Chapter 4

Linking Assessment to Intervention

The Evidence-Based Balanced Practice Model

Response to Intervention

In our CHT model, emphasis is placed on helping a majority of children through an established MTSS approach. In the multi-tiered comprehensive service delivery model discussed in Chapter 1, assessment and intervention are closely intertwined, with ongoing data collection used to inform classroom instruction at all levels. MTSS is needed to help all children achieve in inclusive general education settings through regular screening and progress monitoring of students, and providing instructional supports before academic or behavioral problems become significant. Not only has research shown this procedure improves academic achievement in a majority of children, but it also reduces the number of referrals for formal comprehensive evaluation and potentially reduces the numbers of children who need special education (Wanzek & Vaughn, 2011). The traditional "test and place" model of school psychology practice required children to wait too long before they were provided with needed services (Erchul & Martens, 2010). These children likely experienced increased academic and/or behavioral problems that had become routinized and were highly resistant to subsequent intervention attempts (Koziol et al., 2013), a point we will return to in subsequent chapters.

An effective multi-tiered approach combines standard protocol instruction and interventions at Tier 1, problem-solving consultation at Tier 2, and individualized interventions in Tier 3 (Hale, 2006). This approach should be adopted to meet the needs of all children regardless of disability status (Flanagan et al., 2010; Hale, 2006). While MTSS is necessary to ensure we do everything we can to help a child succeed, if the child continues to struggle, a comprehensive CHT evaluation is necessary to determine if they have a disability and if so, what type and what to do to help them succeed. If such a need is found, the evaluation can assist in designing targeted interventions, consistent with legislative and legal mandates in the United States (Hale, Alfonso, et al., 2010; Wright et al., 2013). Keep in mind that even if a child is identified as having a disability and requiring an IEP at Tier 3, it doesn't necessarily mean placement outside of the general education classroom. Many children experience academic success in inclusive classrooms with IEPs and instructional accommodations.

In our multiyear study in a lower-SES community of diverse elementary children, we saw steady improvements in reading and math achievement using this combined RTI/CHT approach (Hale, Betts et al., 2010). Interestingly, every child who did not respond to Tier 1 and Tier 2, and was referred for CHT evaluation, was determined to have a disability (100% hit rate). However, only 80% of those who were evaluated in the non-RTI school were determined to have a disability resulting in the completion of more evaluations than in the RTI/CHT schools. So RTI accurately determined who needed a CHT evaluation and who had a disability. In the non-RTI school, about 20% of those children referred did not have a disability following a standard psychological evaluation, and the school psychologists completed approximately 15% more evaluations that year, requiring a lot of extra time evaluating children who didn't need time-consuming and expensive evaluations.

Before we leave these promising results, there is one last point about this study worth your consideration (Hale, Betts, et al., 2010). Some have argued that those students who are nonresponders should be considered to have a specific learning disability (SLD), largely on the basis of achievement data and lack of response to intervention (Fletcher et al., 2004). In contrast, this evidence-based balanced practice model (Hale, 2006) showed that the children should be referred for a comprehensive CHT evaluation with assessment in all areas of suspected disability as required by IDEA (2004) and U.S. Supreme Court precedent (Wright et al., 2013). Throughout this multiyear study, we found that comprehensive CHT evaluations for these struggling students showed that only some of them had SLD. Instead, many had *other* neurodevelopmental disorders requiring special education (e.g., ADHD, depression, bipolar, Tourette syndrome, and speech and language disorders). The CHT evaluation showed us what types of processing deficits these children had, and how these interfered with one or more areas of learning. With the needs for the child with a disability identified, Tier 3 special education services could be provided to develop individualized IEPs, targeted interventions, and instructional modifications/accommodations, with most children served in inclusive classrooms (Wright et al., 2013). The fact that nonresponse to intervention can suggest disability, but not the type of disability or subsequent intervention efforts necessary to help a child succeed, is but one of the reasons why the National Joint Committee on Learning Disabilities (NJCLD) has taken a position against using RTI as the sole means for identifying SLD (e.g., Gartland & Strosnider, 2020).

RTI has many benefits for helping children who struggle with learning, but nonresponse to an intervention should not automatically result in a SLD classification decision (Reynolds & Shaywitz, 2009; Gartland & Strosnider, 2020). Instead, a comprehensive evaluation guided by comprehensive cognitive and neuropsychological assessment is needed to determine *why* they didn't respond and *what* to do next to help them (Reynolds & Fletcher-Janzen, 2009). This combination of standard protocol RTI and problem-solving protocol RTI, followed by comprehensive CHT evaluations for nonresponders, uses best practices to provide all children with the best of both worlds (Hale, Wycoff, et al., 2010).

CHT in Assessment and Intervention

With their expertise in data-based decision making, school psychologists will be viewed as indispensable professionals who support the education of all children in schools. Although this is an important role and function of school psychologists, there are always children who do not respond to the interventions and supports in a MTSS model; these children need something *different* than more intensive instruction (Fiorello et al., 2012; Hale, Wycoff, et al., 2010). Clearly, more intensive instruction and arbitrary decisions regarding tier changes and interventions does not work, and can even be detrimental to student achievement, according to a large-scale, multistate, government funded study of over 20,000 students (Balu et al., 2015). In our comprehensive model, a child who does not benefit from these initial interventions is referred for a formal CHT evaluation. The CHT evaluation is needed prior to Tier 3 special education service delivery to pinpoint areas of strength and weakness that can contribute to developing appropriate interventions, consistent with IDEA (Etscheidt & Curran, 2010) requirements and U.S. Supreme Court precedent (Wright et al., 2013). Figure 4.1 shows the cyclical nature of the CHT evaluation process, which you will note is a scientific method approach to assessment and intervention. In practicing it, you will be operating under the scientist–practitioner model of practice—a model we authors endorse without reservation.

In CHT, the referral question, history, and results of previous interventions are examined to develop a *theory* of the problem. If cognitive functioning is related to the academic or behavioral deficit areas in question, the intelligence/cognitive test is used as one of the first-level assessment tools. Using demands analysis, described later, the findings are interpreted to determine possible cognitive strengths and weaknesses. This process is where many psychologists stop. Because of time demands, psychologists in schools typically write their reports and present their findings in a team meeting; they have little contact with the child, parents, or teacher thereafter (unless individual counseling or behavior therapy is offered). *In CHT, however, we see the intelligence/cognitive test as merely a screening tool of psychological processes, not a final assessment of these skills.* Best practices and our CHT model go beyond this screening level assessment to choose additional measures to confirm or refute the intellectual test data. The results are examined in light of the MTSS data, record review, history, systematic observations, behavior ratings, and parent and teacher interviews to gain a good understanding of the child.

Completing the initial assessment is where the CHT process *begins*, not ends. Interventions are subsequently developed using the understanding of the child and the environment during collaborative consultative follow-up meetings with teachers, parents, and/or children. If a child is identified with a disability, these interventions may become part of a child's IEP, or if not, they can help inform Tier 2 problem-solving interventions, with both Tier 2 and Tier 3 service delivery preferably happening in inclusive general education classrooms. Even if a child has an IEP, we still do problem-solving consultation in Tier 3. The ideas of problem identification, problem analysis, intervention development/implementation, and intervention evaluation permeate CHT Tier 3 service delivery, but we recognize the value of cognitive and neuropsychological data in the problem analysis part. From this perspective, all data are relevant in data-based decision making.

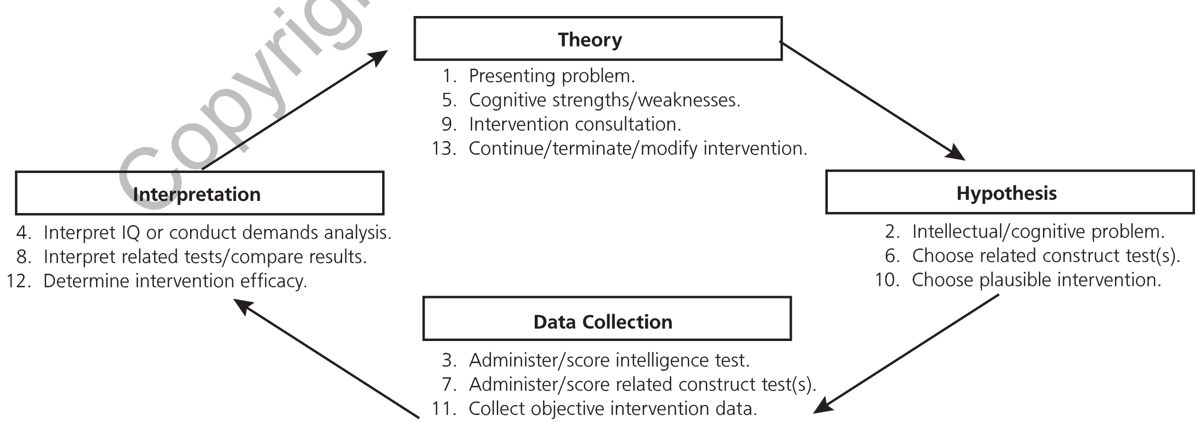

FIGURE 4.1. The CHT model.

Possible intervention strategies are brainstormed in consultation with the teacher; design, resources, and feasibility are considered; and an intervention plan likely to succeed is developed. The systematic intervention is then undertaken and evaluated to determine intervention efficacy. If the Tier 3 progress monitoring data suggest the intervention is not effective, the intervention is revised or recycled until beneficial results are obtained. We recommend using known evidence-based academic and behavioral interventions and single-subject methodology. We also use information about cognitive and neuropsychological functioning in developing and individualizing our evidence-based interventions. Understanding the child's psychological processing strengths and weaknesses in relation to his or her academic and behavioral needs allows us to truly differentiate instruction to ensure children have optimal educational outcomes (Hale et al., 2018).

Conducting Demands Analysis in CHT

Demands analysis is a core component of the CHT model. It is the key to both accurate identification of childhood disorders and to the development of interventions sensitive to individual needs. The demands analysis process that we present here is derived from two assessment traditions and modern neuroscience perspectives on brain functioning. The first tradition is the "intelligent testing" approach, which examines global, factor, and subtest scores based on clinical, psychometric, and quantitative research (Flanagan et al., 2013; Sattler, 2020). When formulating your clinical demands analysis, you must be careful to examine all relevant technical and cross-battery subtest information. As stated earlier, one cannot deny the value of a CHC-based interpretive approach, especially when there is a convergence between CHC and modern understanding of brain functions (Fiorello et al., 2008). Heavily influenced by the Lurian (1973b) approach to neuropsychological assessment, the second tradition consists of the developmental and process-oriented neuropsychological assessment approaches (e.g., Gable et al., 2015; Bernstein, 2000; Lezak, 1995; Libon et al., 2013; Poreh, 2012; Radel et al., 2017; Schneider et al., 2013). Although demands analysis may seem similar to other versions of profile analysis (Kaufman et al., 2016), the major difference is the emphasis on the neuropsychological and cognitive processes necessary for task completion.

As we noted earlier, *you will not see reference to brain structures or systems* in our evaluation reports (unless medical evidence is present in the history). We do use our *knowledge of brain structure and function help guide us in the interpretive process*. We have noted previously that the input and output demands are straightforward; they are the observable and measurable test stimuli and behavioral responses. However, research is clearly demonstrating that the underlying neuropsychological *processing demands* are essential for understanding and helping many children with their learning and behavior problems (Fiorello et al., 2012; Hale, Fiorello, et al., 2010). In other words, the key to case conceptualization, for both identification and intervention purposes, is understanding psychological processes, not input or output. This is in some ways a bane, because we can see input and output, but not processes. Only careful inferences drawn from idiographic interpretation of cognitive and neuropsychological measures can tell us about these processes in individual children, so it is both costly and time consuming. It is also a blessing because it is the clinician's understanding of psychological processes that can be critical for understanding and helping children. This dichotomy is why MTSS is essential practice, because it ensures that any child referred for a CHT evaluation is likely to have a brain-related disability that continues to interfere with academic and/or behavioral functioning despite our best attempts at early intervention. Recall from the very beginning of the book—we have been arguing that you need to *intervene to assess* (and do both well).

For many children and most tests and subtests, a brief demands analysis should be sufficient to examine and test hypotheses about brain–behavior relationships. We have provided you with two forms (Appendix 4.1 and Appendix 4.2) to guide you in interpretative efforts. The form in Appendix 4.2 may even be more helpful as you become more accustomed to demands analysis, because this allows you to add constructs as necessary to reflect the neuropsychological processes underlying a particular subtest or if a child responds in an idiosyncratic manner.

To conduct the demands analysis, identify tests/subtests that represent the child's strengths and weaknesses. Enter them in the appropriate spaces in Appendix 4.2, and for each measure conduct a task analysis of the *input, processing,* and *output* demands. *Input* refers to the stimulus materials as well as the directions, demonstrations, and teaching items. Think about what modality or modalities are needed for the input—for example, whether there are pictures or verbal directions, whether the content is meaningful or abstract, and what other aspects of the content are relevant (e.g., level of English language used or amount of cultural knowledge required). *Processing* refers to the actual neuropsychological processing demands of the task, as discussed in Chapters 2 and 3. Think about the primary requirement of the task, but also the secondary requirements, such as the attention, executive, and working memory demands (often suggested by the test's developers and the CHC cross-battery approach; Flanagan et al., 2013). *Output* refers to the modalities and skills required for responding to the task. Is the output a complex verbal response, a simple pointing response, or a complex motoric response? If oral expression is needed, is syntax important, and is word choice an issue? How is the child's response pattern reflective of the main process demands of the tasks, and how are other processes influencing within- and between-task performances? These are some of the questions you must answer in demands analysis. Always keep in mind sociocultural differences and expectations during demands analysis, since different responses could reflect these characteristics (e.g., use of nonliteral language by indigenous populations, lower emphasis on processing speed in Latino populations).

The form we provide in Appendix 4.1 is merely a tool for you to begin thinking about underlying psychological processes on your cognitive, neuropsychological, *and* achievement tests. Yes, the neuropsychological process approach not only helps you interpret cognitive and neuropsychological tests, *but also achievement tests*, because the same brain processes both! We have included blanks in the last column for you to provide additional subtest input, processing, and output demands. Once you have listed the input, processing, and output demands for all of the child's strengths and weaknesses, it is important to look for commonalities and contradictions among the data using the CHT methodology, thereby avoiding confirmation bias as a result.

As you become familiar with processing demands and have experience in interpreting the processing differences a child demonstrates, you need not always refer to demands analyses sheets. When you are beginning the process, you may want to do it for many measures that are factorially complex (i.e., have a lot of demands at the same time). After completing demands analysis for the measures, you attempt to identify the child's patterns of performance within and between measures and attempt to verify processing strengths and weaknesses using multiple data sources. If you find that one particular processing demand is required on all low-score tests, and it is not needed for the high-score tasks, you would hypothesize that this demand is a weakness for the child. Information from your observations of the child during testing, as well as information provided by the teacher and possibly parents, should also be consistent with any definitive hypothesis to ensure concurrent and ecological validity. The weakness may be a cognitive processing weakness, but it may also be a sensory or motor weakness, a result of emotional interference, or a consequence of limited exposure or background. Enter this information on the worksheet provided in Appendix 4.3.

There is an important issue to consider before we move on. The key to this demands analysis process is generating and evaluating hypotheses. However, care is needed. You have to avoid confirmation bias in test interpretation when you look for support for your hypotheses. When teaching demands analysis, we ask students to generate hypotheses and when just learning the approach they often come up with a single probable hypothesis, because our cognitive tendency is to go to the left hemisphere to come up with the "right" answer. However, to encourage discordant/divergent thought, we say "Even if that explanation seems likely, what else could it be?" We also ask students to always include a null hypothesis, forcing them to consider that the child might not have a processing deficit or disability. This approach ensures we keep the diagnostic door open until we have convincing evidence that our understanding of the child is accurate.

Although the forms in the appendices and several interpretive texts (Flanagan, 2013; Kamphaus & Campbell, 2006; Kaufman et al., 2016; Sattler, 2020) can be helpful in conducting demands analysis, you should not be lulled into a "cookbook" approach when interpreting subtest data—a tendency that often results in erroneous interpretation. You can't have template reports where you describe a factor or subtest processes, and change the wording to reflect good, average, or poor performance on that construct. Recall that the children can use different brain areas and psychological processes to perform any given task, so you can't just say the task measures "construct A" and then change the level of performance descriptor for each child. To guard against this and to foster accurate interpretation, we have provided a checklist in Appendix 4.4. This checklist serves primarily as an aid in clinical judgment, but it could also be used as an informant rating scale.

Let's walk through a demands analysis of the WISC-V Block Design subtest to see what the process looks like. First, consider the input. The task has oral directions and is modeled for children and corrected for those who have difficulty on the first item. The stimulus materials (booklet with two-dimensional visual model and two-color three-dimensional blocks) are abstract, colored shapes, so that verbal encoding is difficult (but not impossible for those compensating for visual–spatial processing weaknesses). The task will be novel for most children (although perhaps not on re-evaluation or as the testing progresses). The processing demands are quite complex and involve both hemispheres and executive/frontal demands. Primarily, Block Design is a right-hemisphere task, since it is visual–spatial (i.e., involves the dorsal stream), is novel, and does not depend on crystallized prior knowledge. There is some bilateral processing, however, because of the bimanual sensory and motor coordination, as well as the part (directional orientation of the blocks—left parietal) and whole (gestalt/spatial—right parietal) coordination (Hain & Hale, 2010; Poreh, 2012; Swanson et al., 2013). Some children may walk themselves through this task using verbal mediation—this behavior should be noted because it might indicate that the visual–spatial skills are not as well developed as would be expected.

There are heavy frontal demands, due to the executive and motor requirements of the task. The frontal demands include planning and organization, analysis and synthesis, self-monitoring and evaluation of the response, inhibition of impulsive responding, and fine-motor and bimanual coordination. This is particularly true if the child uses a trial-and-error match-to-sample approach to the task (i.e., comparing/contrasting model to blocks, quickly flipping them repeatedly until the correct "design" is achieved). Note particularly if the child has more difficulty after the lines are removed from the stimulus book, as this may suggest difficulty with configuration or novelty. Considering the output, Block Design requires fine-motor and bilateral motor coordination, so look for problems crossing the midline or a tendency to use just one hand, or switch hands and not use them together, which suggests possible corpus callosum problems, which are surprisingly common in some neurodevelopmental disorders for which white matter development is a problem. Processing speed can also impact performance on Block Design. Look for slow performance due to overall low tone or lethargy, perfectionism, or inattention/disorganization. Also, if the child is close to completion, let the child complete the design even if time runs out. While you cannot score the design as correct, it tells you that the child's visual–spatial skills are not the problem—timing is.

Although conducting demands analysis may be helpful in understanding patterns of performance, remember that multifactorial tasks can be solved in more than one way, so that the demands analysis may differ from child to child. For instance, a child who uses good executive and psychomotor skills to compensate for a right posterior spatial problem may still do well on Block Design. It is an error if you conclude that the child had adequate visual–spatial–holistic processing skills based solely on this measure. We practitioners have often gone wrong in the past by concluding that the same subtest measures the same thing for all children. For instance, concluding that poor WISC-V Information subtest performance is due to a limited "fund of information" may not be correct if a child has retrieval problems or difficulty due to limited knowledge in just one area, such as science. Concluding that a child has adequate attention, working memory, and executive function because he or she has an average WISC-V Digit Span scaled score, but a Digits Forward score of 10 and a Digits Backward score of 2, would clearly be inappropriate (Hale et al., 2002). In addition, there can be developmental aspects of test items that influence interpretation as well; if a young child scores well on Similarities, but most items are scored 1 point, it may suggest the child has good lexical–semantic knowledge and concordant–convergent thought, but you wouldn't want to say she has good abstract verbal concept formation; whereas another child could get the same score but responds mostly with 2-point responses but the ceiling is achieved earlier. These examples bring us to a potentially disconcerting but a very real conclusion about psychological test interpretation: *the same score does not always mean the same thing for all children, even of the same age.*

Table 4.1 provides you with some sample demands analyses on a few additional subtests, so you can see how the process works. As you become more familiar with using demands analysis to task-analyze subtests, you will eventually become quite comfortable with determining the demands on any subtest. As you gain confidence, and knowledge of brain–behavior relationships, you will be surprised how easy it is to understand the psychological processes required on any task or behavior.

In our graduate child neuropsychology assessment class, we have a final exam item that requires students to do a "mystery test" demands analysis on a test they have not been exposed to in class. Though students find the thought of this daunting, and the task challenging, they typically find that they can identify the key input, processing, and output demands on the test, and this gives them incredible confidence in their budding idiographic interpretive skills. Try this activity yourself. Generalizing these skills to other measures will allow you to expand your use of demands analysis to just about any instrument you are trained to administer. It can also be generalized to almost any behavior. In another graduate school task, students pick an everyday activity (e.g., making coffee, brushing their teeth, greeting a friend) and determine the brain structures and processes involved in each of the steps. Again, this is a novel task students find daunting, but working through it helps them become more comfortable in observing overt behavior and thinking about the neuropsychological processes involved in completing it.

We now turn to a discussion of neuropsychological tests for use in the CHT model. Although many of these tests may be new to you, the demands analyses you perform on cognitive and intellectual measures apply to neuropsychological measures as well. Do not let yourself be overly concerned that these measures are "neuropsychological"; many of them are easier to administer and score than the measures you generally use. For instance, the D-KEFS Color–Word Interference Test requires approximately 5 minutes to administer, and it has brief, simple instructions. Even though it is easy to administer, this test is highly sensitive

TABLE 4.1. Sample Demands Analysis of Selected Subtests

WISC-V Block Design

Input
- Models and abstract visual pictures
- Oral directions—moderate English-language knowledge
- Demonstration/modeling
- Low cultural knowledge and emotional content

Processing
- Visual processing (spatial relations, visualization)
- Perception of part–whole relationships
- Discordant/divergent processing (analysis)
- Constructional praxis
- Bimanual coordination/corpus callosum
- Concordant/convergent processing (synthesis)
- Attention and executive demands: Moderate
- Planning and strategy use
- Inhibition of impulsive/wrong responding
- Novel problem solving: Low to moderate

Output
- Fine-motor response, arrangement of manipulatives
- Timed score with speed bonus; process score without time bonus
- Visual–motor integration

SB-5 Picture Absurdities (Levels 4, 5, and 6—Nonverbal Knowledge)

Input
- Large color pictures
- Oral directions
- Sample item
- High cultural and English-language knowledge

Processing
- Visual scanning
- Perception of objects (ventral stream)
- Crystallized ability for prior knowledge (left temporal)
- Discordant/divergent processing (analysis)
- Attention and executive demands: Low to moderate
- Persistence/inhibition of impulsive responding
- Novel problem solving/reasoning

Output
- Brief oral or pointing response
- One right answer (convergent responding)

WJ-V Visual–Auditory Learning

Input
- Brief oral directions, teaching items, feedback
- Semiabstract figures/symbols
- Moderate cultural and English-language knowledge

Processing
- Visual perception of figures/symbols (dorsal and ventral streams)
- Sound–word/symbol–rebus association

(*continued*)

TABLE 4.1. (*continued*)

- Working memory/learning
- Encoding and retrieval of associative/semantic memory
- Benefiting from feedback
- Inhibition of impulsive/wrong responding
- Syntax knowledge: Helpful
- Attention and executive demands: Moderate to high
- Memory: primary; novel problem solving: secondary

Output
- Brief oral response
- Oral formulation/retrieval

KABC-II NU Pattern Reasoning

Input
- Brief oral directions; sample and teaching items
- Abstract/nonmeaningful figures
- Low cultural knowledge and English-language knowledge

Processing
- Visual scanning and discrimination
- Color processing
- Visual–spatial processing (dorsal stream)
- Part–whole relationships
- Discordant/divergent processing (perceptual analysis)
- Novel problem-solving and inductive reasoning/fluid abilities
- Attention and executive demands: Moderate
- Inhibition of impulsive/wrong responding

Output
- Pointing response
- Multiple-choice format (can solve by elimination/match to sample)

to executive functions and to frontal–subcortical circuit dysfunction (particularly cingulate function), and therefore is an excellent supplement to test hypotheses generated by your initial battery. Of course, you must always remember that no one test can diagnose a specific difficulty or disability.

Assessment Tools for CHT

Fixed versus Flexible Batteries in Hypothesis Testing

One of the biggest debates in neuropsychological assessment is whether to use a fixed test battery (a standard set of tests) or a flexible battery (a set of tests chosen for an individual child) (Decker et al., 2012; Lezak et al., 2012; Riccio & Reynolds, 2013). Fixed batteries predominated early in the field's history, but flexible batteries have become increasingly popular, especially since they tend to be more time and cost effective. Fixed batteries tend to lead to more testing than is needed to address unique child characteristics. One of our biggest complaints about traditional neuropsychological assessment is that children are tested all day long on a battery of tests that may not be needed for interpretation. For instance, if explicit language skills are in the superior range on the intellectual/cognitive task, is it worthwhile to give a measure of verbal learning and memory? In addition, a fixed battery gives examiners the impression that the battery assesses all relevant neuropsychological domains, even though that is not necessarily true (Lezak et al., 2012).

We prefer a flexible-battery approach in the CHT model, because different measures and techniques can be used to address hypotheses developed after initial data gathering and intellectual/cognitive screening tool assessment. You may need one or more measures that look at a particular domain in depth. For instance, if you have a child who has handwriting problems you don't just give your intelligence test and say the child has "graphomotor" or "visual–motor integration" problems; that doesn't tell us what we need to do to help the child and there is actually no direct measure of graphomotor skills on that test! Instead, if you're interested in an apparent visual, somatosensory, motor, or integration deficit, you need to pick and choose measures that tap each of these four possible causes to get a better understanding of the problem and the direction to take for intervention. Be aware of the problem of confirmation bias, however. If you only choose tests to confirm a hypothesis, you are likely to find evidence to support it. Choose tests to disconfirm your hypotheses as well. For instance, when looking at the graphomotor problem you may be thinking about cortical problems, and sure enough you find it is largely a motor problem, but you have to think of multiple causes of motor problems, including subcortical structures such as the basal ganglia and cerebellum (Koziol & Budding, 2009). You should also ensure that you are at least screening all the major functioning areas so as not to overlook something. To further our visual–spatial example, a child may be able to copy a figure when there is a space provided for that copy (i.e., the Developmental Test of Visual–Motor-Integration), but may experience significant problems with the Rey–Osterrieth or Bender–Gestalt tests which are less structured and require the child to organize the information. These measures that are "purported"' to test the same skill, are different, and allow you to think about other aspects of the task that may be interfering with the child's performance (visual organization, attention to detail, etc.).

We are not suggesting that a fixed-battery approach should be completely avoided. Some neuropsychologists prefer a standard set of tests or published test battery, because all the children tested are administered the same tests in the same order, which increases the validity of interpretation in many practitioner's eyes, especially in many forensic settings (Kaufmann et al., 2013; Russell et al., 2005). Fixed batteries can also serve both research and practice needs. Obviously, many children who receive the same measures would be needed for a group-design research project. Missing data are the nemesis for most research projects, because many cases need to be eliminated if they are missing just one variable in a multivariate statistical analysis. Given that fixed batteries have led to many research findings, they are often recognized for strong reliability and validity, and for the use of normative data (Witsken et al., 2008). For clinicians, fixed-battery approaches not only help standardize performance expectations across children, but also allow practitioners an opportunity to develop "head norms" about child performance. It is much easier to interpret a measure after dozens of regular administrations than if it is used sparingly to test hypotheses for individual children. In addition, once demands analyses have been done on the fixed-battery subtests, they may only need to be changed slightly for children who perform them in a unique way. Finally, the use of a fixed battery does not preclude additional hypothesis testing with other instruments. Actually, using an intellectual/cognitive measure (e.g., WISC-V), a fixed neuropsychological battery (e.g., the Halstead–Reitan), and additional hypothesis-testing measures (e.g., Comprehensive Test of Phonological Processing, Second Edition, subtests) might be the ultimate approach for conducting CHT. However, it is important to remember that as the number of measures increases, the likelihood of child performance variability and of Type I error increases as well. It does no good to test a child all day, only to find his "processing deficits" occurred on the late afternoon tests. In other words, if fatigue is too great, performance deficits can be expected and not necessarily indicative of the real problem.

Intellectual Tests for Hypothesis Testing

You may be surprised to find that you are already familiar with many of the tools available for CHT—including the intelligence/cognitive tests discussed in Chapter 1. Although intelligence test subtests are typically factorially complex (Flanagan et al., 2013), there is often a wealth of information published about these measures; their technical quality can be thoroughly evaluated; and you are familiar with their scoring and interpretation. The manuals on these measures come with many statistics to support interpretation, such as reliability, standard deviations, standard error of measurement, correlations, factor analyses, and other validity studies.

To aid in your demands analysis of these and other measures, it is worthwhile to consult *Essentials of Cross-Battery Assessment* (Flanagan, Alfonso, & Ortiz, 2025), which specifies subtest technical characteristics from a Cattell–Horn–Carroll perspective, and Sattler's (2024) *Assessment of Children: Cognitive Applications* text. Similarly, CHT of the skills necessary for academic performance can utilize subtests from several achievement batteries. For instance, the Kaufman Test of Individual Achievement—Third Edition (KTEA-3) includes subtests that can be used to assess fluent retrieval of lexical–semantic information and rapid automatic naming (Associational Fluency and Naming Facility). Recall that your understanding of brain–behavior relationships is important for interpretation of all test data, and in several subsequent chapters you will get a better understanding of the psychological processes involved in academic achievement.

Although these intellectual and achievement instruments are useful in CHT, let us now examine several tests that are often considered "neuropsychological" instruments. It is important to realize that many neuropsychological tests are easy to administer and score, and that they tap many of the constructs already discussed in this book. However, some of these tasks are quite challenging to administer (especially if they tap your particular processing weaknesses). It will take some practice and/or supervised experience depending on your background before you can become proficient in their use for CHT. We do not claim to present an exhaustive list of measures here, just those that we have found to be useful in our practice of CHT. We do not suggest that these measures are better than others, or that measures not included here cannot be adopted in the CHT model. You should complete a demands analysis for measures you are not familiar with and review the extant literature on new tests before you use them. Do not automatically assume that a test measures what we suggest, or what the test authors report in the manual. Although our interpretive information is limited in this volume, you can consult the test manuals and other excellent interpretive texts to aid in your understanding of the measures (Lee et al., 2005; Lezak et al., 2012; Miller, 2019; Fletcher-Janzen & Reynolds, 2010; Riccio et al., 2010; Semrud-Clikeman & Ellison, 2009; Strauss et al., 2006). Your background, training, and experience will determine your need for individual training and supervision on these measures.

Traditional Neuropsychological Test Batteries

We begin our review of instruments by discussing two historically important neuropsychological test batteries (NTBs): the Halstead–Reitan NTB (Reitan & Wolfson, 1993) and the Luria–Nebraska NTB (Golden et al., 1985). Both have versions for children that are downward extensions of the adult batteries. Though we aren't advocating that school psychologists use these batteries, a brief description follows to familiarize you with them. These batteries are often used as "fixed" batteries, and both have a long tradition of use in neuropsychological assessment and research, so there are many supplemental resources and publications to aid in their interpretation.

Halstead–Reitan NTB

Table 4.2 provides an overview of the constructs tapped by the Halstead–Reitan NTB (Reitan & Wolfson, 1993) subtests, and of possible brain areas responsible for performance. Reviews of its empirical evidence and clinical utility can be found in Nussbaum and Bunner (2009), and Ross et al. (2013). The Category Test requires the child to view simple objects on a screen and press a button coinciding with the numbers 1 to 4. The child is not told how to perform the task, but instead receives feedback after each response. (A more recent version of the Category Test is mentioned later in Table 4.9.) For the Tactile Performance Test, the child is blindfolded and presented with an upright form board and shapes. The child places the different shapes in the corresponding holes as quickly as possible, first with the dominant hand, then with the nondominant hand, and then with both. This is an interesting task in that the input is tactile instead of visual to evaluate visual object recognition and memory. However, be aware that some children find being blindfolded disconcerting or even aversive.

The Trail Making Test is a connect-the-dots task, where the child draws a line connecting numbers in order (Trails A), and then alternates between numbers and letters (Trails B), as quickly as possible, tapping executive functions. For the Sensory–Perceptual Examination, a brief screening of visual, auditory, and somatosensory functioning is followed by three somatosensory tasks: finger touching, writing of numbers (older children) or symbols (young children) on fingers/hands, and recognition of shapes, all hidden from the child's view. The Finger Tapping test is a simple measure of motor speed and persistence.

The Halstead–Reitan provides an Impairment Index of brain dysfunction/damage, which ranges from 0 to 10. It has shown good reliability and validity for identifying brain damage,

TABLE 4.2. Characteristics of Halstead–Reitan Neuropsychological Test Battery (NTB) Subtests

Subtest	Constructs purportedly tapped	Brain areas involved
Category Test	• Concept formation, fluid reasoning, learning skills, mental efficiency	• Prefrontal area, cingulate, hippocampus, temporal lobes (associative and categorical thinking)
Tactile Performance Test	• Tactile sensitivity, manual dexterity, kinesthetic functions, bimanual coordination, spatial memory, incidental learning	• Lateralized sensory and motor areas, parietal lobes, corpus callosum, hippocampus
Sensory-Perceptual Examination	• Simple and complex sensory functions	• Lateralized sensory areas (more complex, bilateral?)
Finger Tapping	• Simple motor speed	• Lateralized motor areas
Trail Making Test, Parts A and B (Trails A and B)	• Processing speed, graphomotor coordination, sequencing, number/letter facility (Trails B also requires working memory, mental flexibility, set shifting)	• Trails A: Dorsal stream, premotor area, primary motor area, corpus callosum; Trails B: also prefrontal–basal ganglia–cingulate

and has been used for identifying strengths and weaknesses for children with learning disabilities and other disorders (Russell et al., 2005). However, the test is quite dated, having been last updated in 2004 for adults and children 15 and older (Heaton, 2004). Versions designed for children have inadequate norms (Nussbaum & Bunner, 2009) and users are cautioned about its use with children (D. C. Miller, 2019; Semrud-Clikeman & Teeter Ellison, 2009). Nonetheless, it is important to be familiar with the tests in this battery, as many of the measures have formed the basis for many more recently developed tests.

Luria–Nebraska NTB

The Luria–Nebraska NTB (Golden et al., 1985) consists of 12 scales derived from Luria's (1973b, 1980a, 1980b) approach to neuropsychological assessment, which emphasizes flexible administration and interpretation of measures. Therefore, it is not a true fixed battery per se, but practitioners may have a tendency to administer it as such. The 12 Luria–Nebraska subscales are labeled Motor, Rhythm, Tactile, Visual, Receptive Language, Expressive Language, Writing, Reading, Arithmetic, Memory, Intelligence, and Delayed Memory. The most predictive subscales are those assessing language and achievement, which are adequately assessed using more current instruments, as the norms are quite limited (Semrud-Clikeman, 2001). Although some have lauded the virtues of this battery (Golden, 2004), the traditional examination may take up to 2 days to complete (Golden, Freshwater, et al., 2001), making this instrument impractical for use in the schools, and in any case is seldom used with children anymore (Miller & Wang, 2019; Semrud-Clikeman & Teeter Ellison, 2009). It has not been updated since the 1980s, and several contemporary neuropsychological assessment tools are available (as discussed later) to assess skills similar to those tapped by the Luria–Nebraska domains, and many were designed solely for use with children. Note that A. R. Luria was not involved in the development of the measures.

Neuropsychological/Cognitive Tests for Hypothesis Testing

We now review instruments that assess multiple as well as specific areas of neuropsychological functioning. You may wish to use an entire test at times, but for the most part, you will pick and choose subtests from these batteries for CHT. They are listed in alphabetical order as to not suggest a preference for one over another for use in CHT.

Children's Memory Scale

Since we are often asked to give an indication of a child's capability of learning in the classroom, it is somewhat surprising that more educational administrators don't mandate assessment of learning and memory skills. Designed for use with children ages 5–16, the Children's Memory Scale (CMS; Cohen, 1997) was an excellent measure of learning and memory designed for clinical assessment. It was carefully standardized on a representative sample; however, its norms are now quite dated, leading us to recommend use of a more current measure. Nonetheless, it is not surprising that the CMS demonstrates adequate internal consistency for a memory measure, and comprehensive validity studies support the instrument's construct validity.

The CMS has six core subtests, two each in the Auditory/Verbal, Visual/Nonverbal, and Attention/Concentration domains; the last domain is probably the least useful in CHT. In addition, there are three supplemental subtests, one for each domain. The subtests we typically use are presented in Table 4.3. The reported subtests all have delayed portions for further examination of long-term memory retrieval—an advantage of this measure. A disadvantage is relying on the Auditory/Verbal–Visual/Nonverbal dichotomy for organizing the battery, and the norms are now quite old. For more information about the utility of the CMS, please consider these sources (Hildebrand & Ledbetter, 2001; Kibby & Cohen, 2008; Riccio et al., 2007).

Cognitive Assessment System—Second Edition

The Cognitive Assessment System—Second Edition (CAS2; Naglieri et al., 2014) retains the essential structure of the CAS, assessing cognitive functioning in children and adolescents from 5 to 18 utilizing the authors' planning, attention, simultaneous, and successive (PASS) model (Das et al., 1994). Although it is reported to be based on Luria's model of neuropsychological processing and assessment, as we have seen in Chapter 2, there is no PASS acronym in Luria's model, and the authors recommend nomothetic interpretation of PASS factors, which is essentially not a Lurian-type approach to clinical assessment. On the one hand, the authors' confirmatory factor analysis has been used to support a four-factor model, but cross-battery analyses of the first edition have raised doubt about the model, with findings suggesting that the Planning and Attention factors should be combined (Georgiou et al., 2020; Keith et al., 2001). On the other hand, Planning and Attention have differential predictive validity of outcome, supporting the authors' model. This model would certainly fit with Luria's (1973b) theory, as attention and executive functions are intimately related to the integrity of the third functional unit or frontal lobes (except for cortical tone, which would be the responsibility of Luria's first functional unit). The separation of planning and attention leads to different recommendations, supporting the treatment validity of the PASS model (Das et al., 1996). In addition, the CAS has been shown to have construct validity and diagnostic utility for children with ADHD (Canivez & Gaboury, 2016) and the factor scores predict achievement well (Naglieri & Rojahn, 2004) and have been shown to be related to intervention (e.g., Iseman & Naglieri, 2011).

TABLE 4.3. Characteristics of Children's Memory Scale (CMS) Subtests

Subtest	Constructs purportedly tapped
Auditory/Verbal	
Stories	• Auditory attention, semantic long-term memory encoding and retrieval, sequencing/grammar, verbal comprehension, expressive language
Word Pairs	• Paired-associate task; auditory attention, learning novel word pairs
Word Lists	• Selective reminding task; long-term memory encoding, storage, and retrieval of unrelated words
Visual/Nonverbal	
Dot Locations	• Visual–spatial memory encoding and retrieval (dorsal stream), susceptibility to interference
Faces	• Visual–facial memory encoding and retrieval (ventral stream)
Attention/Concentration	
Sequences	• Rote recall of simple information followed by mental manipulation/executive function items

We like several of the CAS2 subtests for hypothesis testing (albeit the CAS2 authors would not support such use in actual practice). The scale was adequately normed, and most subtests show good technical characteristics. In addition, the test authors have provided us with the first substantial treatment validity studies of any cognitive measure, presented in the PASS Remedial Program (PREP; see Das et al., 1997). The PREP has focused primarily on reading, with training of successive and simultaneous skills leading to improved word recognition and decoding skills. There is also evidence that strategy-based instruction can improve mathematics achievement in students with poor planning skills. We do not think, however, that the CAS2 should be used to measure global intellectual functioning, even though it provides a Full Scale standard score (SS). Absent from the CAS2 is a measure of crystallized intelligence (*Gc*). Although the lack of *Gc* measurement makes the CAS2 a fairer test for people of linguistic and cultural difference than most other intellectual/cognitive measures, it doesn't adequately tap left-hemisphere processes as a result. Therefore, though we feel that the CAS2 is not adequate as a baseline measure of global functioning for most children, it is a good tool for hypothesis testing.

We present the CAS2 subtests that we recommend in Table 4.4. Note that our interpretation is somewhat different from that presented by the test authors. Naglieri and colleagues have conducted some useful intervention research with the original version of the CAS showing the link between PASS processing strengths and weaknesses and intervention choice, and linking assessment to intervention that leads to improved outcomes (Goldstein et al., 2011; Haddad et al., 2003; Iseman & Naglieri, 2011; Naglieri 2002; Naglieri & Johnson, 2000; Tomporowski et al., 2008). This is noteworthy given our earlier complaints about test publishers not examining the treatment validity of their measures

Comprehensive Test of Phonological Processing, Second Edition

The Comprehensive Test of Phonological Processing—Second Edition (CTOPP-2; Wagner et al., 2013), is a unique measure of the cognitive constructs most commonly associated with reading and language disorders. Designed for use with children and youth aged 4–24, it measures phonological awareness, phonological memory, and rapid automatized naming, which have been linked with word recognition, word attack, and other basic reading skills (Wolf, 2001). The CTOPP-2 is composed of 12 subtests, several of which we find useful in CHT. It was normed on a large representative sample, and subtests have good to excellent technical characteristics. Validity studies show the phonological awareness and rapid naming tasks have strong relationships with reading skills.

Table 4.5 outlines the CTOPP-2 subtests and what they measure. The Nonword Repetition subtest is an interesting task that taps phonemic processing and expression skills for nonsense words (e.g., "lidsca"), similar to other visually presented pseudoword tasks for comparison purposes. However, it includes an auditory model (so the child hears the nonword first) and an auditory working memory component (because the child has to recall what he or she heard). This task can be combined with the Blending Nonwords (e.g., "raq" + "di") subtest to help determine whether the phonological breakdown is occurring at the individual-phoneme level or the phonological assembly level. One concern with the CTOPP-2 is the limited assessment of rapid naming. Including rapid naming of more complex letter combinations (e.g., digraphs, diphthongs) and simple words presented two grades below reading level would have been helpful. Although phonological processes have been linked to left temporal lobe functions, rapid naming is typically associated with temporal lobe and frontal–subcortical circuits, as well as cerebellar functions. Further information about brain functions and reading competency are discussed in Chapters 2 and 5. The first and second editions of the CTOPP have been used extensively in research, attesting to its value in

TABLE 4.4. Characteristics of Cognitive Assessment System, Second Edition (CAS2) Subtests

Subtest	Constructs purportedly tapped
Planning	
Planned Number Matching	• Sustained attention, visual scanning, psychomotor speed, noticing pattern and figuring out appropriate strategy
Planned Codes	• Similar to WISC-V Coding, but format allows for strategy use (e.g., filling in by code rather than in order)
Planned Connections	• Substitute for Halstead–Reitan Trails A and B (see Table 4.2), but no separation of scores
Attention	
Expressive Attention	• Substitute for Stroop Color–Word Test (see Table 4.9); inhibition of automatic response (reading words) to name ink color of printed word
Number Detection	• Cancellation task; sustained attention, visual scanning, visual discrimination, inhibition, psychomotor speed
Simultaneous Processing	
Nonverbal Matrices	• Typical *Gf* measure of inductive reasoning; multiple-choice format
Verbal/Spatial Relations	• Receptive language, verbal working memory, grammatical relationships, visual scanning/discrimination
Figure Memory	• Similar to DAS-II Recall of Designs (see Chapter 1, Table 1.2); visual perception, spatial relationships, visual memory, graphomotor reproduction, constructional skills, figure–ground relationships
Successive Processing	
Word Series	• Word span; rote recall of unrelated words
Sentence Repetition	• Rote recall of meaningless sentences; grammatical structure important (ages 4–7)
Sentence Questions	• Similar sentence stimuli to Sentence Repetition, but child answers questions (e.g., "The brown is purple. What is purple?" Answer: "The brown.") (ages 8–21)
Visual Digit Span	• Rote sequential memory using visual stimuli

understanding the relationship of its measures to brain function, reading competency, and even treatment response (Conant et al., 2013; Foorman et al., 2015; Hutton et al., 2020; Kovelman et al., 2012; Lonigan et al., 2009; Leitão & Fletcher, 2004; Marshall et al., 2013; McNorgan et al., 2013; O'Brien et al., 2013; Park & Lombardino, 2013; Pollitt & Harrison, 2021; Pugh et al., 2013; Saygin et al., 2013; Toste et al., 2020).

Delis–Kaplan Executive Function System

No one has inspired and transformed the field of neuropsychology like Edith Kaplan, who could be considered the founder of the neuropsychological process approach (Oscar-Berman & Fein, 2013). Kaplan attempted to bring process assessment into the mainstream by developing

TABLE 4.5. TABLE 4.5. Characteristics of Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2), Subtests

Subtest	Constructs purportedly tapped
Phonological Awareness	
Elision	Phonological perception, segmentation, individual phonemes
Blending Words	Phonological assembly
Sound Matching	Phonological perception, segmentation, individual phonemes
Phoneme Isolation	Phonological perception, segmentation, individual phonemes
Blending Nonwords	Phonological assembly
Segmenting Nonwords	Phonological perception, segmentation, individual phonemes
Phonological Memory	
Memory for Digits	Rote auditory memory
Nonword Repetition	Phonemic analysis, assembly, auditory working memory
Rapid Naming	
Rapid Color Naming	Naming automaticity, processing speed, speed of lexical access, verbal fluency
Rapid Object Naming	Object recognition, naming automaticity, processing speed, speed of lexical access, verbal fluency
Rapid Digit Naming	Number automaticity, processing speed, speed of lexical access, verbal fluency
Rapid Letter Naming	Letter automaticity, processing speed, speed of lexical access, verbal fluency

measures that were both clinically useful and had psychometric integrity (Libon et al., 2013), which is not an easy feat. She will be terribly missed but her legacy and contribution to the field of neuropsychology continues. Her trainees carry on her tradition which is very similar to Luria's approach to clinical assessment and testing of limits to understand brain structure, function, and implications.

One of her legacies is the Delis–Kaplan Executive Function System (D-KEFS; Delis et al., 2001). While it is a prominent measure of executive functions, mediated primarily by the prefrontal cortex, largely the dorsal system that deals with planning, organization, strategizing, monitoring, evaluating, and shifting behavior (Hale et al., 2013), the norms are becoming outdated. A revised version, the D-KEFS Advanced, which will be all-digital, is expected to be available in 2025. The original test was developed and normed on a large representative national sample to assess ages 8–89. Unlike many neuropsychological measures, the D-KEFS has extensive information about technical quality presented in the manual, which facilitates interpretation. Any of the specific tests can be administered separately, making it ideal for use in CHT. Of particular interest is the trail-making task, which allows the examiner to parse out sequencing, executive, and motor demands. It is also the only tool that has a tower task normed with other executive measures. Many of the tasks are versions of tasks with rich histories in neuropsychological assessment, and research supports the validity of the measures (Delis et al., 2004). A great deal of research has explored the psychometric characteristics (e.g., Sevadjian et al., 2011; Fine et al., 2011; Latzman & Markon, 2010), with evidence of both concurrent brain function and clinical utility examined (e.g., Berninger et al., 2008a, 2008b; Corbett et al., 2009; Figueras et al., 2008; Latzman et al., 2010; Lin et al., 2012; Poretti et al., 2013; Strong et al., 2010; Vasilopoulos et al., 2012). Table 4.6 describes the individual D-KEFS tests and the constructs purportedly assessed by each.

NEPSY—Second Edition

The NEPSY—Second Edition (NEPSY-II; Korkman, 2007) is the updated edition of the first *truly* developmental neuropsychological measure designed for children. This latest edition is for children ages 3–16. There are 32 subtests designed to provide a comprehensive evaluation of six functional domains: Attention and Executive Functioning, Language, Memory and Learning, Sensorimotor Functions, Social Perception, and Visuospatial Processing. However, unlike its predecessor, the NEPSY-II does not give factor or index scores, so it is ideal for use in hypothesis testing, and this was the reason the test publisher did not provide a method for calculating global scores. The NEPSY-II subtests and flexible administration format are primarily based on Luria's (1973b, 1980a, 1980b) model. This allows the examiner to pick and choose a subtest to test a specific hypothesis. Since this process requires the user to have enough understanding of brain–behavior relationships to recognize the processes tapped, some beginning practitioners might not consider the NEPSY-II, but we find some measures incredibly useful in CHT.

Although the test is based on a Lurian approach, the test does not break tasks down into primary, secondary, or tertiary skills; nor does the manual readily identify the relationships between subtest performance and the first, second, and third functional units. With many years in development, the NEPSY-II has all the advantages of being published by a major test developer, including an adequate normative sample, subtest technical quality, and ample validity studies. Not all of the NEPSY-II subtests show comparable technical quality, however, so Table 4.7 presents the subtests we have found to be most beneficial in CHT. In addition, though the Language subtests serve as a measure of *Gc*, the NEPSY-II does not adequately measure *Gf* or novel problem-solving skills. Both versions of the NEPSY have sufficient evidence of its use as a measure (e.g., Hayes & Watson, 2013; Horska & Barker, 2010; Schwartz et al., 2013).

TABLE 4.6. Characteristics of Delis–Kaplan Executive Function System (D-KEFS) Subtests

Subtest	Constructs purportedly tapped
Sorting Test	Problem solving, verbal and spatial concept formation, categorical thinking, flexibility of thinking on a conceptual task
Trail Making Test	Mental flexibility, sequential processing on a visual–motor task, set shifting
Verbal Fluency Test	Verbal fluency
Design Fluency Test	Visual fluency
Color–Word Interference Test	Attention and response inhibition
Tower Test	Planning, flexibility, organization, spatial reasoning, inhibition
20 Questions Test	Hypothesis testing, verbal and spatial abstract thinking, inhibition
Word Context Test	Deductive reasoning, verbal abstract thinking
Proverb Test	Metaphorical thinking, generating versus comprehending abstract thoughts

TABLE 4.7. Characteristics of NEPSY-II Subtests

Subtest	Constructs purportedly tapped
Attention/Executive Functions	
Auditory Attention and Response Set	Sustained auditory attention, vigilance, inhibition, set maintenance, mental flexibility
Design Fluency	Visual–motor fluency, mental flexibility, graphomotor responding in structured and unstructured situations
Animal Sorting	Ability to formulate basic concepts and to transfer those concepts into action
Clocks	Planning and organization and visual–perceptual and visual–spatial skills
Inhibition	Ability to inhibit automatic responses
Statue	Motor persistence and inhibition
Language	
Phonological Processing	Similar to WJ-IV *Ga* subtests (see Chapter 1, Table 1.6); auditory attention, phonological awareness, segmentation, assembly
Comprehension of Instructions	Receptive language, sequencing, grammar, simple motor response
Repetition of Nonsense Words	Auditory presentation of nonsense words; phonemic awareness, segmentation, assembly, sequencing, simple oral expression
Oromotor Sequences	Oromotor programming
Speeded Naming	Rapid semantic access
Word Generation	Verbal productivity
Memory and Learning	
List Memory	Remember list of unrelated words over multiple learning trials; one delayed trial after interference list
Memory for Designs	Visual-spatial memory; also requires maintenance of rules
Memory for Faces	Select previously viewed photo from an array
Memory for Names	Learn the names of line drawings of children's faces over multiple trials
Narrative Memory	Recall of orally presented narratives; recall of details and inferential comprehension
Sentence Repetition	Rote auditory recall; grammatical knowledge
Word List Interference	Rote repetition of unrelated words, with each set of two followed by recall of both sets; working memory
Social Perception	
Affect Recognition	Matching photos expressing the same feeling: happy, sad, fear, anger, disgust, neutral
Theory of Mind	Understand how others are feeling, understand false beliefs; also requires verbal comprehension and memory
Sensorimotor Functions	
Fingertip Tapping	Simple motor speed, perseverance
Imitating Hand Positions	Visual perception, memory, kinesthesis, praxis

(*continued*)

TABLE 4.7. (*continued*)

Visual–Motor Precision	Visual–motor integration, graphomotor coordination without constructional requirements
Manual Motor Sequences	Motor imitation
Visual–Spatial Processing	
Design Copying	Visual perception of abstract stimuli, visual–motor integration, graphomotor skills
Arrows	Spatial processing, visualization, line orientation, inhibition, no graphomotor demands
Block Construction	Similar to WISC-V Block Design (see Tables 1.5 and 4.1)
Geometric Puzzles	Mental rotation, visual–spatial analysis, and attention to detail
Picture Puzzles	Visual discrimination, spatial localization, spatial localization, and visual spanning
Route Finding	Visual-spatial relations and directionality

Process Assessment of the Learner—Second Edition: Test Battery for Reading and Writing and Test Battery for Mathematics

The artificial distinction between "ability" and "achievement" in SLD identification led us down a clinical path that ignored their interrelationship. To look in more detail at the processes involved in reading and writing, the Process Assessment of the Learner—Second Edition: Test Battery for Reading and Writing (PAL-II RW; Berninger, 2007b) is available to complement regular standardized achievement testing. Individual subtests can be administered and interpreted, making this test ideal for CHT. There are also intervention materials available for both individual and classroom implementation based on Berninger's extensive research findings. The PAL-II RW includes measures of phonological processing; orthographic coding; rapid automatized naming; phonological decoding; and integration of listening, note taking, and summary writing skills. Although the PAL-II RW is used for examining academic skills, it focuses on the psychological processes associated with these skills, making it especially useful for linking assessment to intervention.

The Process Assessment of the Learner—Second Edition: Test Battery for Mathematics (PAL-II Math; Berninger, 2007a) measures the development of cognitive processes that are related to math. The PAL-II Math includes measures of numeral writing, numeric coding, quantitative working memory, spatial working memory, rapid naming, and graphomotor integration. Like the PAL-II RW, the PAL-II Math's user guide contains resources and recommended interventions based on the assessment results.

Berninger is one of the most accomplished researchers in the field, and has used an extensive evidence base to establish the validity of the PAL-II (e.g., Abbott et al., 2010; Altemeier et al., 2008; Berninger et al., 2006; Berninger et al., 2010; Berninger & Dunn, 2012; Berninger et al., 2013; Berninger & O'Malley May, 2011; Berninger & Niedo, 2012; Berninger & Richards, 2010; Niedo et al., 2013; Brooks et al., 2011; Richards et al., 2009). Not only does the work cited here show her impressive work to highlight the validity of the PAL-II, but she also documents its utility with research in the field and laboratory, showing changes in brain functioning after intervention. Not only is the PAL-II a useful test for use in CHT, this depth of literature across disciplines is worth reading to help you understand the relationship between test data, brain functioning, classroom achievement, and brain-based treatment response.

Feifer Assessments of Reading, Mathematics, and Writing

Well known in the school neuropsychology field, Steven Feifer has developed a family of instruments designed for the neuropsychological assessment of processes underlying achievement. These relative newcomers are well designed from a acomprehensive view of the neuropsychology of learning. Unfortunately, the norms are grade-based, making direct comparison to other measures that are mostly age-based (as we recommend) difficult, but the excellent coverage of processing components makes them very useful for CHT. The Feifer Assessment of Reading (FAR; Feifer, 2015) has several subtests assessing each of phonological skills, fluency, and comprehension. The more unique and therefore most useful subtests for CHT assess visual perception, orthographic processing, semantic concepts, print knowledge, and morphological processing. The Feifer Assessment of Mathematics (FAM; Feifer, 2016) has sections assessing verbal retrieval and linguistic components of math, procedural knowledge, and semantic components. The semantic index, in particular, includes many useful subtests for CHT, including measures of number sense, magnitude representation, visual–spatial and conceptual components of math, and high-level problem solving. The Feifer Assessment of Writing (FAW; Feifer, 2020) assesses graphomotor skills with several subtests, two measures of spelling contributing to a dyslexia index, and an executive index. As you will see in later chapters, writing is one of the most cognitively complex tasks we undertake, and the executive component is considerable. The FAW subtests will be useful in CHT for students with writing difficulties.

Test of Memory and Learning—Second Edition

The Test of Memory and Learning—Second Edition (TOMAL-2; Reynolds & Voress, 2007) is in many ways a more comprehensive measure of learning and memory than the CMS. The current edition has been normed for children and adults ages 5–59. (Note that the TOMAL-3 is being normed as this book is being written.) The TOMAL consists of eight core and six supplemental subtests, and two delayed-recall subtests. It was carefully standardized, and the norms are representative of the U.S. Census population. Reliabilities tend to be quite strong across ages, especially for the composite scores. However, further support for its use in memory assessment can been found in subsequent studies reported in the literature. Table 4.8 provides an overview of the TOMAL-2 subtests we find useful in CHT. The Delayed Recall Index includes delayed recall from the Memory for Stories, Word Selective Reminding, Facial Memory, and Visual Selective Reminding subtests. As with the CMS, one of the difficulties with the TOMAL-2 is its breakdown into verbal and nonverbal memory domains. It is not surprising that the second edition has strong technical quality and evidence of clinical utility (e.g., Brooks & Iverson, 2012; Dehn, 2010; Fuentes et al., 2012; Lajiness-O'Neill et al., 2010; Sutton et al., 2011; Meekes et al., 2013; Riegler et al., 2013; Thaler et al., 2012; Thaler et al., 2010; Till et al., 2013).

Wide Range Assessment of Memory and Learning—Third Edition

The Wide Range Assessment of Memory and Learning—Third Edition (WRAML-3; Adams & Sheslow, 2021) was the first child memory scale on the market, having been first developed in the 1980s. Like the other measures reviewed here, it examines verbal and visual memory, and includes an attention/concentration index score. Additional examination of delayed recall, working memory for children nine and older, and recognition are possible. For verbal memory, rote, sentence, and story memory are tapped. For visual memory, both abstract and meaningful memory are assessed. These tasks are challenging yet interesting for children,

TABLE 4.8. Characteristics of Test of Memory and Learning, Second Edition (TOMAL-2), Subtests

Subtest	Constructs purportedly tapped
Verbal Memory Index	
Memory for Stories	See CMS Stories (Table 4.3 lists this and other CMS subtests)
Word Selective Reminding	Similar to CMS Word Lists, but no interference task
Paired Recall	See CMS Word Pairs
Digits Forward	Auditory rote memory, sequential recall, attention
Digits Backward	Similar to WISC-V/WJ-IV versions; more demands on attention, working memory, executive functions than Digits Forward
Letters Forward	Auditory rote memory, sequential recall, attention
Letters Backward	Working memory, attention, executive functions
Nonverbal Memory Index	
Facial Memory	See CMS Faces; good ventral stream measure
Visual Selective Reminding	Visual analogue to word selective reminding, with dots; dorsal stream, visual–motor coordination, praxis without visual discrimination
Abstract Visual Memory	Visual discrimination of abstract symbols, recognition memory
Visual–Sequential Memory	Visual discrimination of abstract symbols, sequencing, praxis
Memory for Location	See CMS Dot Locations; good dorsal stream measure
Manual Imitation	Short-term visual–sequential memory, praxis
Object Recall	Visual and verbal presentation of objects with verbal recall over multiple trials.

making the WRAML-3 a possible alternative to the CMS and TOMAL-2. It is fairly easy to administer and score. It has a large normative sample and adequate technical characteristics. The WRAML-3 is a popular test and well-liked by children taking it. The Third Edition is much improved in terms of its psychometric characteristics and clinical utility was established in the second edition (e.g., Atkinson et al., 2008; Burton et al., 2012; Goldstein et al., 2014; Lajiness-O'Neill et al., 2013; McKnight & Culotta, 2012). Given this is a new measure, it will be important to see how well it fares in independent research.

Supplemental Neuropsychological Measures for Hypothesis Testing

Table 4.9 presents a number of other neuropsychological measures we have found useful in CHT. Although some are specifically for use with children, others listed in this table have a long history of use in neuropsychological assessment of adults, and most have been adequately extended downward for use with children. These instruments measure a variety of cognitive or neuropsychological constructs, and many have been found to be sensitive to brain functions and dysfunctions. They can be used to test initial hypotheses or validate hypotheses derived from previously discussed measures. Some measures, such as the Rey–Osterreith Complex Figure (a visual–spatial–graphomotor task) and the California Verbal Learning Test—children's version (a language task), could be listed under other table subheadings. We have put the measures in the domains that are most likely to serve our CHT purposes.

TABLE 4.9. Supplemental Measures for Hypothesis Testing

Subtest	Constructs purportedly tapped
Attention Memory/Executive Function	
Children's Category Test (Boll, 1993)	See Halstead–Reitan Category Test (Table 4.2)
Wisconsin Card Sorting Test (Heaton et al., 1993)	Executive functions, problem solving, set maintenance, goal-oriented behavior, inhibition, ability to benefit from feedback, mental flexibility, perseveration
Tower of London (Shallice, 1982)	Planning, inhibition, problem solving, monitoring, and self-regulation
Stroop Color–Word Test (Golden et al., 2002)	See CAS-2 Expressive Attention (Table 4.4)
Rey–Osterrieth Complex Figure (Meyers & Meyers, 1995)	Visual–motor integration, constructional skills, graphomotor skills, visual memory, planning, organization, problem solving
Conners Continuous Performance Test 3 (CPT-3; Conners, Sitarenios, & Ayearst, 2018)	Computerized measure of sustained attention, impulse control, reaction time, persistence, response variability, perseveration, visual discrimination
Hale-Denkla Cancellation Task (Hale et al., 2009)	Attention, concentration, visual scanning
California Verbal Learning Test—Children's Version (Delis et al., 1994)	Verbal learning, long-term memory encoding and retrieval, susceptibility to interference
Comprehensive Trail-Making Test—Second Edition (CTMT2; Reynolds, 2020)	Attention, concentration, resistance to distraction, cognitive flexibility/set shifting
Behavior Rating Inventory of Executive Function, Second Edition (BRIEF2; Gioia et al., 2015)	Parent and teacher rating scales of behavioral regulation, metacognition; includes clinical scales assessing inhibition, cognitive shift, emotional control, task initiation, working memory, planning, organization of materials, and self-monitoring; includes validity scales assessing inconsistent responding and negativity
Tests of Variable Attention (TOVA; Leark et al., 2008)	Computerized measure of sustained and selective attention
Sensory–motor/Nonverbal skills	
Developmental Test of Visual–Motor Integration, Sixth Edition (Beery & Beery, 2010)	Visual–perceptual skills, fine-motor skills, visual–motor integration
Purdue Pegboard (Tiffin & Asher, 1948)	Fine-motor skills, bimanual integration, psychomotor speed
Grooved Pegboard (Kløve, 1963)	Complex visual–motor–tactile integration, psychomotor speed (compare to simple sensory–motor integration)
Judgment of Line Orientation (Benton & Tranel, 1993)	See NEPSY-II Arrows (Table 4.7)
Language Measures	
Oral and Written Language Scales, Second Edition (Carrow-Woolfolk, 2011)	Listening comprehension, oral expression, written expression; not limited to single-word responses
Comprehensive Assessment of Spoken Language, Second Edition (Carrow-Woolfolk, 2017)	Language processing in comprehension, expression, and retrieval in these categories: lexical/semantic, syntactic, supralinguistic, pragmatic

(*continued*)

TABLE 4.9. (*continued*)

Clinical Evaluation of Language Fundamentals—Fifth Edition (CELF-5; Wiig, Semel, & Secord, 2013)	Assesses receptive and expressive language with the core subtests, but also allows assessment of language structure, language content, and memory; includes standardized observations in the classroom and assessment of pragmatic language skills, in addition to individual assessment
Test of Language Development—TOLD-5, Primary and Intermediate; Newcomer & Hammill, 2019)	Primary version assesses phonology, semantics, and syntax; Intermediate version assesses semantics and syntax
Receptive auditory/Verbal skills	
Wepman Auditory Discrimination Test—Second Edition (Wepman & Reynolds, 1987)	Auditory attention, phonemic awareness, phonemic segmentation, phoneme position (primary/medial/recent)
Peabody Picture Vocabulary Test, Fifth Edition (PPVT-5; Dunn & Dunn, 2018)	Receptive vocabulary (visual scanning/impulse control); conormed with EVT-3 (see below)
Token Test for Children, Second Edition (TTFC-2; McGhee et al., 2007)	Receptive language, auditory working memory, direction following without significant cultural content
Expressive auditory/Verbal skills	
Controlled Oral Word Association Test (Spreen & Benton, 1977)	See NEPSY-II Verbal Fluency (Table 4.7)
Boston Naming Test (Goodglass & Kaplan, 1987)	Expressive vocabulary, free-recall retrieval from long-term memory versus cued-recall retrieval (semantic/phonemic)
Expressive Vocabulary Test, Third Edition (EVT-3; Williams, 2018)	Expressive vocabulary (picture naming); conormed with PPVT-5 (see above)

Behavioral Neuropsychology and Problem-Solving Consultation

Utilizing Assessment and Consultation Skills

Now that we have reviewed the assessment part of our model, let's integrate it with consultation. Notice the heading above. Isn't *behavioral neuropsychology* an oxymoron? No, because we believe that these two orientations should be seen as integrated, not as antithetical. In the past, consultation was seen as something a school psychologist would do before a comprehensive evaluation, or instead of a comprehensive evaluation. In contrast, we see consultation as an integral part of everything you do in all the tiers! Data collection is important in consultation too, and the fact that you are doing standardized assessments doesn't mean you can't also do problem-solving consultation. Data collection is important for understanding and serving children, but how you use it is probably even more important for your practice of neuropsychology. We are suggesting that these two functions of school psychologists can be combined to make both stronger. You can bring assessment data into the consultation data-gathering phase when this is appropriate, linking interventions to the child's strengths and needs. The CHT emphasis on *ecological validity* and *treatment validity* is what sets our model apart from other test interpretation models, which have largely focused on testing for identification or diagnostic purposes. As Miciak and Fletcher (2020) recommend, assessment data should lead to intervention, a point we wholeheartedly agree with—your testing is not only about *understanding*, but it is also about *doing*—linking assessment data to intervention to improve outcomes for all children is best practice.

Most referrals for problem-solving consultation concern academic problems, and most of those academic problems are reading difficulties. Although general consultation on reading instruction may be helpful, combining this knowledge with information about the multiple determinants of the child's problem can have important effects on the intervention you and the teacher choose, and on the success the child experiences as a result of your efforts. Consultation is intended to be a collaboration between equals, but the fact that the consultant is there to help the consultee solve a problem has the potential to make the power relationship unequal. We have to guard against a traditional mental health consultation approach that does not work well in schools. There's a tendency for consultees to defer to the "expert" neuropsychologist consultant, agree with the consultant during meetings, but then not feel ownership of the interventions developed during consultation. The problem-solving collaborative consultation approach must be an equal partnership because without consultee ownership many of these interventions will not be fully implemented or implemented with integrity (Erchul & Martens, 2010). We believe that the power issues within the consultative relationship must be acknowledged and dealt with directly. Both school psychologists and teachers feel that expertise and informational power are essential in making changes with teachers (Owens et al., 2018). You are using your expertise and knowledge to help solve a problem, influence a teacher to make changes, and support and develop the teacher's skills (Erchul & Martens, 2010). By gaining knowledge and skills in instructional leadership, systems issues, collaborative team building, and academic and behavioral interventions, you can be an important source for the consultee, and guide him or her toward solutions without being coercive.

Consultation begins with the premise that the consultant works with the consultee (usually the classroom teacher) to solve a client's (the teacher's student's) problem. It is also assumed that both professionals have specific expertise to bring to bear on the problem. In our view, your knowledge of neuropsychological and cognitive functions, neuropsychological assessment, the academic and behavioral interventions literature, and intervention-monitoring methodology should be the core of expertise that you bring to the relationship. The teacher's knowledge of the student's classroom performance, awareness of effective and ineffective teaching techniques for the child, and professional experience as a teacher form the core of his or her expertise as the consultee. Fully acknowledging the expertise of the consultee is one part of building rapport. This consultee's expertise is also necessary if an appropriate problem solution is to be found. An intervention plan that takes into account available resources and the interventions the teacher is already trying should have greater applicability and effectiveness (Miller et al., 2019).

The following problem-solving consultation model is a summary of content and skills presented by several experts combined with our CHT model (Erchul & Martens, 2010; Brown-Chidsey & Andren, 2012). This model is especially relevant in the Tier 2 approach (problem-solving RTI approach) in an attempt to serve children before comprehensive evaluation is ever considered, and in Tier 3 following comprehensive CHT evaluation. The only difference is that in Tier 3 the cognitive and neuropsychological processing characteristics are also considered in problem analysis.

Stages of Problem-Solving Consultation

Problem Identification

During the initial interview, the consultant (you) and the consultee identify a target behavior for intervention. The behavior must be defined in an observable, measurable way. It is typical for teachers and parents to report summative judgments ("He's depressed") rather than an operational behavior that is related to that summative judgment (slow in completing work, does not interact with peers). Information is needed about how often and when the behavior occurs, so information about frequency, duration, and severity need to be considered. To do this well, you often need to consider multiple behaviors before choosing the target behavior. It is important to list and operationalize all behaviors, and then hierarchically order them from most to least significant in the eyes of both the consultee and consultant. Many models suggest picking the most significant problem, but in reality this may not be feasible or appropriate. It is sometimes more important to pick a target behavior that is readily changeable so both the child and consultee experience success. It does no good to pick a significant problem that is so routinized that change is unlikely or very difficult, thus discouraging all those involved in the process. Instead, you should consider task-analyzing (breaking the problem behavior into subcomponent parts), and then using a shaping process (reinforcing successive approximations of target behaviors) to encourage success before taking on the more significant problem. In addition, consultation and behavioral research have shown us that if we just suppress a behavior it will manifest in another way, so we also need a positive replacement behavior we want to increase. The main idea here is you want to punish (extinguish) problem behaviors, and reinforce positive replacement behaviors. To do this, we need to think about strategies such as differential reinforcement of alternative (DRA), differential reinforcement of incompatible (DRI), or differential reinforcement of other (DRO) behaviors. As a reminder, DRA involves decreasing an unwanted behavior while at the same time reinforcing an alternative, more acceptable, behavior. DRI is when a behavior that is incompatible with the target behavior is reinforced. For example, a child who pulls out her hair is instead engaged in squeezing a stress ball, which is reinforced. This gives the child a distraction which requires her to use her hands for something else besides hair pulling. DRO is simply reinforcing the behavior that you wish to shape, while ignoring or extinguishing the problem behavior.

CHT problem identification is somewhat more complex than in other problem-solving models. CHT includes data collected from MTSS interventions, observations, interviews, and cognitive, neuropsychological, academic, and behavior assessment results. As noted previously, it is important that the teacher be consulted to determine that your findings have concurrent and ecological validity.

Problem Analysis

Even if you have conducted a CHT evaluation, a more in-depth study of the target behavior is made during problem analysis, possibly including a functional behavioral assessment (FBA) and/or curriculum-based measurement (CBM). A handbook for conducting a functional assessment in schools, such as Steege et al. (2019) will be helpful at this stage. For assistance with academics, the work of Hosp et al. (2016) is a useful resource. An FBA should include a review of the prior intervention data collected, and an interview with the teacher to identify possible causes, antecedents, and consequences of the behavior. Most functional assessments focus on obvious causes for the behavior, such as stimuli that precipitate problem behavior, or consequent events such seeking attention or escaping from a task. Although discussing functional behavior analysis would seem counterintuitive in a school neuropsychology book, this is hardly the case since the environment plays an important role, not only in determining the antecedent and consequent events associated with a given behavior, but also in how these interact with brain function.

The CHT process will provide information about the student's cognitive processing strengths and weaknesses to use in developing working hypotheses, such as processing difficulties, memory problems, language deficits, or difficulty with unstructured situations. As part of the problem analysis, a review of interventions that have already been attempted and their effectiveness is also helpful. Although CHT includes functional analysis in this stage, it relies on much more information from numerous data sources and integrates these sources with our understanding of the child's individual neuropsychological strengths and weaknesses.

While considering an FBA for academic problems is potentially useful if the problem is a performance deficit instead of a skill deficit, in many cases a skill problem is important to identify. If the child does not have the skill to complete the task, then that skill should be taught first. Both survey and specific level assessments can be useful in determining the level and pattern of academic skills deficits, especially using the validated CBM tools discussed later in this chapter. Your CHT evaluation may have adequately linked the processing deficits associated with the achievement problem, but the CBM can give you a more fine-grained analysis of the problem. What CHT acknowledges, which is absent in the either/or perspective regarding skill or performance deficit, is that an executive problem (a skill deficit) can look like a performance deficit, because the child does the skill sometimes and not others, due to poor executive control. A more detailed explanation of academic error patterns can be found in Chapter 5 (reading), Chapter 6 (mathematics), and Chapter 7 (written expression). Recall that error analysis is the key to understanding academic skill deficits, and when linked with your understanding of brain function, you can design individualized interventions for children, which is the next problem-solving step.

Plan Development/Implementation

Following the problem analysis step, a working hypothesis needs to be developed conjointly by the consultant and the consultee as a basis for the intervention. In our experience, this can be a very challenging stage for many beginning consultants and even advanced consultees. School psychologists may be better versed in interventions such as cognitive-behavioral therapy (e.g., Kendall, 2011) or social skills interventions (Gresham, 2016) than academic interventions. One of the things we recommend is that every school psychologist buy special education books for teachers, such as *Students with Learning Problems* (Mercer & Pullen, 2010) or *Positive Behavioral Supports for the Classroom* (Scheuermann, Hall, & Billingsly, 2011). These books are useful for two reasons; they give you a good idea of what teachers are exposed to in their training, and they also give you a lot of great ideas for evidence-based interventions to pull out during the consultation process, and subsequently to address the child's needs.

During this phase, the consultant and consultee recognize there are multiple factors associated with intervention development and implementation that require careful consideration. This process begins with a good operational definition and specification of the goal, which includes the behavior objective (also written in problem identification), content, materials, conditions in which the behavior will occur, and the criterion for acceptable performance. The plan takes into account the student's characteristics and behavior, the classroom ecology, the resources available, and the teacher's style and preferences. Working together, the consultant and consultee brainstorm many possible interventions, and then choose the intervention that is likely to be effective and can be plausibly implemented. Based on the child, consider instructional or intervention strategies and modifications, evaluate resources for implementation, and consider reinforcers or consequences to increase the student's behavior or performance. For instance, some reinforcers like social interaction may be best for some students, but others may prefer video game time.

Be sure to write these intervention plans with extensive behavioral detail, and data collection methods. Whenever possible, using extant rating scales or academic probes for data collection is a good choice, but on many occasions you may have to develop your own. Finally, determine how long it might take for the child to show significant improvement. The length of intervention in relation to the severity of the problem is a difficult consideration. If the problem is significant, and the intervention period is too short, the intervention will not be successful. If the problem is mild and a long intervention period is considered, the child and teacher can become frustrated as improvement is so gradual and the intervention may be terminated prematurely. Feasibility is a key variable here. In our experience teachers get excited about the intervention effort, and sometimes pick strategies or measurement systems that may not be feasible. For instance, teachers may want to do a frequency count of a target behavior by marking every time it occurs on a sheet during the class, but this often is not feasible. Instead using an intermittent count may be more helpful.

Finally, consider how to ensure the intervention is implemented with integrity and how you will evaluate it. Even though the consultant attempts to work with the teacher to develop the plan, it may need recycling because any of the details just noted might not have been feasible. What is the best intervention for children with academic or behavioral needs? The answer is simple: The best intervention is the one that works.

Plan Evaluation/Recycling

The key to successful interventions is ensuring fidelity of interventions and treatment integrity (Harn et al., 2017). Several times during the intervention it is important for the consultant to "check in" with the consultee to see how the intervention is progressing and review the data collected. These repeated informal connections help identify if more formal consultation meetings are needed, and they ensure the interventions are implemented with integrity. Periodic checks and data collection/graphing can help the consultant and consultee evaluate intervention effectiveness (i.e., data collection/analysis and data interpretation) and determine when changes are needed. There are numerous methods for evaluating interventions via within-subject experimental designs, several of which we will review later in the chapter. If the intervention is successful, either the intervention is extended, or it is discontinued if the target has been reached.

If minor revisions appear necessary, the consultee makes them at this time, and he or she decides on an additional meeting to evaluate the revised intervention. If different or more intensive interventions appear necessary (i.e., a new working hypothesis), a new intervention can be attempted, or additional special education support services may be needed. This evaluation process is also important as the instructional supports begin to be removed and the child begins to function completely within his or her natural environment with natural consequences. In most instances, however, there are minor and on occasion even major revisions during recycling. Recall it is the consultee who takes ownership of the intervention, and so the consultee may need to recycle the intervention several times before we get a good response. The theory–hypothesis–data collection/analysis–data interpretation cycle continues until the problem appears to be under natural stimulus–consequence control. As you can see, the CHT model is not really about testing per se; it is about a way of practice that combines the best techniques of problem-solving consultation with comprehensive evaluations and multiple data sources.

Practicing Behavioral Neuropsychology

Since we are suggesting that you combine neuropsychological assessment with behavioral methods, In-Depth 4.1 and Table 4.10 review the basics of behavioral interventions for those readers who may not recall the details. As part of the problem-solving model, you need to recognize that antecedent and consequent actions affect the child's learning and behavior

(Crone et al., 2015), and also that cognitive and neuropsychological processes interact with these determinants. Having this understanding allows you to use what cognitive psychologists have called *stimulus–organism–response* (S-O-R) psychology, in which stimulus and response are still important, but the organismic variables (i.e., child neuropsychological processes) help you determine what the best intervention is and how to carry it out. This is not unlike Bandura's (1978) reciprocal determinism model, where one considers the behavior, cognitions, and environment in both understanding the child and in developing interventions that meet the child's needs. The behavior techniques become especially useful in designing the intervention, determining intervention efficacy, and managing contingencies.

IN-DEPTH 4.1. Review of Behavioral Psychology Principles

RESPONDENT CONDITIONING TECHNIQUES

Respondent conditioning is a method of eliciting behavior by manipulating a stimulus. An example of a conditioned stimulus is the teacher's turning on and off the light to cue a child's transition behavior. Behavioral examples might include anxiety about tests or speaking in class, or fear when the teacher raises his or her voice. Common interventions, including relaxation training and systematic desensitization, may be used to treat anxiety responses in students. However, more broadly conceived, variations in stimuli can lead to different behaviors (e.g., varying spacing or size of letters during reading, using simultaneous visual and auditory teacher instructions, using an adapted pencil for sensory problems for writing, tapping on a desk to cue on-task behavior, etc.). Modeling and discriminative stimuli designed to elicit operant behaviors, though not considered respondent techniques, can both be related to stimulus–response psychology.

OPERANT CONDITIONING TECHNIQUES

Operant conditioning is a method of affecting behavior by manipulating the consequences of that behavior. Behaviors that are followed by reinforcing consequences (either presentation of something positive or removal of something negative) will tend to recur. Behaviors that are followed by punishing consequences (either presentation of something negative or removal of something positive) will be less likely to recur, as indicated in Table 4.10. One of the best uses of operant technology is the "Premack principle," in which a less reinforcing behavior is reinforced by a more reinforcing one (e.g., providing computer time after a certain level of reading accuracy is obtained). Positive reinforcement can include natural consequences (these are preferable) or secondary ones (e.g., tokens, points). A good use of negative reinforcement is reducing the workload if a child demonstrates mastery on an assignment.

People are often confused about the difference between *positive reinforcement* (presenting something positive) and *negative reinforcement* (removing something negative). Why do children have tantrums? Not only because they are positively reinforced for having tantrums, but also because their parents are negatively reinforced—they get peace and quiet by giving in to the children. Most interventions in school should use *positive reinforcers*, and these can even be used to teach children *not* to do something, so (we hope) you don't have to use punishment. You identify an alternative behavior, preferably one that is incompatible with the negative behavior, and reinforce that behavior (i.e., differential reinforcement of other, alternative, or incompatible behavior). For example, Taniqua is always running in the halls. Instead of punishing her for running, reinforce her for walking. In some cases, a child may not be able to do the target behavior. In these situations, reinforcing successive approximations of target behaviors, or "shaping," is what we have to do with academic and behavioral deficits.

Is there a place for punishment in the schools? If a child is always being punished at school, it becomes aversive, something to avoid; it may even eventually lead him or her to drop out. A particular teacher who, or a subject that, is punishing may also be seen as aversive. There is another problem with punishment, though: The child isn't actually learning a replacement behavior. We prefer to use school interventions to teach children how to do something, rather than just to suppress negative behavior. If you must use punishment, we recommend that you use negative punishment that involves taking away something positive (either *time out from reinforcement* or *response cost*) combined with differential reinforcement. For example, if Kyle is aggressive on the playground, you can use negative punishment by having him sit on the sidelines and miss 5 minutes of recess, but you must also use positive reinforcement when you see Kyle playing nicely.

As you will recall from training, the schedules of reinforcement influence how a skill will be learned and maintained. Continuous reinforcement is good for skill acquisition, but this acquired skill will also be extinguished quickly, so intermittent reinforcement on a variable-ratio or interval scale is more appropriate. Think about slot machines; infrequent payoffs can maintain betting behavior for a long time. The same thing can happen in a classroom. If a teacher slips and accidentally reinforces an unwanted behavior, that behavior will be maintained longer.

Developing and Evaluating Interventions

After cycling through the first eight steps of CHT evaluation and refining a theory as to what will help the child, the next step is to utilize behavioral strategies that are combined with specific empirically supported instructional methods to help the child learn—through either remediation, accommodations, or both. In Chapters 5–7, we offer a number of interventions for academic skills problems. Some problems transcend academic domain boundaries, and the comorbidity among academic learning and other disorders is quite high. To help you understand the relationship between neuropsychological functioning and academic domains, we have provided a worksheet in Appendix 4.5. This worksheet may be useful in your examination of the academic issues associated with a child's neuropsychological functioning. This ensures that when you identify the cognitive pattern of performance, you are relating it to the academic pattern of performance seen on testing and in the classroom, which should help guide intervention planning and implementation. Taking what you know about the child's current level and pattern of performance, academic interventions, problem-solving consultation, and behavioral techniques, you can design, implement, and evaluate an intervention for him or her.

In CHT, we recommend using single-subject (within-subject or single-case) research

TABLE 4.10. Reinforcement and Punishment

	Provide	Remove
Positive consequence	Positive reinforcement	Negative punishment (response cost)
Negative consequence	Positive punishment	Negative reinforcement

Note. Shaded boxes *increase* the preceding behavior; unshaded boxes *decrease* the behavior.

designs to evaluate the effectiveness of interventions. We believe that practitioners should collect child performance data on a regular basis to ensure that interventions are effective. In this way, progress monitoring in the MTSS model permeates all parts of the balanced practice model and CHT process, so you are truly a data-based problem solver (Burns & Gibbons, 2013). We recommend that similar models be used to evaluate any intervention, whether it is behavioral, academic, cognitive, or socioemotional. These interventions must be individualized based on the teacher, student, and classroom.

In the next section, we review the most useful designs for evaluating school-based interventions, illustrating each intervention model with examples. Keep in mind that we are presenting the ideal research designs to demonstrate the effectiveness of an intervention scientifically; in real life, you may have to modify these designs to be more acceptable and less cumbersome, even if doing this provides less experimental control. Graphing behaviors is an excellent way of demonstrating progress to teachers and students, though, so we do encourage you to use them in progress monitoring. Nothing speaks louder than data in our opinion, and since psychologists are often the data management specialists in schools, this is a good role for you to take on. These visual illustrations show teachers, parents, and administrators the fruits of your intervention efforts. In addition, since graphing can be time consuming, some children can learn to graph their own behavior, and research supports their accuracy in doing so.

Research Designs for Evaluating Interventions

All of the research designs we discuss require two basic concepts. One is that you must have some way of *measuring the outcome* you want. Behaviorists generally call this "taking data," but you can think of it as "progress monitoring" or "checking up on the intervention." You can't simply say, "Jimmy's doing better"; you must have some way to *show* that the child is doing better. The outcome measure you choose depends on the target behavior and the goal of the intervention. Data might include information that the teacher already collects (i.e., authentic data—homework completed, spelling test score, office referrals or detentions, absences, etc.). You might collect information as part of the intervention itself (e.g., math worksheets, CBM probes of reading fluency, flashcards placed in correct and incorrect piles). Students can also collect and chart their own data, which reduces the load on the teacher, as described next.

During consultation, you and the teacher can also develop a data collection plan that interferes very little with the teacher's routine (e.g., child self-monitoring, using a wrist counter, completing end-of-the-period or end-of-the-day checklists). If it is too demanding, the likelihood of successful implementation is unlikely. For behavioral interventions, rating scales can be useful, but in some cases systematic observations can be used to evaluate progress by observing the target behavior directly, using event, duration, latency, and partial- or whole-interval recording. With observational data collection, it is important to use a randomly selected peer at baseline to establish a discrepancy with the target child. Finally, scatterplots are effective for teachers if they are recording behaviors (e.g., recording a + or – for a class period) themselves. Finally, students can take their own data, such as marking a checkmark on a sheet of paper, usually for either occurrence or nonoccurrence of the behavior. Table 4.11 presents some suggestions for outcome measures that can be useful in the classroom.

The second basic concept is that you must have a *baseline* measurement, in addition to measuring the behavior during the intervention. Teachers are generally familiar with just measuring the outcome of teaching, such as giving a test at the end of a chapter. In order to

TABLE 4.11. Examples of Outcome Measures for School-Based Interventions

Outcome area	Possible measures
Several behaviors	• Pre- and post-ratings on a brief behavior rating form. • Daily report card with ratings for day. • Systematic observation using event, duration, latency, partial- or whole-interval recording.
Negative classroom behavior (e.g., calling out, getting out of seat, yelling, aggression)	• Measurement of rate via tally marks, golf wrist counter, or pennies/paper clips transferred from pockets. • Student self-monitoring of behavior on sheet or card.
Serious negative behavior	• Count of office referrals or detentions.
Positive classroom behavior (e.g., raising hand, giving correct answers)	• Measurement of rate or student self-monitoring as above. • Observational data as above.
Attention, on-task behavior	• Periodic classroom observations. • Child self-monitoring of skills.
Academic work completion	• Worksheets or other permanent products. • Measurement of accuracy, rate, or both.
Homework completion	• Completed homework. • Daily report card signed by parent and/or teacher.
Academic skills accuracy	• Correct–incorrect flashcards kept in separate piles by student or peer. • Worksheets graded in percentages correct and recorded in grade book.
Academic skills fluency (speed and accuracy)	• Progress monitoring probes (e.g., DIBELS, AimsWEB).
Academic skills comprehension	• Pre- and posttest with alternate forms.

evaluate the effectiveness of an intervention, you have to measure the child's performance at the start (without the interventions), and then keep measuring as you implement the intervention to see how the child's performance changes. Without having a baseline for comparisons, you won't know whether the child's improvement is really due to the intervention. In describing some of the intervention models that follow, we use the letter *A* to refer to the baseline condition. The other letters (i.e., *B*, *C*) represent whatever interventions you implement.

ABAB/ABAC Designs

The ABAB design is used when you have picked one intervention and you want to see if it works better than the baseline condition (i.e., better than what the teacher would normally do). It is also sometimes called a "reversal design," because you do the intervention, then reverse to baseline for a short while, then do the intervention again. While this is a good way to show that the intervention is really what is affecting the child's performance, and can be published in behavioral journals, it doesn't work well for a situation where your intervention actually teaches the child something new. For example, if you teach a child to break a word into syllables to sound it out, you can't "unteach" that for the reversal phase. It also is not appropriate to do a reversal phase if the behavior you are trying to reduce is harmful to the child or others. For instance, if you are using time-out to reduce hitting, it would be unethical to do a reversal phase. As a result, this design is best for situations where you want to change the *rate* at which a child does something that he or she already knows how to do. For an example of an ABAB design, please see Case Study 4.1 and Figure 4.2.

The ABAC design allows you to compare two different interventions to see whether they are different from the baseline, and to see which is better at changing the child's behavior. Similar to the ABAB design, you first collect baseline data, then implement the first intervention (B), then reverse to baseline, and finally implement the second intervention (C). For instance, after taking baseline data on multistep math addition item accuracy (A), you can determine whether a child is more accurate if he or she draws lines between columns (B), or follows a step-by-step algorithm sheet on how to complete the problems (C). Case Study 4.2 and Figure 4.3 provide an example of an ABAC design.

Multiple-Baseline Design

A multiple-baseline design is useful when you expect the child's learning to be cumulative, so you don't want to reverse success. This design can teach children to display target behaviors across settings, people, or behaviors. For instance, if staying on task is the target behavior, you first seek on-task behavior in one class, then another, and so forth. In this design, you

Case Study 4.1. Jared's Impulsive Calling Out

An 8-year-old boy diagnosed with ADHD, Jared, was described by his teacher as extremely impulsive. The behavior that she identified as most problematic was Jared's calling out in class. Systematic observation data suggested that the teacher typically accepted Jared's answer when he called out, but then she often reminded him to raise his hand the next time. After discussing the baseline data with the teacher, we decided that she would use a wrist counter to count whenever Jared called out during whole-group instruction.

Figure 4.2 presents the results for the ABAB intervention designed to reduce his inappropriate call-out behaviors. During the first week, the teacher collected the baseline data. She counted Jared's call-outs without doing anything different about them, and this information was charted. The next week, the teacher continued to count Jared's call-outs, but she ignored him immediately after each call-out, practicing negative punishment. She only acknowledged Jared if he raised his hand first and did not call out, which is differential reinforcement. Notice that, at first, Jared's call-outs increased. This is called an *extinction burst*—a very common finding when a previously rewarded activity is being ignored. After that, Jared's call-outs began to decline. The teacher then returned to baseline for a short time (accepting call-out answers and reminding him to raise his hand), and the call-outs became frequent again. After a few days of this, the intervention was reintroduced. As you look at Figure 4.2, you should notice a few things. Each phase is separated by lines and labeled, so the baseline and intervention phases are clear. Within the baseline phase, Jared was calling out very frequently; the average was about 20 times per day. During the first intervention phase, his call-outs increased at first and then began to decline. As soon as the reversal to baseline took place, they increased again to about 20 times per day. During the second (and final) intervention phase, call-outs declined to an average of only 8 times per day. You can clearly see that the intervention was what was affecting Jared's behavior (this is called *establishing functional control*), because every time the intervention was implemented, he changed his behavior.

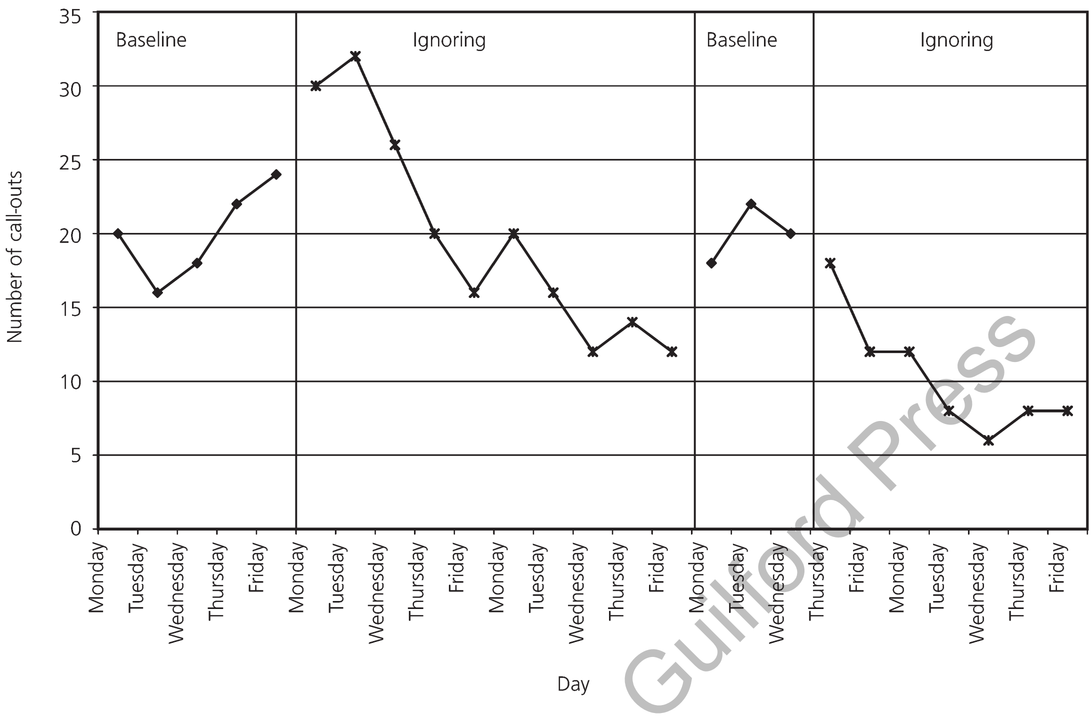

FIGURE 4.2. Jared's calling out.

Case Study 4.2. Increasing Mia's Reading Speed

Mia was a 9-year-old girl who was pleasant, cooperative, and hardworking. However, she was a slow, choppy reader, and her teacher sought support in helping Mia to read more fluently. Mia was in a small reading group with three other children, and the teacher worked individually with her for 15 minutes every day, but she was still struggling. The teacher now had an aide in class and wanted to know what the aide could do with Mia. Based on the CHT evaluation information, one of us (Hale) found that Mia had good phonemic awareness skills, and her phonemic segmentation and blending were not problems, but her word finding and rapid naming skills were quite poor. Hale met with the teacher, and they thought of two possible interventions for Mia: one where the aide would use flashcards to improve Mia's speed at identifying words, and one where the aide would read orally with Mia to increase the fluency of her reading. They decided that CBM of reading fluency, using daily 1-minute probes, would be a good outcome measure. As can be seen in Figure 4.3, her fluency was quite low at baseline (A). During the first intervention phase (B), the aide pronounced each word for Mia; Mia repeated it; Mia and the aide then practiced with the flashcards for about 10 minutes; and they finished with another 1-minute CBM probe. After this intervention, the teacher returned Mia to the baseline condition (A), but the aide continued to take CBM probes during this time. Finally, the second intervention phase was introduced (C). This intervention involved the aide's reading the passage to Mia one time with expression and fluency, and then their reading it together in tandem for about 10 minutes. Again, the sessions ended with another 1-minute CBM probe. As you can see from looking at Mia's chart, the flashcard drill improved her fluency over baseline, but the tandem reading was much more effective. This is not to say that tandem reading is a better intervention for all children; it just appeared to be better for Mia.

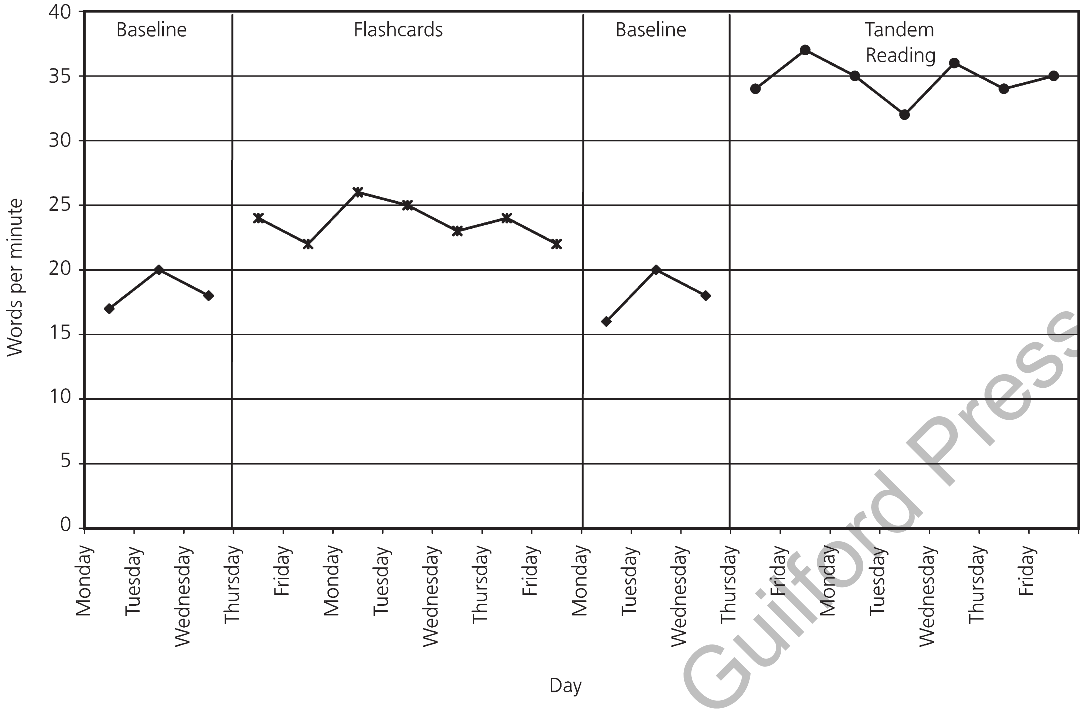

FIGURE 4.3. Mia's reading fluency.

collect baseline data in two or more subjects or at two or more times during the day. Then you start the intervention in one subject or at one time during the day, while continuing to take baseline data at the other time(s). Later, you introduce the intervention in the other subject or at the other time. If the child's performance changes in each setting only when the intervention is in place, you will know that the intervention is responsible for the change. An example of this design can be found in Case Study 4.3 and Figure 4.4.

Pre- and Posttest Design

A pre- and posttest design is useful when the teacher, student, or you can't collect data every day, but you want to measure the effectiveness of an intervention via direct observation, test, or rating scale. Although it is more difficult to establish functional control of the behavior at any given time, it is an easier method of data collection and is more likely to be acceptable to teachers. Keep in mind that if the intervention or data collection methods are too difficult or time consuming, they are unlikely to happen with integrity. Additionally, some data are better than no data, so a pre–post design may be optimal in some situations where resources are limited. For this design, it is important to choose a test (preferably one with alternate forms) or a rating scale that can be given repeatedly with minimal practice effects. There is increasing evidence that direct behavior ratings, including daily report cards, are a reliable and valid way to frequently assess behaviors in school (Miller et al., 2019). The pretest results become your baseline, and then you test again after implementation of the intervention to judge its effectiveness. Observations and brief rating scales can be used repeatedly if you choose to gather multiple data points during the intervention. Case Study 4.4 and Figure 4.5 provide an example of how to use a pre- and posttest design.

Case Study 4.3. Ellen's Accuracy Problem

Ellen was a 7-year-old girl who presented as a fast, careless worker. She reportedly completed her seatwork as fast as possible, without worrying about the accuracy of her responses. One of us (Fiorello) met with Ellen's teacher, and we decided to try to increase Ellen's accuracy by using rewards for correct responding. The teacher used Ellen's number correct on her seatwork papers to measure the outcome. She made sure that there were exactly 10 questions on each worksheet in math and spelling, and noted in her grade book the number correct for each day. For the first week, the teacher collected baseline data in both subjects for each day, and these data were charted on a multiple-baseline graph (see Figure 4.4). After collecting a week of baseline data, the teacher explained to Ellen that she could earn 1 point for each spelling word she copied correctly during seatwork, and the points could be traded for free time at the end of the morning classes. At the same time, Ellen's math work was kept in the baseline condition, with no rewards offered. As you can see from Ellen's chart, her spelling accuracy improved when rewards were added, but her math remained inaccurate. The next Monday, the teacher explained that the point system would apply to math as well, and as you can see from the figure, Ellen's accuracy in math improved thereafter.

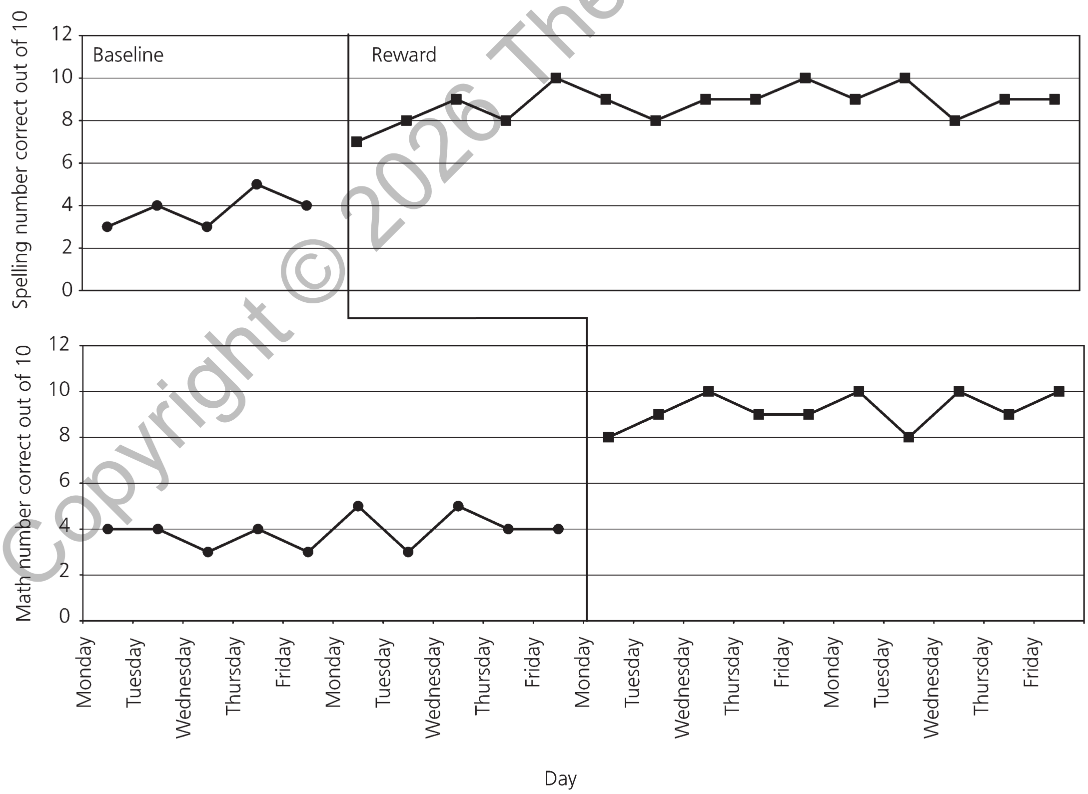

FIGURE 4.4. Ellen's accuracy.

Case Study 4.4. Herman's Auditory Processing

Herman was a boy with a common problem: a history of frequent ear infections (otitis media) and poor auditory processing. He was having difficulty learning the letter sounds in his kindergarten class. His teacher referred him to the reading specialist, who arranged for Herman to complete a 6-week computer-based auditory processing and phonics program. Before Herman began the program, one of us (Fiorello) was called in to develop a method for monitoring the efficacy of the program. We agreed that Fiorello would administer the CTOPP-2 and CBM of the alphabet sounds and would chart his scores, as depicted in Figure 4.5. After 6 weeks, Fiorello administered both tests again. Since the CTOPP-2 has age-based SSs, you can see that Herman's auditory processing improved over the course of the program. In addition, charting his improvement in letter sound knowledge helped the teacher compare Herman to other children, to guide her expectations for his curricular progress.

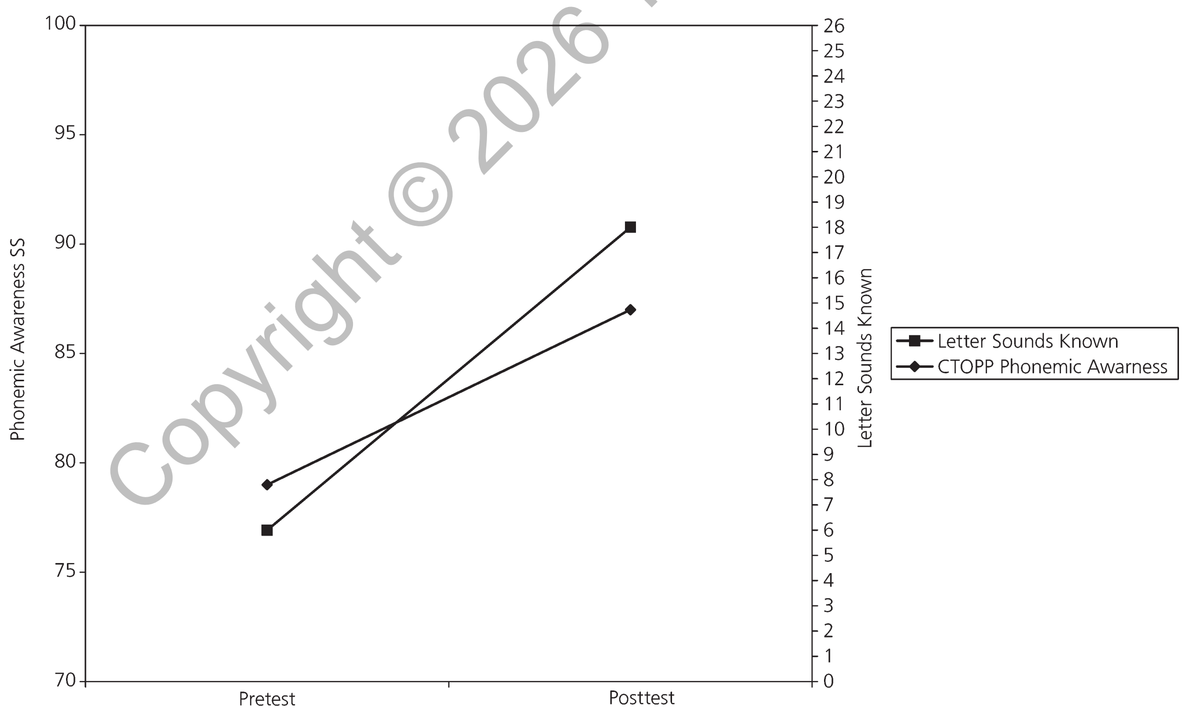

FIGURE 4.5. Herman's auditory processing and letter-sound knowledge.

CBM Progress Monitoring

CBM is useful for evaluating the effectiveness of instructional interventions on reading, mathematics, and writing. A brief probe is completed for several days during baseline, and then repeated every 1–2 days following the intervention session. To use this method, you have to determine the performance discrepancy (the child's functioning relative to peers), the goal for intervention (where you want the child to be after a period of time–usually where peers will be at that time), and the length of the intervention, all of which are somewhat subjective and dependent on a number of factors, including the severity of the child's problem and the cognitive functioning of the child.

CBM data are plotted to gauge progress over time, hence the name progress monitoring of performance. An *aimline* is drawn between the current functioning and the goal that has been set for the student. The beginning of the line is determined by the child's baseline performance or behavior; the end of the line is determined by where the child should be, compared to his or her peers, and how long it will take for the child to "catch up" once the intervention is in place. Unfortunately, there are no explicit guidelines for "how long it should take." For instance, if the child is 2 years behind, saying that he or she will make it up in a month is unrealistic, and would produce a very steep slope, ensuring in essence they would not respond to the intervention. Conversely, it is inappropriate to give a child too long to catch up, even though it might make them look like a responder throughout the process. After you establish an aimline, a *trendline* is drawn and recalculated regularly. The trendline shows the rate of improvement over time. If the trendline is below the aimline for several data points, the intervention should be adjusted or changed, or possibly you have set too high a goal for the child. Case Study 4.5 and Figure 4.6 highlight the use of CBM progress monitoring.

Case Study 4.5. Beverly's Limited Expressive Language

When one of us (Fiorello) was called in to consult with Beverly's teacher, Beverly was having considerable difficulty with expressive language, primarily because she spoke very little during conversations with her teacher and peers. CHT results revealed difficulty with word retrieval, oral fluency, and expressive syntax. Data collection with an audio recorder began, and Beverly's oral fluency at baseline was found to be only 23 words per minute on average (see Figure 4.6). Her teacher set a goal of 45 words per minute, and the teacher and Fiorello decided that a peer tutoring program would be implemented. The teacher picked a child who was not only friendly with Beverly, but also talkative, social, caring, and supportive. Each time the two children would get together, they would discuss a topic of interest. To facilitate this process, the teacher brainstormed possible topics with them before the intervention. As you can see, the peer tutoring improved Beverly's oral fluency at first, but on Days 10, 11, and 12, Beverly's fluency scores fell below the aimline. When three data points fall below the aimline, a decision point is reached. This means that it is time either to adjust or change the intervention, or to readjust the aimline. In Beverly's case, this ensured that goals would be set at a level where they could realistically be attained, while still ensuring that Beverly was making appropriate progress. It was decided that Beverly's goal might have been a little ambitious; however, she was making progress in the program and was developing a good relationship with the peer.

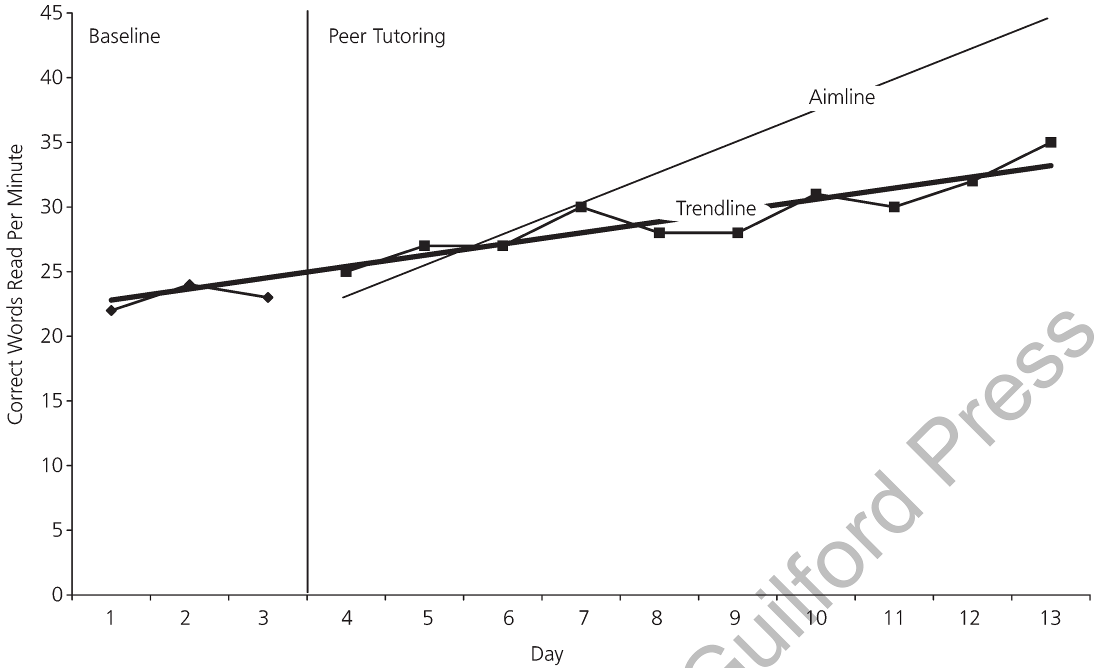

FIGURE 4.6. Beverly's CBM chart.

Multiple-Intervention Design

Before we leave our section on behavioral neuropsychology and problem-solving consultation, it is important to recognize that not all intervention designs discussed will fit nicely with the needs of a child, teacher, or parent. Certainly you want experimental control and good outcome data, but beyond that, you have to be sensitive to the needs of all parties, or the intervention effort will not be effective. Interventions that are easy are preferred, but they may not be effective. Others may be labor-intensive and have good experimental control, but because they are so cumbersome, treatment adherence or integrity is limited. This is where you, as the consultant, must work with the consultee to take into account the nature of the problem, the environmental determinants of the problem, and the resources available to affect behavior change. Schools that are implementing MTSS may already have progress monitoring procedures in place; in other cases, you will have to develop a data collection system that is not intrusive. We have found that pretest–posttest designs, and ongoing monitoring using direct behavior ratings or CBM probes, are the most acceptable to teachers. Case Study 4.6 and Figure 4.7 provide an example of alternative treatments for a child who does not respond easily to interventions.

Linking Assessment to Intervention: A Case Study

Considerations and Caveats

Now that we have given you a good understanding of assessment practices and measures, brain–behavior relationships, and consultation and intervention techniques, the next step

Case Study 4.6. Coping with Gary's OCD

Gary was a student diagnosed with OCD. His classroom teacher's main concern was Gary's incessant questioning about assignments during seatwork. Gary typically asked for clarification of the directions, and the meaning of individual items. The teacher wanted to decrease Gary's questioning and increase his on-task behavior. She agreed to count Gary's questions with a wrist counter during the seatwork period in her class. As can be seen in Figure 4.7, Gary's baseline average was a little over 10 questions per period. We decided to try a number of interventions, starting with the easiest to implement and gradually adding more intrusive ones. This called for a variation on the ABAC design, where the interventions were cumulative (it might be called an A-B-BC-BCD design). First, the teacher developed a checklist for completing seatwork, and she taught Gary to use it to answer his own questions. She then laminated it and let him check off each item for himself. During this intervention, Gary's questions decreased slightly, to an average of about eight per period. The next intervention added was a set of five tokens that Gary had to use to ask questions. He would turn in one token every time he asked a question; any question after that would not be answered. Gary's questions decreased again, eventually settling at five per period. At this point, the teacher added one more intervention: She provided Gary a reward—a choice of activity during the last 5 minutes of class—if he had one token left at the end of the period. This lowered Gary's questions to four immediately. If the teacher had felt that even fewer questions would be allowed (based on what was normally acceptable in class, perhaps one or two), she could have gradually increased the number of tokens necessary for a reward.

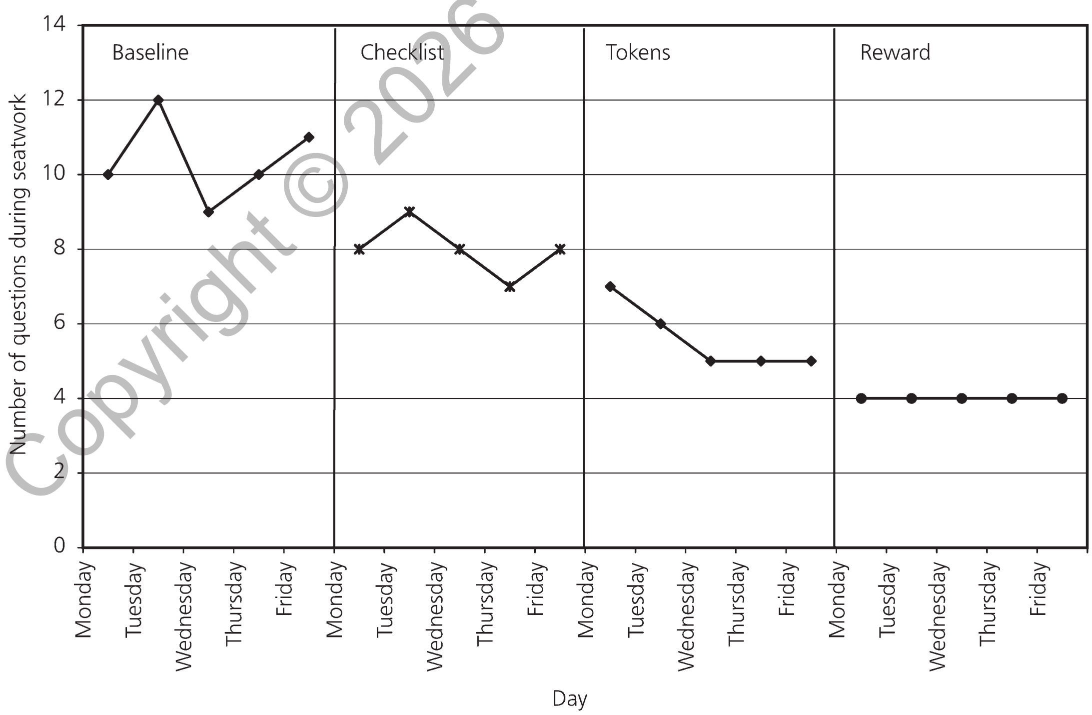

FIGURE 4.7. Gary's teacher questions.

is to bridge the gap between these apparently disparate areas of psychology. We provide you with one more case study, and detailed information in Chapters 5–8, in an attempt to make assessment information meaningful for individualized interventions for children with unique assets and deficits. You may be disappointed to find that we don't offer you diagnostic–prescriptive advice in the following chapters. We feel that this is where the early researchers on aptitude–treatment interactions went astray: Not all children learn the same way, even if they show similar neuropsychological profiles, so we don't oversimplify by saying, "If you have this disorder, then do this intervention." There is no guarantee that the first attempt at intervention will be successful, even with CHT, because most times we need to recycle or "tweak" the intervention until response is achieved. That is the thing about CHT, its recursive scientific method approach ensures hypotheses are generated and tested using data, with new hypotheses developed, and data collected and evaluated, until response is achieved.

To paraphrase an old adage, some interventions work for some children some of the time, but no interventions work for all children all of the time. You may feel confident that you have a good understanding of a child's psychological and neuropsychological strengths and weaknesses, but if you don't have ecological and treatment validity, then your results are of questionable value. Even if you have a good handle on the problem and the findings have ecological validity, the intervention you and the teacher choose may be ineffective. Don't dismiss the original findings; rather, try to understand why the intervention the teacher thought would be effective was not, and try to modify it or try another intervention. This recycling of interventions is necessary, whether you use a CHT approach or a regular behavioral consultation method. We provide you with assessment and intervention information about various learning and behavior disorder subtypes, but it is up to you to use CHT with the techniques presented in this chapter to individualize interventions for the children you serve.

Cognitive Hypothesis Testing for Scott's Motor Problem

Case Study 4.7 and Figure 4.8 present the completed CHT worksheet (see Appendix 4.3) for Scott, a student referred for "motor problems" in the classroom. We have purposely picked Scott's case because it highlights the use of CHT without the use of "neuropsychological" tests. We do this so that you can become familiar with the CHT procedure while using tests you already know. This case also demonstrates that CHT and neuropsychological analysis of the data can occur with typical cognitive/intellectual measures. In later chapters, we will provide you with several reading, mathematics, written language, and emotional/behavior disorder case study examples that use CHT and the neuropsychological tests described earlier in the chapter.

As you can see from Scott's case, the original "theory" about motor problems was not quite right, as the deficit appeared to be related to visual–spatial dorsal stream functions, or poor perceptual feedback to the motor system. The process would have continued with this case had all results come back negative. For instance, if we had seen signs of somatosensory difficulty, like differences in writing pressure or difficulty with pencil grasp, we might have done further testing in this area. Similarly, if we had seen difficulty with crossing the midline or using both hands together, we might have done more assessment. But we found enough testing and ecological validity evidence to support our hypothesis, and now have evidence that our CHT evaluation had ecological and treatment validity.

Although Case Study 4.7 and Figure 4.9 suggest that Scott's intervention was effective, it should be noted that Scott was receiving occupational therapy during this time, so the positive results could have been related to this intervention. Obviously, as time went on, both interventions may have had a positive and complementary effect. This is not a good empirical practice per se, as we don't want two interventions going on at the same time. But in real life, students will be receiving multiple interventions, and it may not be feasible to evaluate the effectiveness of each individually. The bottom line is that we need to help children, and if they get better and we have data that shows it, we are better off as a result. It might not get our case study published in the *Journal of Applied Behavior Analysis,* but it will lead to successful outcomes, and that is why this is a practitioner book, not a research one. Now that we have the methods to link assessment to intervention in multiple tiers of service delivery, the remainder of this book will focus on the neuropsychological aspects of specific academic and behavior problems experienced by the children we serve, and the interventions to help them achieve success.