states that the assessment is valid and reliable for discriminating between individuals with SLD and other groups, including validity of .9 (or .7-.8 with corroborating evidence from other sources of data).
- Use tests that create cluster scores or a score derived from multiple sub-tests
- Avoid deviations from the standard administration of any standardized test that invalidate the score for eligibility and placement decisions. Non-standard administration includes, for example:
 - Not using a tape recorder for a subtest when required by the standard administration directions in the testing technical manual.
 - Testing in a classroom full of students.
 - Extending the allotted time for a subtest.
 - Completing the math calculation section with a calculator.
- Testing of limits may occur after ceilings are reached, and may provide valuable information for the design of instruction and to reveal a student's thinking strategies or processes.
- Administer a standardized test according to procedures outlined in the administrative manual. Do not administer testing sessions subtest by subtest, occurring on different days. This will invalidate the score.
- Administer assessments that ensure students with impaired sensory, manual or speaking skills accurately reflect aptitude or achievement level rather than the impairments (unless those skills are the factors the tests purport to measure).

Refer to Selection of Assessments in Chapter 11 for more on ethical principles that guide conduct of assessment practices.

The following list of standardized measures of achievement may be helpful to teams in the selection of tests useful for filling in gaps and determining underlying skill deficits represented within the student's curriculum.



Quality Practices in Collecting Information Processing Data

Minnesota Rule 3525.1341 requires documentation that a disorder occurs in multiple settings impacting one or more of the basic psychological processes. Teams are required to use multiple sources of data to illustrate that the disorder in basic psychological processes is manifested in the imperfect ability to listen, think, speak, read, write, spell or to do mathematical calculations. For this reason, it is recommended that once a disability is suspected, teams should create a hypothesis of basic psychological processes that are likely impacted. Teams should use a combination of the informal data gathered from the tools listed above plus student work samples collected during intervention and standardized measures to validate and document the disorder and its impact on achievement.

Quality practices indicate that a normative deficit in information processing ability is necessary for being identified as having an SLD. For efficiency, teams will be served by using observation data, work samples, and interview results to direct the selection of instruments and methods. Teams should validate the suspected disorder in basic psychological processes by using standardized measures to determine the normative weakness and strengths (both relative and normative). Use of norm-referenced tools, such as standardized assessments, and observations, rating scales, etc., allows teams to establish a threshold of functioning that is more reliable than informal checklists and interviews.

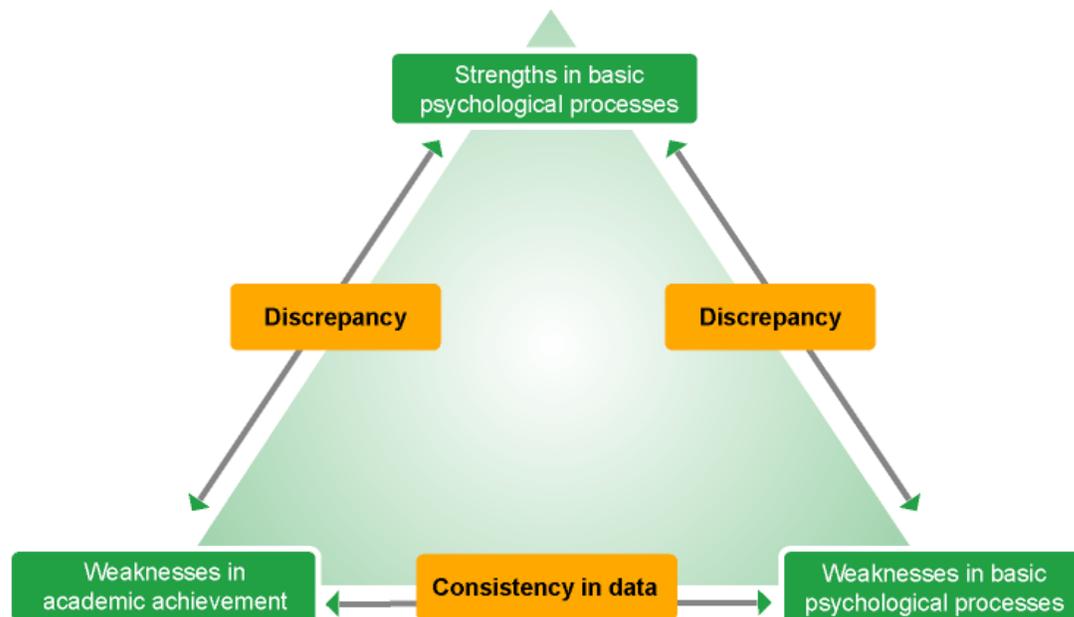


Figure 8-3. Relationship between Basic Psychological Processes and Achievement

Adapted with minor changes in terminology to be consistent with language in Minnesota Rule from Hale, J. B., Flanagan, D. P., & Naglieri, J. A. (2008). *Alternative Research-Based Methods for Reauthorized Federal IDEA 2004 Identification of Children with Specific Learning Disabilities. Communique.*

Teams should look for an empirical relationship between poor achievement and normative weaknesses in basic psychological processes. To further differentiate students with SLD from those with general learning difficulties, teams should expect to find normal functioning in those abilities/processes not strongly related to academic deficits. Include assessment of attention, memory and executive functions for older elementary students because they are critical elements in middle and high school success. Limitations in executive functions and memory have increasing impact on academic achievement throughout middle and high school.

Examples of Standardized Measures of Basic Psychological Processes Used by Psychologists or Specially Trained Personnel

Important: Teams may not be familiar with all the assessments in the following lists. Professionals have an obligation to familiarize themselves with tests that they may not regularly use. Testing manuals, peer-reviewed articles or independent reviews conducted by Burros Mental Measurements Yearbook may be helpful in determining the intended uses, strengths, limitations, and interpretive guidance for otherwise unfamiliar tests.

The following lists of assessments are not equal in their ability to address evaluation questions. It is the obligation of the professionals selecting and administering the tests to use the most appropriate test for each student and referral concerned. The list is not intended to be exhaustive. Teams may choose other assessment tools or versions with updated norms as long as they are adequate measures of the abilities being tested (see each test manual for intended uses, strengths, limitations and interpretive guidance). Professionals have an obligation to be trained in and knowledgeable about the tests they are administering.

Multiple Abilities Measures

- Cognitive Assessment System for Children (CAS).
- California Verbal Learning Test-II Children's Version.
- Differential Ability Scales (DAS II).
- Kaufman Assessment Battery for Children 2nd Ed (KABC-II).
- NEPSY: A Developmental Neuropsychological Assessment.
- Process Assessment of the Learner II (PAL-II): Test Battery for Reading and Writing.
- Wechsler Intelligence Scale for Children - 4th Edition Integrated (WISC-IV Integrated).
- Woodcock Johnson Test of Cognitive Abilities (WJ III Cog).

Processing Speed

- Rapid Automatized Naming/Rapid Alternating Stimulus Test (RAN/RAS).

Executive Functions

- Behavior Rating Inventory of Executive Functions (BRIEF).
- Delis Kaplan Executive Function System.

Phonological Processing

- Comprehensive Test of Phonological Processing (CTOPP).
- Lindamood Auditory Conceptualization Test.
- SCAN-C Test for Auditory Processing Disorders in Children Revised (SCAN-C-R).
- Test of Auditory Processing Skills.

Visual Processing

- Benton Visual Retention Test Revised (BVRT).
- Developmental Test of Visual Perception for Adolescents and Adults.
- Test of Visual Perception Skills.

Orthographic Processing

- Test of Irregular Word Reading Efficiency (test for orthographic processing), Nancy Mather.
- Kauffman Test of Educational Achievement (KTEA-II) Word Recognition Fluency Test.
- Test of Silent Word Reading Fluency.

Working Memory

- Wechsler Memory Scale for Children 3rd Ed (WMS-III).
- Working Memory Test Battery for Children,
- Woodcock-Johnson Interpretation and Instructional Interventions Program (WIIP) .

Oral Motor Production Processing

- PAL-II Rapid Naming Tasks.
- Delis-Kaplan Word Reading Task.
- CTOPP Rapid Naming Task.
- NEPSY II-Speeded Naming.
- KTEA-II Naming Facility.

NOTE: The following tests may be used to confirm professional judgments; however, they are not technically adequate for documenting a deficit in basic psychological processes.

- Wisconsin Card Sorting Test, Revised and Expanded.

- Ray Osterreith Complex Figure Drawing - How an individual organizes information, executive processing.
- Conners Continuous Performance Test 3rd Edition.
- Ross Information Processing Assessment.

Note: At this time none of the following tools are technically adequate to determine whether areas of psychological processing are below average relative to same age peers. Some trained practitioners will find them helpful in making professional judgments or interpreting other standardized test results.

- Psychological Processing Checklist (PPC). This tool was designed as a screener for developing interventions but should not be used as a sole source of data for the eligibility determination.
- Learning Disabilities Diagnostic Inventory (LDDI). Independent reviews suggest that when used as a screener this tool may miss up to 43 percent of students that are truly SLD.
- Neuropsychological Observation Checklist (Hale & Fiorello, 2008).

Empirical Relationship between Achievement and Information Processing Ability

This table presents each area of inadequate achievement empirically linked with areas of likely information processing deficits. Where appropriate the Cattell - Horn - Carroll (CHC) abilities (for an explanation of CHC abilities [see the *Cattell - Horn - Carroll \(CHC\) Cognitive Abilities-Achievement Meta-Analysis* project at <http://intelligencetesting.blogspot.com>\) have been included so that assessment teams may more easily select the appropriate domains to assess.](http://intelligencetesting.blogspot.com)

Table 8-3

Domains of Achievement and Their Related Information Processing Abilities.

Skill	Information Processing Ability / Stage of Development
Oral Language	<ul style="list-style-type: none"> • Working memory. • Processing speed.
Listening Comprehension	<ul style="list-style-type: none"> • Auditory working memory. • Processing speed. • Auditory short-term memory.
Basic Reading Skills	<ul style="list-style-type: none"> • Phonetic coding (Ga) phonological awareness - very important in elementary years. • Naming facility and associative memory (Glr) - very important during elementary years. • Memory span (Gsm) - important, especially when evaluated within the context of working memory. • Perceptual speed (Gs) - important across all ages, particularly in elementary school. • Orthographic processing (Gv) - important especially in early elementary years. Indicated by poor visual tracking and/or motion sensitivity. • Successive processing—(Dehn, 2006). • Verbal working memory - best predictor of ability to identify letters for young students (Webster, Plante, & Couvillion, 1997).
Reading Fluency	<ul style="list-style-type: none"> • Naming facility and associative memory (Glr) - very important during the elementary years. • Phonetic coding (Ga) phonological awareness - important during elementary years. • Perceptual speed (Gs) is important across all ages, particularly in elementary school.

Skill	Information Processing Ability / Stage of Development
Reading Comprehension	<ul style="list-style-type: none"> • Language development, lexical knowledge, and listening ability (Gc) become increasingly important with age. • Inductive and general sequential reasoning (Gf) play a moderate role in reading comprehension. • Morphological awareness is showing some influence in the late identified reading disabilities. • Working memory. • Self-monitoring.
Written Expression	<ul style="list-style-type: none"> • Inductive and general sequential reasoning (Gf) impacts the fluency aspect of writing as well as more general writing ability across all ages. • Phonetic coding (Ga) or phonological awareness is very important during elementary years (primarily before age 11) for both basic writing skills and written expression. Automaticity in spelling lays the foundation for higher level writing skills. • Naming facility (Glr) or rapid automatic naming has demonstrated relations with written expression, primarily the fluency aspect of writing. • Memory span (Gsm) is important to writing, especially spelling skills whereas working memory has shown relations with advanced writing skills (e.g., written expression). • Perceptual speed (Gs) is important across all ages, particularly in elementary school. • Orthographic processing and morphological awareness [there is increasing support for these two abilities and their impact on spelling and basic writing abilities]. • Executive processing and planning (Dehn, 2006). • There is limited evidence that lexical knowledge, language development and general information also contribute to written expression; however, more research needs to be conducted.
Math Calculations	<ul style="list-style-type: none"> • Inductive and general sequential reasoning (Gf) are consistently very important at all ages. • Memory span (Gsm) is important especially when evaluated within the context of working memory. • Perceptual speed (Gs) is important across all ages, particularly in elementary school.

Skill	Information Processing Ability / Stage of Development
Math Problem Solving	<ul style="list-style-type: none"> • Inductive and general sequential reasoning (Gf) are consistently very important at all ages. • Language development, lexical knowledge, and listening ability (Gc) become increasingly important with age . • Perceptual speed (Gs) is important across all ages, particularly in elementary school. • Memory span (Gsm) is important, especially when evaluated within the context of working memory. • Visual Processing (Gv) may be important primarily for higher level or advanced mathematics (geometry, calculus). • Long-term memory capacity (Glr) may be important in predicting mathematical problem-solving accuracy beyond that predicted by memory span and processing speed. More research studies are needed in this area.

--Adapted from Flanagan, Ortiz, Alfonso, & Mascolo (2007). For an explanation of CHC abilities [see the Cattell - Horn - Carroll \(CHC\) Cognitive Abilities-Achievement Meta-Analysis project at http://intelligencetesting.blogspot.com](http://intelligencetesting.blogspot.com). The CHC codes are in parentheses following the names for each ability.

Collecting Data on Cognitive or Intellectual Functioning

Historically, the rationale for assessing cognitive or intellectual functioning was to determine that a lack of achievement was unexpected. Teams using a system of SRBI to document eligibility must ask whether it is appropriate to fully assess cognitive or intellectual functioning. Each comprehensive evaluation must consider the student as a single case. Some have suggested that measures of adaptive behavior used in determining Developmental Cognitive Disability (DCD) provide adequate data to rule out questions of inadequate intellectual ability. **However, this practice is not recommended.** Data from adaptive measures may overestimate a student's abilities, as they are better measures of social maturity and independence than cognitive potential.

Brief intellectual assessments provide convergent data for average or above intellectual performance and may be helpful in circumstances where bias or uncertainty regarding student abilities exists. Therefore, weigh the cost/benefit of conducting a brief or full cognitive/intellectual assessment. One consideration is whether brief measures adequately gauge IQ over more comprehensive measures. Brief IQ measures do not provide an adequate measure of information processing.

The *Cattell-Horn-Carroll Theory of Intelligence (CHC)*, currently the most comprehensive and empirically supported theory of the structure of cognitive/intellectual functioning and academic abilities, is appropriate for discrepancy calculation. Derived from synthesizing hundreds of factor analytic studies, CHC is the guiding framework from which all intellectual assessments have been based since 2000.

The structure is composed of broad and narrow abilities. Broad abilities are "basic constitutional and long standing characteristics of individuals that can govern or influence a great variety of behaviors in a given domain" (Carroll, 1993 p. 634). Narrow abilities

“represent greater specialization of abilities, often in quite specific ways that reflect the effects of experience and learning, or the adoption of particular strategies of performance” (Carroll, 1993 p. 634). The *Planning, Attention, Simultaneous, and Successive (PASS) Theory* also provides a reasonable alternative.

Cross-Battery assessment has guided the organization, selection, and interpretation of assessments across intelligence batteries. Use Cross-Battery to generate scores of intellectual functioning and information processing.

Important: The following lists of assessments are not exhaustive or meant to convey an approval for identification. Teams may not be familiar with all the assessments in the following lists. Professionals have an obligation to familiarize themselves with tests that they may not regularly use. Testing manuals, peer-reviewed articles or independent reviews conducted by Burros Mental Measurements Yearbook may be helpful in determining the intended uses, strengths, limitations, and interpretive guidance for otherwise unfamiliar tests.

Examples of Standardized Measures of Intellectual Abilities used by Psychologists or Specially Trained Personnel

- Cognitive Assessment System (CAS).
- Comprehensive Test of Nonverbal Intelligence (CTONI).
- Differential Abilities Scales-II (DAS-II) (Appropriate for students with cultural or linguistic differences and young children. Research indicates this test has the smallest differences across ethnic groups.)
- Kaufman Adolescent and Adult Intelligence Test (KAAIT).
- Kaufman Assessment Battery for Children 2nd Edition (KABC-2).
- Reynolds Intellectual Assessment Scales (RIAS) (Research indicates scores may be inflated compared with more traditional measures, see Edwards & Paulin, 2007).
- Stanford-Binet V (SB-V).
- Wechsler Adult Intelligence Scale (WAIS-III).
- Wechsler Intelligence Scale for Children 4th Edition (WISC-IV).
- Wechsler Pre-school and Primary Scale of Intelligence 3rd Edition (WPPSI-3).
- Woodcock Johnson III Test of Cognitive Abilities (WJ-III Cog.) (Most comprehensive measure of CHC theory.)



Quality Practices in Designing Comprehensive Evaluations for Young Children

This section focuses on young students approximately kindergarten age and older advancing on to elementary settings. Students participating in systematically implemented research-based interventions are typically identified in the latter part of the kindergarten year.

A comprehensive assessment is essential for the identification of SLD in young students and calls for the careful consideration and selection of assessment instruments. The assessment of young students is complicated by factors related to their personalities and development. While significant research exists on identifying children at high risk for SLD at ages 3 and 4, literature for assessing and diagnosing preschool students is fraught with caveats.

Consider the following during selection of tools and sources of data in order to meet SLD criteria:

- Validity and reliability of existing data.
- Data necessary for:
 - Ruling out alternative explanations for the inadequate achievement.
 - Understanding the underlying causes of inadequate achievement.
 - Meeting remaining eligibility criteria and designing instruction.
- Tests designed to collect standardized assessment data. Use screening batteries or narrow measures of achievement when comprehensive measures are lacking; however, do not use them to calculate the discrepancy. Teams may decide to continue to administer these types of assessments because they do provide relevant instructional information.
- Norming group information, available scores, reliability and validity of data, and the use and interpretation of the scores for the student's age level (Salvia & Ysseldyke, 1988).

Include team members who are knowledgeable about Early Childhood Special Education (ECSE) eligibility assessment and SLD. The ECSE evaluation may include assessment of developmental levels in the following skill areas: cognitive (pre-academic), social, motor, adaptive and communication. Use this information to identify areas of weakness or underlying causes of learning difficulties and in forming a hypothesis that guides the assessment process.

Cautions in the Assessment of Young Students

The following are examples of potential problems in the assessment of young students.

- A wealth of systematic research-based interventions exists in K-3, yet the research base prior to kindergarten continues to emerge. Peer norms for progress monitoring of interventions delivered in pre-school settings are not yet established. Without adequate documentation of systemic implementation of SRBI practices, determination of eligibility from intervention data collected prior to kindergarten is likely not possible.
- Be cautious when interpreting scores resulting from standardized tests of intellectual ability and academic achievement for students under age 5. Careful analysis of test norming information is critical, and consulting technical manuals is imperative.
- Include consideration of maturation and development through observing the student's behavior in typical settings, such as home, school and community. Staff must be knowledgeable and experienced not only in early childhood development but also in the use of anecdotal records, behavior rating scales, and functional assessment.
- Pay attention to the student's developmental history, including appropriate medical information such as birth trauma, low birth weight, lack of oxygen, etc. In addition, evaluate present performance levels in speech and language development, motor skills, social competence, conceptual development and abstract reasoning abilities.
- Various manifestations of learning disabilities may be seen in the same student at different ages and as a result of different learning demands. Learning disabilities are often first manifested as specific deficits in language and speech development, and other behaviors for some pre-school students. Marked discrepancies in abilities may be temporary and are resolved during the course of development or within the application of the intervention. For other young students, marked discrepancies persist within and among domains or show continued poor response to well-designed interventions, necessitating the student's referral for special education assessment.
- If a young student receives a raw score of zero on a valid and reliable standardized achievement test, there is typically a corresponding standard score. If the team questions the validity of the derived standard score, further assessment may be necessary using a supplemental test. Results are reported as standard scores, have a standard deviation of ± 15 , and used are to compute a severe discrepancy.
- When making discrepancy calculations, use the Full Scale IQ score or *General Ability Index* score. The following caution from Sattler (1988) remains relevant to this day and requires professionals' consideration: "Generally, whereas IQs obtained prior to 5 years of age must be interpreted cautiously, IQs tend to remain relatively stable from kindergarten on ... The IQ of any given student may change as much as 20 points, but for most children measured intelligence remains relatively stable after 5 years of age ... In spite of high test-retest correlation in assessing individuals it is necessary to conduct frequent and periodic testing if test scores are to be used for guidance or placement decisions".

Setting the Stage for Assessing Young Children

During the assessment process, create a hypothesis about the area of achievement and information processing weakness. Use only valid and reliable data to make the eligibility decision. Be aware of effects on the validity of test scores and testing procedures used, since young students:

- Are sensitive to their surroundings and therefore may be easily distracted.
- Are influenced by their comfort level with the assessor.
- Should be assessed in a variety of situations.
- May have rapid developmental change.
- May have a limited interest in being assessed.
- Experience more rapid neuro-biological changes than older students.
- Have limited communication skills that may interfere with their understanding and/or responses.
- May be distractible and have a short attention span that could affect their responses.
- May have separation issues with parents making assessment difficult.
- May be noncompliant or have poor understanding of social relationships that may affect performance (McLean, Bailey, and Wolery, 1996).

Transitioning from Developmental Delay to SLD

Students who are aging out of Developmental Delay (DD) to a specific disability category under Minnesota Rule 3525.1350, subp. 3 Part B must have a re-evaluation to determine eligibility for a categorical disability prior to the child's seventh birthday.

Important: When a student already receives special education services but is identified under a new disability category, the team will conduct a re-evaluation in terms of due process, but the student must meet initial eligibility criteria for the new disability category. This means that a student with a developmental delay must meet initial SLD eligibility criteria set forth under Minnesota Rule 3525.1341 in any one or more of the eight areas. A student is not required to meet eligibility criteria in all areas of academic need in order to receive specially-designed instruction. Teams must ensure the evaluation is sufficiently comprehensive to identify all of the child's or student's special education and related services needs, whether or not commonly linked to the disability category in which the child has been classified according to 34 CFR 300.304 (c)(6).

If students are transitioning from Developmental Delay to SLD eligibility and the school is exercising its choice to use criteria ABD for the re-evaluation, then data gathered on response to interventions provided in special or regular education program may be used. Students receiving services for DD should already have their progress monitored as part of the IEP in each area of academic, social/emotional, or behavioral concern. This data can be

used as part of documentation of eligibility for criteria ABD.

The Early Childhood transition process should ensure that:

- Team members have adequate information about the young student (if possible, by the beginning of the school year), to feel confident providing services in an elementary setting. Obtain data such as progress monitoring from pre-referral interventions if the student demonstrates inadequate achievement in a skill possibly related to a developmental delay, but which is not served by special education services. For example, provide reading intervention services to students experiencing delays in language development.
- Teams involved in the transition process may include members from ECSE, SLD, Early Childhood and Family Education (ECFE), community preschool staff, kindergarten staff. Parents must participate.
- Calendar timelines allow for appropriate planning, assessment, and IEP development as long as the student has transitioned from Part C to Part B by their seventh birthday.



Quality Practices in Designing Comprehensive Evaluations for Learners with Cultural and Language Differences

Few tasks are more difficult for school psychologists than evaluating the cognitive abilities and intellectual functioning of individuals who are culturally and linguistically diverse. The inadequate training many have to assess the abilities of such individuals can be one reason for the disproportionate representation of minority groups in special education and the possible misidentification of some of these students as having a disability. Likewise, inappropriate evaluation can also lead to under-representation, so that some individuals who have disabilities and are in need of services are not identified.

Compared with assessments of English-speaking students raised in mainstream U.S. culture, the process of assessment of students with language or cultural differences is anything but straightforward. Among other things, it is hampered by the lack of appropriate tools. The requirement in IDEA 2004 represents an intent to draw attention to the goal of non-discriminatory assessment. Valid and reliable assessment tools, methods, procedures, and processes employed in the evaluation of diverse students carries with it some degree of bias. Additionally, individuals administering and interpreting the results may carry hidden biases. Although the intentions of an evaluation are reasonably clear, nondiscriminatory assessment requires special procedures and care.

Goals of Nondiscriminatory Assessment

The framework from Ortiz (2002) makes it clear that nondiscriminatory assessment is more than selecting the “right” test or providing native language evaluation. The emphasis is placed on working in a systematic manner because reducing bias is accomplished only when actions are taken in an appropriate way and in an appropriate sequence. When attempts to reduce the discriminatory aspects of evaluation are marred by modifications or changes in the normal evaluative process, the results cannot be readily interpreted. Although the focus of this section is on intellectual assessment, particularly the use of standardized tests, in the course of such evaluations, remember that testing forms only one part of the overall framework for conducting nondiscriminatory assessment.

Nondiscriminatory assessment is viewed in the larger sense as a process designed to reduce disproportionate representation, the actual goal has more to do with differentiating cultural and linguistic difference from a disability under IDEA. It is important to understand that the focus of nondiscriminatory assessment rests on the issue of fairness and equity and should not be seen as methods that are simply intended to promote more racial balance in special education. In this sense, true nondiscriminatory assessment may be used for all students, not just those who are culturally and linguistically diverse. Professionals engage in these practices because they result in better evaluations and consequently better decisions about educational programming, not because they meet legal requirements or change the ethnic composition of students in special education.

Providing the type of evaluation that is necessary and required is too often seen as the search for the “right” tool or the “best” method. Because of the obvious nature of communication, most of the attention given to attempts at reducing bias in assessment is related to language. A great deal of concern is paid to methods that will provide an evaluation that is conducted in the student’s native language. This notion is perhaps reinforced by another specification in IDEA 2004 that requires agencies to “provide and administer assessments in the student’s native language, including ensuring that the form in which the test is provided or administered is most likely to yield accurate information on what the student knows and can do academically, developmentally, and functionally, unless it is clearly not feasible to provide or administer the assessment in this manner.”

This mandate actually expands the old provision for nondiscriminatory assessment but the wording regarding “native language” often misdirects evaluation efforts toward native language assessment as the primary strategy for providing a fair evaluation. Language usually is the presenting problem but the cultural aspects of evaluation must be paid at least equal attention. In fact, it has been suggested that cultural issues, not linguistic ones, represent the most important factors in being able to conduct fair assessments and that evaluation in the student’s native language often does little to reduce actual bias (Flanagan & Ortiz, 2001; Rhodes, Ochoa, & Ortiz, 2005).

Framework for Intellectual Assessment with ELL Learners

This section provides a framework to plan and carry out a nondiscriminatory evaluation of intellectual ability for ELL students.

Important: Examples of screening measures have been provided in the following tables as illustrative examples for districts. Although many of the following measures have been reviewed by the National Center for Student Progress Monitoring, **examples are not endorsed by the Minnesota Department of Education and are subject to change.**

Framework Summary

1. Review existing information on the student’s language background, language proficiency, culture, and educational history. Use tools and questions found in the *Reducing Bias in Special Education Assessments* manual or in Chapter 7: Suspicion of Disability of this manual.
2. Based on information on language proficiency and prior education, plot results on the Multidimensional Assessment Model for Bilingual Individuals (MAMBI). Identify the modes of intellectual assessment that are most likely to yield fair estimates of ability.
3. Use the *Culture-Language Test Classification matrix (C-LTC)* to select the most appropriate instruments (or subtests if using a Cross-Battery approach).

4. For instruments administered in such a way that standardization is valid, use the *Culture-Language Interpretive Matrix (C-LIM)* to plot and interpret results. See Chapter 9: Interpretation of Data.
5. Use at least one additional procedure (i.e., an optional mode of assessment recommended on the Multidimensional Assessment Model for Bilingual Individuals (MAMBI)) and/or testing-of-limits procedures.

Framework Details

1. Review and collect background information.

Formal assessment of intellectual ability is not the first step in the evaluation process. Teams should engage in a series of data-gathering efforts before using standardized tests. The information to be sought prior to the evaluation of cognitive abilities is crucial in setting the context for interpreting results fairly.

Of the various types of information collected, the most important are those which relate to the student’s level of acculturation and English language proficiency (conversational and advanced language capabilities as compared to native speakers). Teams often overestimate the level of acculturation or English language proficiency of students. Background information gathered should be used to determine how “different” the student is from the mainstream because the degree of difference impacts the expectations for performance on tests, such as “slightly different,” “different,” or “markedly different.”

2. Select assessment mode using MAMBI.

Psychologists may use the *Multidimensional Assessment Model for Bilingual Individuals* (MAMBI; Ochoa & Ortiz, 2005) to select appropriate assessment methods and materials. MAMBI is designed to provide guidance on the “most appropriate” modality of assessment, and the use of native language, English-only, nonverbal, or bilingual tests and methods. “Most appropriate” is to the method that is likely to yield the most fair and non-discriminatory estimates of actual ability assuming that standardization is maintained in the administration of the test.

MAMBI assists in balancing and integrating decision factors when using tests. It brings together the important variables in decisions such as students’:

- Current level of language proficiency both in English and the native language.
- Current grade placement.
- Current or previous educational program.

The integration of these factors using the MAMBI make it easier to determine the best assessment and least discriminatory modality for assessment as well as what other modes might provide valuable information.

Use of the MAMBI requires the ability to place the student into one of three categories for each language: minimal (Cognitive/Academic Language Proficiency levels 1-2), emergent (CALP level 3), and fluent (CALP levels 4-5). It then generates a “language profile.” These levels correspond to the student’s ease in performing classroom tasks as follows:

Minimal	CALP Levels 1-2	Classroom tasks are impossible or extremely difficult.
---------	-----------------	--

Emergent	CALP Level 3	Classroom tasks can be done with support.
Fluent	CALP Levels 4-5	Classroom tasks can be done with decreasing support and at a level of mastery that is similar to native speakers.

Case Example:

Juan D. tests “minimal” in the native language (L2) and “emergent” in English (L1), which gives him a Language Profile 4 (L1 minimal/L2 emergent). The preliminary stages of assessment reveal that this fourth grader received formal education in English (with or without ESL support). Interpretation of the section of the MAMBI that corresponds to his language profile and educational information indicate that the nonverbal assessment is the best modality most likely to yield the fairest estimate of his ability. This makes sense primarily because his language development is slightly better in English than Spanish, but both are limited in development. Using only verbal tests is unfair in either language. However, because of Juan’s better development in English, testing in L2 (English) may be valuable, but results would be more biased than those obtained from a nonverbal approach.

See Appendix for an example of the MAMBI. A wide variety of tests are available to measure a broad range of abilities in English (or L2). Native language tests may not yet be available when native language testing (L1) is recommended. When this happens, use a translator or interpreter. Remember to consider the BVAT, which is currently available in 16 languages, and various other tests in Spanish (Bateria III; WISC-IV Spanish), for testing L1. See Additional Procedures for more information.

Use a language-reduced test and administration format to accomplish nonverbal assessment using pantomime (such as with the UNIT) and other similarly language-reduced instruments, such as the Nonverbal Index from the KABC-II, the Leiter-R, C-TONI, etc.

Important: Use of nonverbal tools and methods does not automatically render the results valid. Less biased interpretation of the results from any test, irrespective of the modality, requires use of other procedures such as the C-LIM described later in this section.

Bilingual Assessment vs. Assessment of Bilinguals

A true “bilingual assessment” is carried out by a bilingual professional with access to valid assessment tools in all of the languages spoken by the student and administers these tools in a bilingual (or multilingual) manner. However, this rarely occurs due to a lack of appropriate tools and bilingual practitioners. Even when it is accomplished, no guidelines and standards for what constitutes best practices in true “bilingual” evaluation exist. Often the term bilingual is used when the evaluation is in fact monolingual in nature. Assessment of students in their native language only is not “bilingual.” When using a native language instrument, maintaining standardization is only necessary if the student’s background matches the norming sample and the assessor meets the professional and linguistic requirements. If these conditions are not met, testing of limits procedures should be liberally employed in order to evaluate and estimate the individual’s abilities in the fairest manner possible.

Most assessments are conducted in English using English-language tools, known as “assessment of bilinguals.” Adhere to the standardized instructions and administration guidelines because C-LIM can only be used to analyze and interpret the results when standardization is maintained.

3. Select Instruments using the Culture-Language Test Classification (C-LTC)

The Culture-Language Test Classification (C-LTC) and Culture-Language Interpretive Matrix (C-LIM) seamlessly integrate with MAMBI and provide an additional means of reducing bias in the assessment of intellectual ability.

Use C-LTC after choosing the assessment modality using the MAMBI to “hand pick” the tests that measure the constructs of interest with the least amount of cultural loading or linguistic demand and bias leading to fairest evaluation of the student’s abilities. Because it is impossible to assess all cognitive abilities with tests that are low in culture and low in language loading, use C-LIM to analyze test results and reduce bias in interpretation.

The C-LTC categorizes subtests of commonly used instruments along two dimensions: degree of language skill demanded by items, and degree of cultural knowledge required for successful task completion. Subtests are rated as low, medium, or high on both dimensions. Quick examination of the C-LTC shows a range of linguistic and cultural demand among both verbal and nonverbal subtests.

MAMBI helps evaluators select the assessment modality. C-LTC helps to select the fairest tests within that modality. Culture-Language Interpretive Matrix (C-LIM) helps interpret the results obtained from that modality. C-LIM was designed primarily for tests administered in English (including non-verbal administrations) integrated with any test or battery, and is not dependent on the CHC Cross-Battery assessment or the MAMBI. However, maintain standard protocol when using C-LIM.

Research is emerging on the use of these tools developed in other languages.

Native Language Assessment and Use of Interpreters

Valid psycho-educational instruments are available in only a few languages other than Spanish and are either translated or redeveloped and normed in the U.S. or elsewhere with monolingual or bilingual populations.

For students who are English Language Learners, a team may administer English-language tests with the help of an interpreter who:

- Interprets directions for the student.
- Interprets practice items or provides additional practice items.
- Interprets actual test items.
- Records responses given in native language to items that are posed in English.

Administration of an English-language test in native language through an interpreter is neither “bilingual assessment” nor “assessment of bilinguals” (see page 40), and is more structurally similar to assessment of bilinguals because the norms of an English-language test are based on English speakers. Comparisons of performance are made relative to this population, not a native-speaking one. Because the student is accorded native language instructions, this advantage makes it difficult to compare performance against similar individuals who were not provided the same benefit.

When an interpreter is required, administer the evaluation in English first and then in the native language. Follow standard protocol closely for the English test. Flexibility in administration is permitted for the native language administration and testing of limits. The interpreter’s primary role is to translate instructions and responses for both verbal and nonverbal tests. However, because the standard protocol is violated during ongoing translation, interpreters may help with meaning or purpose of a task to ensure best performance.

Conducting assessments in this manner allows student performance on the first administration to be analyzed for cultural and linguistic influences. When followed by administration of the same test in the native language, a comparison can be made between performance on the former and the latter. Individuals with learning difficulties are unlikely to appreciably change their performance so that any observed “practice effects” can be attributed to either better comprehension of the language (due to the change in administration) or intact ability that benefited from the prior practice. In either case, it provides valuable diagnostic information relative to whether the individual has a disability—the central question to any evaluation.

Choosing Instruments

KABC-II likely is a good instrument for assessment since it contains a relatively wide-range of abilities represented on the battery as a whole and provides composite scores that follow the C-LIM principles. It also provides fairer estimates of performance as a function of the student’s “difference.”

Use the following subtests of the KABCII under these conditions:

- **Fluid Crystallized Index (FCI)** - Student is slightly different in terms of language and culture. FCI is based on all age-appropriate subtests in the battery.
- **Mental Processing Composite (MPC)** - Student is moderately different, MPC eliminates the most highly culturally loaded and linguistically demanding subtests from the results (no Gc).
- **Nonverbal Index (NVI)** - Student is markedly different. The best estimate of performance, further reducing the inherent cultural loadings and linguistic demands of the component tests.

KABC-II is a good choice when native language tests are unavailable; it allows significant mediation, explanation, and practice of the task prior to the administration of the actual items. Even when native language tests are available, the KABC-II provides variable composites systematically related to cultural and linguistic issues as well as flexibility in administration that would make it a good fit for English-first, native-second administrations. The availability of a native-language test that is parallel to an English-language version (i.e., WJ III/Bateria III or WISC-IV/WISC-IV Spanish) accomplishes the same goals, but these may not be as flexible in administration and may not have the advantage of composites.

Additional Procedures

IDEA 2004 reiterates that identification of students with disabilities must be based on use of multiple assessment procedures. This principle is even more important when evaluating students of diverse cultural and linguistic backgrounds. It is recommended that at least two procedures to evaluate intellectual ability be used, such as:

- Standardized administration.
- One or more of the secondary modes of assessment recommended on the MAMBI.
- Testing-of-limits procedures, which could include the assistance of an interpreter.

The purpose of multiple procedures is to confirm the evaluation results and to explore questions and issues that emerging from the initial assessment. For this reason, the initial assessment should be based on the recommendations in the MAMBI and Culture-Language Test Classification (C-LTC), and then use the Culture-Language Interpretive Matrix (C-LIM) to interpret the results, when appropriate. Select any appropriate additional procedures based on these preliminary results.

Next Steps

This chapter provided guidance on the process of gathering data for conducting a comprehensive evaluation. It also listed an array of tools teams may choose from to assist these efforts. The next chapter will discuss how to integrate and interpret the multiple sources of data that teams collect during a comprehensive evaluation, so that they can develop a coherent picture of student performance leading to an eligibility determination.

Table 8-4

Guiding Questions

Guiding Question	Existing Data	Information Needed
How has the team determined the student has had sufficient access to high quality instruction and the opportunity to perform within grade-level standards?		
What supplemental efforts, aligned with grade-level standards, were implemented to accelerate the student's rate of learning and level of performance?		
What, if any, modifications or accommodations are being made within core instruction to enable the student to access content standards?		
What has and has not worked to increase access and participation in core instruction (the general education environment)?		
What educational performance/achievement continues to be below grade-level expectations?		
What factors limit performance? What supplemental efforts have been successful in mediating the impact? What about the student's profile leads the team to suspect a disability and the need for special education and related services?		
How is the student limited from making progress toward grade-level standards?		

Appendix

The Multidimensional Assessment Model for Bilingual Individuals (MAMBI) Created by Ochoa & Ortiz, 2002

Instructional Program/History	Currently in a bilingual education program, in lieu of or in addition to receiving ESL services								Previously in bilingual education program, now receiving English-only or ESL services								All instruction has been in an English-only program with or without ESL services											
	K - 4				5 - 7				K - 4				5 - 7				K - 4				5 - 7							
Assessment Mode	NV	L1	L2	BL	NV	L1	L2	BL	NV	L1	L2	BL	NV	L1	L2	BL	NV	L1	L2	BL	NV	L1	L2	BL	NV	L1	L2	BL
Language Profile 1 L1 minimal/L2 minimal	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓		✓*	✓	✓				✓		✓	✓
Language Profile 2 L1 emergent/L2 minimal	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓	✓	✓*	✓								
Language Profile 3 L1 fluent/L2 minimal		✓				✓	✓			✓				✓	✓													
Language Profile 4 L1 minimal/L2 emergent	✓				✓	✓			✓		✓		✓		✓										✓		✓#	
Language Profile 5 L1 emergent/L2 emergent	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓#	✓	✓				✓		✓	✓
Language Profile 6 L1 fluent/L2 emergent		✓				✓	✓			✓	✓			✓	✓													

Language Profile 7 L1 minimal/L2 fluent						
Language Profile 8 L1 emergent/L2 fluent						
Language Profile 9 L1 fluent/L2 fluent						

CALP Level 1-2 = minimal proficiency; CALP Level 3 = emergent proficiency; CALP Level 4-5 = fluent level of proficiency.

NV = assessment conducted primarily in a nonverbal manner with English language-reduced/acclturation-reduced measures.

L1 = assessment conducted in the first language learned by the individual (i.e., native or primary language).

L2 = assessment conducted in the second language learned by the individual, which in most cases refers to English.

BL = assessment conducted relatively equally in both languages learned by the individual (i.e., the native language and English).

■ = combinations of language development and instruction that are improbable or due to other factors (e.g., Saturday school, foreign-born adoptees, delayed school entry).

✓ = recommended mode of assessment that should take priority over other modes and which would be more likely to be the most accurate estimate of the student's true abilities.

✓ = secondary or optional mode of assessment that may provide additional valuable information, but which will likely result in an underestimate of the student's abilities.

✓* = this mode of assessment is not recommended for students in K-2, but may be informative in 3-4; results will likely be an underestimate of true ability.

✓# = this mode of assessment is not recommended for students in K-1, but may be informative in 2-4, however, results will likely be an underestimate of true ability.

Rubrics for Reading Prosody

Example 1: Multidimensional Fluency Scale for Reading Prosody

Area	1	2	3	4
Expression and Volume	Reads as if just trying to "get words out." Little sense of trying to make text sound like natural language. Tends to read in a quiet voice.	Begins to use voice to make text sound like natural language in some areas but not in others. Focus remains largely on pronouncing words. Still reads in a quiet voice.	Makes text sound like natural language throughout the better part of the passage. Occasionally slips into expressionless reading. Voice volume is generally appropriate throughout the text.	Reads with good expression and enthusiasm throughout the text. Varies expression and volume to match his or her interpretation of the passage.
Phrasing	Reads in monotone with little sense of phrase boundaries; frequently reads word-by-word.	Frequently reads in two- and three-word phrases, giving the impression of choppy reading; improper stress and intonation fail to mark ends of sentences and clauses.	Reads with a mixture of run-ons, mid-sentence pauses for breath, and some choppiness; reasonable stress and intonation.	Generally reads with good phrasing, mostly in clause and sentence units, with adequate attention to expression.
Smoothness	Makes frequent extended pauses, hesitations, false starts, sound-outs, repetitions, and/or multiple attempts.	Experiences several "rough spots" in text where extended pauses or hesitations are more frequent and disruptive.	Occasionally breaks smooth rhythm because of difficulties with specific words and/or structures.	Generally reads smoothly with some breaks, but resolves word and structure difficulties quickly, usually through self-correction.
Pace	Reads slowly and laboriously.	Reads moderately slowly.	Reads with an uneven mixture of fast and slow pace.	Consistently reads at a conversational pace; appropriate rate throughout reading.

NAEP's Integrated Reading Performance Record Oral Reading Fluency Scale

Level 4	Reads in primarily large, meaningful phrase groups. Although some regressions, repetitions, and deviations from text may be present, these do not appear to detract from the overall structure of the story. Preservation of the author's syntax is consistent. Some or most of the story is read with expressive interpretation.
Level 3	Reads primarily in three or four word phrase groups. Some smaller groupings may be present. However, the majority of phrasing seems appropriate and preserves the syntax of the author. Little or no expressive interpretation is present.
Level 2	Reads primarily in two-word phrases with some three- or four-word groupings. Some word-by-word reading may be present. Word groupings may seem awkward and unrelated to larger context of sentence or passage.
Level 1	Reads primarily word-by-word. Occasional two- or three-word phrases may occur, but these are infrequent and/or do not preserve meaningful syntax.

From *Listening to Children Read Aloud* by U.S. Department of Education, National Center for Education Statistics. 1995, Washington, D.C.

References

This section lists reading and tools for assessing cultural and linguistically diverse learners, and articles, books and publications on nondiscriminatory assessment. To assist the reader the references have been clustered by topic.

General Suggested Readings

Carroll, J. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, U.K. Cambridge University Press.

Rhodes, R. L., Ochoa, S.H., & Ortiz, S. O. (2005). *Assessing Culturally and Linguistically Diverse Students: A Practical Guide*.

Flanagan, D.P., Ortiz, S.O., & Alfonso, V.C. (2007). *Essentials of Cross-Battery Assessment*. John C. Wiley and Sons.

Suggested Readings for Cognitive and Intellectual Functioning

Carroll, J. (2005). The three stratum theory of cognitive abilities. In D.P. Flanagan, J. L. Genshaft, & P.L. Harrison, (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 69-76). New York Guilford.

Edwards, O. & Paulin, R. (2007). *Journal of Psychoeducational Assessment* Vol. 25 (4).

Flanagan, D. Oriz, S. & Alfonso, V. (2007). *Essentials of Cross-Battery Assessment Second Edition*. Kaufman, A. & Kaufman, N. Series Editors. John Wiley and Sons, Inc.

Flanagan, D.P., Ortiz, S., Alfonso, V., & Mascolo. (2006). *The Achievement Test Desk Reference: A Guide to Learning Disability Identification, (2nd Ed)* Boston: Allyn & Bacon.

Suggested Readings for Conducting Information Processing Assessments

Dehn, M. (2006). *Essentials of Processing Assessment*. Hoboken, NJ: John Wiley and Sons, Inc.

Flanagan, D.P., Ortiz, S.O., & Alfonso, V.C. (2007). *Essentials of Cross-Battery Assessment*. John C. Wiley and Sons.

Hale, J. & Fiorello, C. (2004). *School Neuropsychology: A Practitioners Handbook*. New York: The Guilford Press.

McGrew, K. (2008). *Introduction to CHC Theory*. Retrieved from <http://www.iapsych.com/CHCPP/CHCPP.HTML>.

Tools

Multidimensional Assessment Model for Bilingual Individuals (MAMBI).

Culture-Language Test Classification matrix (C-LTC).

Culture-Language Interpretive Matrix (C-LIM).

Articles

Ortiz, S. O. (2001). Assessment of Cognitive Abilities in Hispanic Children. *Seminars in Speech and Language, 22*(1), 17-37.

Ortiz, S. O. (1999). You'd never know how racist I was, if you met me on the street. *Journal of Counseling and Development, 77*(1), 9-12.

Ortiz, S. O. & Flanagan, D. P. (1998). Enhancing cognitive assessment of culturally and linguistically diverse individuals: Application and use of selective Gf-Gc Cross-Battery assessment. *The School Psychologist, 52*(1), 6-9.

Books

Cummins, J. C. (1984). *Bilingual and Special Education: Issues in Assessment and Pedagogy*. Austin, TX: Pro-Ed.

Flanagan, D. P. & Ortiz, S. O. (2001). *Essentials of Cross-Battery Assessment*. New York: Wiley Press.

Flanagan, D. P., McGrew, K. S. & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and Gf-Gc theory: A contemporary interpretive approach*. Boston, MA: Allyn & Bacon.

Rhodes, R., Ochoa, S. H. & Ortiz, S. O. (2005). *Assessment of Culturally and Linguistically Diverse Students: A practical guide*. New York: Guilford Press.

Valdes, G. & Figueroa, R. (1994). *Bilingualism and Testing: A special case of bias*. Norwood, NJ: Ablex Publishing.

Chapters and Other Publications

Ortiz, S. O. (in press). Comprehensive Assessment of Culturally and Linguistically Diverse Students: A systematic, practical approach for nondiscriminatory assessment. In C. R. Reynolds and E. Fletcher-Janzen (Eds), *Special Educator's Almanac*. New York: Wiley Press.

Ortiz, S. O. (in press). Cognitive-behavioral school interventions: Multicultural issues. In Christner, R., Menutti, R., & Freeman, A. (Eds.) *Cognitive Behavioral Interventions in Educational Settings*. New York: Brunner-Routledge Publishing.

- Ortiz, S. O., & Lella, S. A. (2005). Cross-cultural Assessment. In S. W. Lee (Ed.), *Encyclopedia of School Psychology*. Thousand Oaks, CA: Sage.
- Ortiz, S. O. & Ochoa, S. H. (2005). Intellectual Assessment: A nondiscriminatory interpretive approach. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment*, 2nd Edition (pp. 234-250). New York: Guilford Press.
- Ortiz, S. O. & Dynda, A. M. (2005). The use of intelligence tests with culturally and linguistically diverse populations. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment*, 2nd Edition (pp. 545-556). New York: Guilford Press.
- Ortiz, S. O. (2004). Nondiscriminatory Assessment in Schools. In C. Spielberger (Ed.), *Encyclopedia of Applied Psychology*, Vol. X, San Diego, CA: Academic Press.
- Flanagan, D. P. & Ortiz, S. O. (2004). Gf-Gc Theory of Intelligence. In T. S. Watson & C. H. Skinner (Eds.), *Encyclopedia of School Psychology*. New York: Kluwer Academic/Plenum Publishers.
- Ortiz, S. O. (2004). Bilingual Multicultural Assessment with the WISC-IV. In A. Kaufman & D. P. Flanagan (Eds.) (pp. 245-254). *Essentials of WISC-IV Assessment*. New York: Wiley Press.
- Ortiz, S. O. (2002). Best Practices in Nondiscriminatory Assessment. In A. Thomas & J. Grimes (Eds.) *Best Practices in School Psychology IV* (pp. 1321-1336). Washington, DC: National Association of School Psychologists.
- Ortiz, S. O. & Flanagan, D. P. (2002). Best Practices in Working with Culturally Diverse Children and Families. In A. Thomas & J. Grimes (Eds.) *Best Practices in School Psychology IV* (pp. 337-351). Washington, DC: National Association of School Psychologists.
- Ortiz, S. O., McGrew, K. S. & Flanagan, D. P. (1998). Gf-Gc Cross-Battery Interpretation and Selective Cross-Battery Assessment: Referral Concerns and the Needs of Culturally and Linguistically Diverse Populations (pp. 401-444). In K. S. McGrew and D. P. Flanagan (Eds.). *The Intelligence Test Desk Reference (ITDR): Gf-Gc Cross-Battery Assessment*. Boston: Allyn & Bacon.
- Ortiz, S. (2008). Best Practices in Nondiscriminatory Assessment. *Best Practices in School Psychology V*. (vol. 2 p 661). National Association of School Psychologists (NASP).