CHAPTER 6

Working Memory Assessment Strategies

Several years ago, after the release of the Wechsler Intelligence Scale for Children Fourth Edition (WISC-IV; Wechsler, 2003), the Wisconsin Department of Public Instruction began receiving phone calls from concerned school psychologists who had been using the new scale. The WISC-IV was different from its predecessor, the WISC-III, in that it had a Working Memory Index and more working memory and processing speed subtests were included in the computation of the Full Scale IQ (FSIQ). Some school psychologists were reporting that the additional working memory and processing speed subtests were "pulling down" the FSIQ of students referred for learning disabilities, preventing the students from qualifying for learning disability services. At the time, a discrepancy model was being used to determine eligibility; thus, higher FSIQ scores were needed to produce a significant discrepancy between FSIQ and achievement scores. Some of the callers asked if they could omit the working memory and processing speed scores from computation of the FSIQ. Later, the publisher of the WISC-IV (Wechsler, 2003) would offer that option by providing norms for a 6-subtest FSIQ called the General Ability Index (GAI).

It is ironic that one of the strongest indicators of a learning disability, a low working memory score, was thought to be preventing the identification of learning disabilities. Since then, the discrepancy model has been abandoned, providing an opportunity for appropriate diagnostic use of low working memory scores. Even before the discrepancy model met its demise, some psychologists (e.g., Naglieri, 1999) were suggesting that we should look for consistency between processing scores and achievement, instead of discrepancy. That is, if a processing weakness is accounting for an academic skill deficit, then both scores should be low, instead of related cognitive processing scores being significantly higher than the academic skill score. For example, when working memory and basic reading skills scores are both low, there is evidence of a learning disability in basic reading skills.

Proponents of the Response-to-Intervention (RTI) model now being applied in most states argue that working memory and other processing scores are just as irrelevant as FSIQ because they provide no information about the existence of a disability or about interventions. This author agrees that the FSIQ and the discrepancy approach were ineffective diagnostic methods and that the FSIQ provided little direction about course or outcomes of intervention. However, given the extensive evidence documenting the relationships between working memory and specific learning disabilities (see Chapter 5), working memory and other related processing scores can provide valuable information in the determination of learning disabilities. This author also believes that treatment for working memory impairments is an appropriate component of a disabled student's educational plan (see Chapter 9). Advocates for RTI might respond to these arguments by saying that the cause of the learning disability is irrelevant and that educational interventions should focus on proven academic methods. They also believe that RTI can be used to identify students with learning disabilities. In RTI programs, students who do not obtain academic benchmarks receive increasingly intensive educational interventions. Students who do not respond to these evidence-based interventions are determined to have a learning disability. Unfortunately, for students with learning disabilities, RTI programs may be delaying the specialized interventions and individualized education that they need. In many instances, low working memory scores could have predicted academic failure and the need for special education (Gathercole, Brown, & Pickering, 2003). From this author's experience, it is frequently a deficit in working memory, or some other related process, that prevents students with learning disabilities from responding well to interventions. As such, their failure to respond to regular education interventions could have been predicted.

If working memory tests measure a capacity that is crucial for cognitive functioning and fundamental for academic learning, then working memory measures should be a central part of every assessment of cognitive abilities, especially when learning difficulties are a referral concern. An abundance of empirical evidence confirms that deficiencies in working memory skills are implicated whenever learners have difficulty acquiring any academic skill (e.g., Gathercole et al., 2005). Given the strong relationships working memory has with all areas of academic learning, working memory scores can provide valuable diagnostic information. Even if working memory scores are not used for diagnostic purposes, they can provide a better understanding of the learner's cognitive strengths and weaknesses. Once a learner's processing strengths and weaknesses have been identified, this knowledge can be used in designing individualized interventions and appropriate educational programming. Educational interventions are likely to be more effective when they take into account the specific processing weaknesses associated with the learning problem. For example, there are many potential processing problems that can account for a delay in basic reading skills. An intervention for a student with a phonemic awareness deficit should be distinct from an intervention designed for a student whose phonemic awareness is normal but whose working memory is deficient. Thus, assessment of working memory and related cognitive processes should be conducted with treatment in mind.

Although a comprehensive assessment of working memory and related cognitive processes is recommended when students are referred for learning problems, the informal methods and standardized tests should vary somewhat, depending on the specific referral concerns, the age of the student, and the measurement tools available. As such, there is not a standard battery for testing working memory; assessment procedures should be individualized for each case. The selection of informal procedures and tests should be determined by the hypotheses generated to account for the specific learning problems. Systematic planning before conducting an assessment is likely to increase the efficiency of the process and the usefulness of the results. Similarly, there is no standardized interpretative procedure that applies to all working memory test results. Also, the meaning of any given test score varies, depending on the examinee's other abilities and characteristics. Furthermore, the level of functioning indicated by test scores needs to be corroborated by other assessment data. In essence, the interpretation of test results should be based on both statistical analysis and clinical judgment.

In addition to recommendations for planning and organizing selective, cross-battery testing, this chapter describes informal assessment methods and materials that include interviews, observations, and records review. General guidelines for selecting and using working memory subtests, tests, factors, and batteries are offered (see also Chapters 7 and 8). To provide the reader with a knowledge base for subtest selection, a variety of working memory assessment paradigms are explained in detail, along with their intended use. The information will allow readers to determine which memory systems and which working memory components are tapped by various types of memory subtests. Furthermore, there are recommendations for assessing cognitive processes that are closely associated with working memory. Most importantly, this chapter contains detailed procedures for analyzing and interpreting working memory test results. The analytical and interpretative procedures apply to any and all working memory test batteries, especially those constructed in a cross-battery fashion. (Test-specific interpretative suggestions are provided in Chapters 7 and 8.) In addition, this chapter will confront the challenges associated with working memory assessment.

Working Memory Assessment Challenges

Psychologists and related practitioners must resolve several challenges when conducting a formal assessment of working memory. To begin with, there are no recently normed test batteries that are designed specifically for an in-depth assessment of working memory. (At least not any that are normed in the United States.) Several cognitive and memory scales include measures of short-term and working memory. Unfortunately, the majority of these scales only tap certain aspects of working memory; for example, some only tap the phonological and verbal dimensions of working memory while ignoring visuospatial components. These half-measures force the practitioner to administer additional instruments in a cross-battery fashion (discussed later in this chapter) whenever a comprehensive assessment is desired. Other batteries classify all types of temporary storage and processing either as short-term memory or as working memory, leaving the practitioner the responsibility of trying to separate the two. Cognitive batteries and memory scales seldom differentiate well between short-term memory and working memory, even when subtest and composite names indicate such. For example, digit span, the most widely used measure of temporal memory, has traditionally been used to produce a scaled score that combines digits forward and digits backward. Unfortunately, this procedure confounds assessment of two distinct memory components: phonological short-term memory (measured by digits forward) and executive working memory (measured by digits backward). Not surprisingly, digit span is sometimes categorized as a short-term memory subtest and sometimes classified as a working memory subtest.

For practitioners who wish to differentiate among a client's processes and components within short-term and working memory, selective testing and clinical interpretation is necessary. In nearly all batteries, working memory subtests do not discriminate well between subprocesses within working memory. One reason is that span measures taken during concurrent processing tasks are used to measure the capacity of executive working memory instead of attempting to directly measure an executive function, such as inhibition. Consequently many working memory subtests are appraising multiple aspects of working memory and short-term memory. Moreover, no standardized tools are available that explicitly attempt to assess Baddeley's episodic buffer or the ability to maintain long-term representations in an active, quickly retrievable state, an essential aspect of most contemporary theories. Also, there are no measures that allow direct assessment of specific working memory operations, such as the ability to recode visuospatial input into verbal code.

Another concern with existing cognitive and memory instruments is that, despite the perceived importance of strategies, there is only one scale that attempts to assess their influence: The Swanson Cognitive Processing Test (S-CPT; Swanson, 1995; reviewed in Chapter 8) attempts to broadly assess the use and impact of memory strategies on working memory performance. Span procedures, the most common form of measurement, provide little indication of strategy selection, use, or efficiency. In fact, digit span does not even correlate with strategy selection and efficiency (Imbo & Vandierendonck, 2007). Of course, examinees are often utilizing strategies, such as subvocal rehearsal. Following the paradigms of laboratory researchers, some test authors have designed tasks that prevent strategy use by introducing irrelevant interference. Sometimes, the intent of interference is to measure *pure* short-term memory span. *Pure* short-term span is thought to indicate the capacity of short-term memory when it receives no assistance from working memory. To obtain a pure span, distracting interference is introduced that prevents rehearsal and further processing of the stimuli. It is not surprising that such restricted measures are poor predictors of academic performance. In the real world, short-term and working memory operations typically function in conjunction with concurrent processing and strategic applications.

From the perspective of contemporary theories of working memory and the applied model proposed in this text, another major concern with existing measures is their complete failure to assess the interaction between long-term and working memory, as well as long-term memory's contributions to short-term and working memory (Masoura, 2006). For example, almost no existing measures provide a means of determining how much recall can be attributed to short-term storage and how much to long-term storage, a problem that might be solved by the opportunity to contrast retention of words with that of nonwords. Moreover, omitting any attempt to factor in the influence of long-term memory or long-term retrieval results in an incomplete picture of total working memory capacity. The amount of information currently active in working memory may be greater than indicated by the length of a short-term memory span; that is, memory span may not provide an accurate appraisal of total working memory capacity (Richardson, 1996b). Similarly, some tasks identified as long-term memory subtests may actually be measuring working memory components.

To complicate matters further, existing assessment tools, as well as texts on the subject, provide little guidance on when to assess and how to interpret related processes. For example, discovery of a phonological short-term memory impairment in a preschooler should mandate testing of phonological processing. Furthermore, there are multiple interpretation challenges when trying to make sense of data gathered during a multidimensional assessment of working memory. For instance, how does a six-digit span impact learning differently than a five-digit span? From a broader perspective, it is difficult to determine the ramifications short-term and working memory scores have for academic learning and interventions.

Finally, one has to wonder how often we are actually measuring working memory capacity. The typical working memory subtest is a dual-task paradigm that requires examinees to concurrently perform a secondary task that prevents strategy use and adds to the working memory load. The outcome of adding to the working memory load is typically a reduction in short-term memory span, such as the decline in span from digits forward to digits backward. In one sense, it is paradoxical because we believe that working memory capacity is greater than that of short-term memory capacity alone. On the other hand, the effect of the secondary task is to mostly prevent working memory from assisting short-term memory through rehearsal. So, the resulting span may be nothing more than an indication of pure short-term memory span. What does pure short-term memory span tell us about overall working memory performance and capacity? Would it not make more sense to measure working memory capacity by presenting examinees with a complex task and allowing them to freely use all of their resources and strategies to preserve information? Wouldn't that be a more accurate appraisal of real-world working memory functioning? Fortunately, the additional processing required during many complex test activities does not totally disrupt efforts to maintain information. Nonetheless, preventing working memory from full normal functioning when trying to evaluate it seems counterintuitive. Experimental assessment paradigms originally designed to determine the extent of resource sharing between memory systems seem to have questionable validity when the purpose is to measure the overall level of an individual's working memory capacity.

Perhaps working memory measures should not be so harshly criticized; all types of psychological testing present challenges. And all types of test performance are subject to various influences. For example, at first glance, a digits forward subtest appears to be a narrow measure of temporary phonological storage. Nevertheless, it cannot be assumed that only passive storage is being tapped, as some examinees may be using sophisticated executive strategies to extend the span. Clearly, there is a need for additional assessment tools that are designed to specifically measure and differentiate among critical aspects of working memory. Until such products become available, this chapter and the following two chapters are intended to provide practitioners with appropriate guidance, particularly in test selection and the interpretation of results.

Distinguishing Between Short-Term and Working Memory Measures

Most of the standardized working memory subtests found in contemporary test batteries have their origins in experimental research paradigms. The theoretical basis for the majority of these research paradigms and associated subtests is Baddeley's (1986) tripartite model. Because Baddeley views the short-term stores as subsidiary components of working memory, many standardized measures take the same perspective, resulting in "working memory" subtests and factors that do not discriminate between short-term memory and working memory. Despite the lack of differentiation, most test authors adhere to the traditional definition of working memory as processing while trying to retain information in the short-term. Given this definition, a subtest is considered a measure of working memory whenever mental processing is required while trying to maintain information, and a subtest is considered a measure of short-term memory whenever storage but no additional processing is required. The assessment challenge is not that some measures of working memory contain elements of short-term storage, but that test batteries often do not provide measures that isolate short-term memory. The ability to measure short-term memory and working memory separately is important because the functions are different.

Even when test batteries have separately labeled subtests and factors, the titles may not accurately reflect what is being measured. For example, many subtests identified as short-term memory tasks may be primarily measuring working memory. To determine whether a subtest is primarily measuring short-term memory or working memory, consider the task and ask:

1. Is passive, serial recall of the information the only requirement? If so, the task can be classified as short-term memory.
2. Is any manipulation or transformation of the information required? If so, the task is essentially measuring working-memory functions.
3. Is any concurrent or intervening processing required, such as the insertion of an interference task? If so, the task clearly involves working memory.
4. Are both storage and processing required? If so, the task is primarily measuring working memory.
5. Does the task involve the concurrent retention of both visuospatial and verbal information or the recoding of one modality into the other? If so, it is primarily measuring working memory.

Short-Term and Working Memory Testing Paradigms

Nearly all measures of short-term and working memory developed to date involve the measurement of span. In general, memory span is defined as the maximum amount of sequential information an individual can remember accurately (Gathercole, 1999). In measurement, span is the number of items the examinee can recall after a short interval; usually, the items must be recalled in serial order. A span procedure typically begins with only one or two items to remember. When the individual responds correctly to enough trials at a given level, the amount of material to be remembered is increased, usually by one item at a time. Short-term memory span is the number of items retained when no concurrent processing is required, whereas working memory span is the number of items recalled after processing the same or other information. Span activities can be classified as either *simple span* or *complex span*. Simple span is presumed to measure short-term memory, whereas complex span is considered a measure of working memory. Measures of simple span require only the passive retention of information. Examples of simple-span tasks include the serial recall of letters, digits, words, or nonwords.

Complex span activities require effortful processing of information while trying to retain a list of items for a short interval (Bayliss, Jarrold, Baddeley, & Gunn, 2005). Complex spans are considered measures of working memory because they require the involvement of executive working memory. The classic example of a complex-span task is the reading span task (Daneman & Carpenter, 1980) in which the examinee reads aloud a sentence, responds to a simple question about it, and then later must sequentially recall the final word of each sentence. Other complex span activities include listening span, operation span, and counting span. Whereas simple-span tasks are used to measure phonological short-term memory, complex-span activities measure verbal and executive working memory. Complex-span tasks require the coordination of storage and processing—coordination is one of the primary functions of executive working memory. Although verbal working memory plays a role, individual differences in complex-span tasks seem to be primarily due to differences in the executive component of working memory (Conway et al., 2002).

Complex-span tasks are usually constructed in a manner that is intended to replicate real-world memory functioning, such as that which occurs while reading for comprehension. This may be why complex spans generally have higher correlations with academic learning and higher cognitive functions than simple spans do. There is no denying that complex-span activities put executive working memory to the test. It is the responsibility of executive working memory to cope with the challenge of retaining information while minimizing the impact of the interference that has been introduced. If the interference cannot be inhibited, rapid switching between processing the interference and rehearsing the information may preserve some of the span. Therefore, the span resulting from complex-span activities does allow us to make inferences about the effectiveness of executive working memory. A sub-average complex span is one indication of a deficiency in executive working memory.

In the following sections, several common measurement paradigms are categorized and described (see Table 6.1). All of these have been used in experimental research and the majority have been incorporated into one or more standardized test batteries. Additional working memory assessment paradigms may be found in Strauss, Sherman, and Spreen (2006).

Phonological Short-Term Memory

Forward Digit Span

Nearly everyone in education and psychology is familiar with the most frequently used measure of short-term and working memory. Joseph Jacobs, a London schoolmaster who wanted to measure the mental capacities of his pupils, developed the first digit span test in the 1880s. It was later incorporated into Binet and Simon's 1905 intelligence scale and has been a mainstay of cognitive assessment ever since. Forward digit span measures phonological short-term memory, whereas backward digit span is categorized under executive working memory. Compared with forward span, backward span usually yields a reduction in span of at least one digit, with more of a reduction with older subjects. In most standardized procedures the examiner says the digits at the rate of one per second, although some tests double the presentation rate to reduce the opportunity for rehearsal. Some applications also explicitly prevent examinees from chunking the numbers when responding.

Table 6.1 Common Measurement Paradigms for Memory Components

Short-Term Phonological	Short-Term Visual	Working Verbal	Working Visuospatial	Working Executive	Long-Term Retrieval
Forward digit span; Letter span; Word span; Nonword span	Forward block-tapping span; Visual digit span	Memory for sentences; Memory for stories; Listening span; Reading span; Operation span	Backward block-tapping span; Counting span	Backward word span; Backward digit span; Computation span; Trail-making; Stroop; Trail-making; *N*-Back; Random Generation	Retrieval fluency; Rapid automatic naming

Letter Span

Almost identical to digit span in structure and administration, letter span, which also has a backwards option, is ideal for use with examinees who are deficient in mathematics skills. Delayed mathematics development can influence the recall of digits, producing a score that underestimates overall phonological short-term storage capability. Letter span also keeps the processing close to the phonological level because letters activate fewer meaning-based long-term representations than words.

Word Span

As the name implies, word span is a series of words the examinee must recall in order. Like digits, they are typically presented at the rate of one per second. The words should be unrelated and categorical groupings should be avoided so that verbal working memory and long-term representations have less impact on performance. Also, the words should be relatively short, typically one or two syllables in length. Because of the influence of total articulation time on retention, spans with a greater number of syllables are more difficult to maintain than spans with fewer syllables.

Nonword Span

Nonwords, also know as pseudowords or nonsense words, are particularly ideal material for narrowing the assessment to simple phonological short-term memory span. With nonwords, the examinee cannot rely substantially on long-term semantic memory to supplement recall, as such items have no long-term representations other than basic phonetic properties. Consequently, when articulation time is equivalent, nonword span is typically shorter than word span. The best nonwords are those that bear almost no resemblance to recognizable syllables or words, and the number of syllables in each nonword should be limited to two. Also, sequences of nonwords should not include any items that rhyme. Individuals with phonological processing problems often have even more difficulty with nonwords than actual words. Interestingly, nonword span is a better predictor of vocabulary development than is word span (Gathercole, 1999).

Verbal Working Memory

Phonological short-term memory handles verbal information when only a few verbal items are involved and no transformation of the information is required. In contrast, verbal working memory is required when the information is long and complex, is more meaningful, needs to be manipulated, or when long-term semantic memory plays a significant role during recall. The purest measures of verbal working memory are those that do not introduce interference, a dual-processing task, or a secondary processing task. Thus, the complex spans classified here as verbal working memory tasks also tap executive working memory. Verbal working memory span tasks also depend on knowledge and processes beyond working memory; for example, many tasks involve verbal ability and some tasks require quantitative ability (Conway et al., 2003).

Memory for Sentences

Memory for sentences may be the purest form of verbal working memory, as it does not involve any dual processing that would enlist executive working memory processes. It also has the benefit of being distinct from phonological storage because meaning-based encoding will occur with sentences, resulting in spans that are significantly longer than spans for series of unrelated words.

Memory for Stories

Another short-term retention activity that is a relatively pure form of verbal working memory involves the retelling of brief stories. Immediately after hearing a story, the examinee is directed to retell as much of the story as he or she can remember. Complete and sequential recall is not required; points are awarded for each key element recalled and paraphrasing is usually allowed. Even more so than memory for sentences, success at this meaning-based task will depend heavily on support from activated long-term semantic memory structures.

Reading Span

Reading span, the complex-span task originally developed by Daneman and Carpenter (1980), has been a prototype for many verbal and executive working memory measures. Reading span typically requires the examinee to read a series of sentences and then sequentially recall the final word of each sentence. The task has been used in a variety of forms; for example, one version limits rehearsal by requiring a simple response to each sentence (Duff & Logie, 2001). Sentences that are more complex are thought to lead to a greater demand on working memory resources and a consequent reduction in span.

Listening Span

With listening span, the examiner reads a series of sentences to the examinee, and then the examinee recalls the final word of each sentence. The task can be made more challenging by inserting a question, typically a verification question, the examinee must answer before the next sentence is read. For example, the sentence might be "Cats bark at dogs," followed by the question "Is that true?" A cloze procedure can also be used with listening span (Siegel & Ryan, 1989). In this version, the examiner reads aloud sets of short sentences in which the final word is missing—for example, "Apples are red and bananas are ———." Examinees are instructed to complete the sentences. When all the sentences in the set have been completed, the examinee must repeat, in order, the words he or she used in completing the sentences. The word used to complete each sentence should be obvious, but selection of the correct word is not required.

Operation Span

The operation span task is another complex-span task similar to reading span. Subjects read aloud a mathematical equation and then state whether or not the equation is correct before reading a word aloud. Another variation is to have examinees solve a simple math problem before being exposed to the stimulus word. After completing a set of such items, subjects must recall the words in the correct sequence (Engle, 2002).

Visuospatial Short-Term Memory

Visual Digit Span

Rather than hearing orally presented digits, the examinee views printed digits. Although some tasks display the entire set of digits simultaneously, it is best to uncover one digit at a time at the rate of one per second. Sequential presentation seems reasonable when ordered recall is required, and it reduces the opportunities for chunking. Moreover, a direct comparison with auditory digit span is more valid when the method of presentation is similar.

Block-Tapping Span

The classic Corsi block-tapping task, or variations thereof, is often used to assess visuospatial short-term and working memory. The Corsi block task, which consists of an array of nine randomly placed blocks, is the visuospatial equivalent of digit span. Thus, the forward span is a measure of visuospatial short-term memory, whereas the backward span measures visuospatial working memory and executive working memory. To administer the task, the examiner taps or touches a set of two to nine blocks in a preselected random sequence at the rate of one per second. Examinees must then recreate the tapping sequence. In models that distinguish visual and spatial memory, the Corsi task is thought to primarily measure spatial span (Baddeley, 1996). In normal individuals, Corsi block span is typically about two items less than aural digit span.

Visuospatial Working Memory

Counting span involves counting a series of visual arrays and subsequently recalling the totals in order. An informal version can be constructed with index cards. The examinee counts the dots on the card (usually an amount not exceeding 10), turns the card over, and reports the number of dots. Then a second card is added and, after counting the dots on it, the examinee reports the number of dots that were on the first card followed by the number on the second card. More cards are added to increase the difficulty level. A variation that prevents grouping of dots for quick counting is to intersperse red dots with black dots and have the examinee count only the red dots.

Executive Working Memory

The dual-task technique is the classic method for assessing executive working memory. Dual-task activities require the subject to simultaneously perform two tasks—the primary and the secondary. The primary task is the short-term maintenance of stimuli. The secondary task is designed as interference with the purpose of disrupting any strategies that would facilitate maintenance of the information in the primary task. The introduction of interference assures the involvement of executive working memory. The notion behind the dual-task paradigm is that there is a limited pool of working memory resources that the primary and secondary tasks must share. Without the full amount of resources usually available, performance on the primary task (retaining the information) is presumed to decrease. Dual-task measures apply well to working memory functioning in the real world. For example, in the classroom, students must continually deal with distracting interference, some of it internally generated. In order to introduce the most interference, the secondary tasks should be in the same modality. For example, the maintenance of visuospatial information is disrupted by concurrent visuospatial tasks but not by secondary verbal tasks (Olive, 2004). Because of the additional complexity, memory spans measured with executive processing components are often shorter than their uninterrupted, straightforward memory counterparts (phonological short-term memory or verbal working memory). For individuals with an executive working memory deficiency, their difficulty will become more pronounced as the complexity of dual-processing increases. Individuals who do not display a decrement in span usually take longer to complete the secondary task, probably because they are shifting back and forth between the two tasks. In addition to the following tasks, all of the complex span tasks listed under verbal working memory involve executive working memory to some degree.

Computation Span

Computation span is another variation of complex span (De Jong, 1998) in which the examinee must retain digits while making simple computations. The examinee orally reads simple addition and subtraction problems and gives the answers aloud. After each computation, the examiner orally presents a digit. To prevent rehearsal, the examinee is required to begin the next problem immediately after presentation of the digit. After the series of computations is complete, the examinee must recall the digits in order.

Star Counting

Star counting is a task that especially requires inhibition. In the first part, the examinee is directed to count stars in rows of differing amounts, with the number at the beginning of the row showing the total number of stars in that row. However, a plus or minus sign is inserted at varying locations between the stars. These signs indicate the direction, forward or backward, in which subsequent stars have to be counted. The final count is the answer to the item. In the second part, more inhibition is required, as the meaning of the signs is reversed, with a plus meaning backward counting.

Stroop Task

In experimental research, the classic Stroop task has been repeatedly used to measure working memory, specifically the executive ability to focus attention and to inhibit overlearned responses or irrelevant information. The most challenging Stroop task is to read a list of color words that are printed in ink colors incongruent with the printed word (e.g., the word *red* in green ink).

Trail-Making

Cognitive psychologists have used trail-making tests to assess the ability to switch between operations or retrieval strategies. This paper-and-pencil task usually consists of numbers or letters on a page that must quickly be connected in numerical or alphabetical order. Difficulty may be increased by combining numbers and letters and requiring the examinee to connect them in an alternating fashion (e.g., 1-A-2-B).

N-Back

The *n-back* task is another popular cognitive neuroscience paradigm that has been widely used. The general procedure requires the examinee to respond to a stimulus, for example a number, when it matches a previously presented item that is a predetermined number of items back. The presentation of stimuli and reporting by the examinee is continuous (Cicerone, 2002). For example, using a deck of cards in a

2-back task, the participant is required to name the card that was exposed two cards prior to the one currently exposed. So if the exposure sequence is "10," "2," "6," "8," the subject says "10" when the "6" is exposed and "2" when the "8" is exposed, and so on. The *n*-back task is often used to measure executive functioning. Recently, the validity of the *n*-back task was challenged by Kane, Conway, Miura, and Colfiesh (2007), who argue that the *n*-back task may not be a valid measure of working memory as it does not correlate significantly with other measures of working memory span.

Random Generation

Random generation is another task that is typically employed with a concurrent storage task. In this approach, the subject must randomly generate numbers, letters, or words from a semantic category while avoiding repetitions. With letters, subjects sometimes are required to generate triads while avoiding acronyms, natural letter sequences, or words.

Verbal-Spatial Association

A new working memory measurement paradigm was recently proposed by Cowan et al. (2006). They developed an experimental task that requires verbal-spatial associations. The task involves remembering the location of names presented on a computer screen. Such a task is thought to measure working memory for abstract information. The task is also unique in that it creates a challenging working memory measure without introducing an unrelated processing task that disrupts normal working memory functioning or prevents retention strategies. The new paradigm may also be the first to tap Baddeley's episodic component of working memory. Thus, the task is similar to real-world functioning in that cross-modal associations are required.

Long-Term Retrieval

Retrieval Fluency

Tasks of this nature are sometimes referred to as *associational fluency* or *verbal fluency* tasks. These activities are intended to measure the examinee's speed of long-term memory retrieval. During these effortful searches, semantic categories, such as food, are commonly used, but initial phoneme sounds are also used sometimes. Working memory plays a role during directed searches; for example, working memory determines whether retrieved items meet search criteria.

Rapid Automatic Naming

Rapid Automatic Naming is a timed long-term retrieval task found in several types of scales, including achievement batteries. Typically, the examinee is required to say the names of symbols or pictured objects. Proficiency at the task is influenced by several factors, including processing speed and the strength of the associations stored in long-term memory. Rapid Automatic Naming probably requires fewer working memory resources than retrieval by category.

Paired Associate Learning

As learning is inextricably linked to memory, most memory batteries include learning subtests that involve repeated trials of the same information. These tasks require the examinee to learn pairings of unrelated words or pairings of words with symbols. The activities are distinct from short-term and working memory paradigms in that there are repeated trials during which corrective feedback is provided. Such subtests are intended to measure how efficiently the examinee can learn novel material. Although long-term retention is necessary, working memory is thought to play an essential role because new information must be encoded meaningfully—a function of working memory. The benefit of including such tasks in working memory testing is that they allow clinicians to assess how the individual's working memory is influencing learning. Nonetheless, for the sake of consistency, paired associate and related learning tasks should be interpreted as measures of learning, not working memory per se or even as the long-term retrieval function associated with working memory (see working memory component definitions in Chapter 3). Also, delayed-recall subtests should be interpreted as measures of long-term storage, a capacity that is distinct from short-term and working memory capacities.

Hypothesis-Driven Assessment of Working Memory

Psychologists and related professionals who conduct psychological and educational evaluations often come across individuals who have been suffering from a disability but have never been diagnosed. Sometimes these individuals have been evaluated previously but the disability was missed or another disorder was diagnosed instead. For example, young children with learning and behavior problems are sometimes diagnosed with ADHD when they actually have a learning disability. Even when children are correctly diagnosed with a specific learning disability, the underlying cause of their learning problems usually remains unknown or is misunderstood. The occurrence of misdiagnoses and the limited understanding of *why* an individual has learning problems can often be attributed to assessment procedures that missed the mark or were too superficial. Such assessment procedures usually consist of "standard battery" procedures (i.e., the same procedures and the same set of tests are used in every case, regardless of the specific referral concerns). An alternative that can reduce misdiagnoses and increase in-depth understanding of the individual is to adopt a hypothesis-driven approach to psychoeducational assessment that results in an individualized assessment plan for every referral.

Prior to selecting informal procedures and standardized tests, evaluators should generate and select hypotheses that account for the specific referral concerns. *Referral* *hypotheses* are statements that explain or account for the presenting problem. These hypotheses get at the underlying cause of the problem or at *why* the problem is occurring. In a sense, referral hypotheses point toward the suspected deficit. Whether we express them or not, we all generate hypotheses about other people's problems; for example, when we suspect someone is performing poorly because he or she lacks motivation, we have, in effect, hypothesized that the individual lacks motivation. This text promotes hypotheses for working memory and related processing problems but evaluators should not restrict hypotheses to these domains; behavioral instead of cognitive hypotheses can also be chosen.

The generation and selection of appropriate memory and processing hypotheses begins with a careful examination of the presenting problems. Especially when it comes to assessing memory, the referral reasons need to be examined and clarified so that all relevant concerns are investigated. This preassessment step may require a preliminary interview with the client or person making the referral to determine the precise nature of the concerns. Simply relying on the initially stated presenting problems is often ineffective because the individual seeking an assessment or intervention often fails to report important behaviors or performance. The client or the person making the referral may posit some hypotheses of their own. If these hypotheses are appropriate, they should be considered and included in the assessment plan. For example, a parent will sometimes hypothesize that poor instruction is the cause of the child's reading problems. Such a hypothesis should be considered but will not be included in a psychoeducational assessment. The generation and selection of referral hypotheses can be completed once all of the referral concerns have been examined and clarified. One structured approach to accomplishing this task is to use the *Working Memory Assessment Plan* form provided in Appendix B. When learning disabilities are a possibility, hypothesis-driven assessment planning keeps the focus on working memory components that have the strongest relationships with the areas of academic deficiency. For example, whenever basic reading decoding skills are a concern, the first memory-based hypothesis should be that the examinee has a significant weakness in phonological short-term memory. After logical or empirically based hypotheses worthy of investigation have been selected, the examiner selects assessment methods (e.g., observations, interviews, and standardized tests) that will allow the testing of each hypothesis. The case of Joey (introduced in Chapter 5) will be used to illustrate hypothesis-driven assessment planning (see Table 6.2). In Joey's case, extreme difficulty in basic reading decoding is one of the primary concerns. Hypotheses with a high likelihood of explaining this difficulty are: (a) Joey has difficulty decoding because of weak phonological processing; (b) Joey has difficulty decoding because of weak phonological short-term memory; and (c) Joey has difficulty decoding because of weak long-term storage and retrieval. For some reminders of plausible memory and processing hypotheses see Tables 5.2 and 5.3 in Chapter 5.

After the memory and processing hypotheses have been selected, the next step is to determine the best assessment procedure for testing each hypothesis. In many cases,

Table 6.2 Working Memory Assessment Plan

Examinee's Name: Case Study of "Joey" DOB: ——— Age: 9 Grade: 4 Testing Dates: ——— Completed By: ———

Referral Concerns	Memory or Processing Hypotheses	Memory Factors or Subtests	Observations	Interviews	Other Method
1. Difficulty decoding words when reading	1.a. Phonological processing weakness 1.b. Phonological STM weakness 1.c Long-term storage and retrieval weakness 1.d Visuospatial STM or WM weakness	1.a. WJ III Phonemic Awareness cluster 1.b. WJ III Memory for Words 1.c. WJ III Rapid Picture Naming and Retrieval Fluency 1.d WISC-IV Integrated Spatial Span Forward and Backward	1.a. Observe while reading in the classroom	1.a. Question parent about rhyming 1. b.c.d. Question teacher about signs of weaknesses	1.a.b.c. Review special education records
2. Difficulty with reading comprehension	2.a. Fluid reasoning weakness 2.b. Executive WM weakness 2.c. Long-term storage and retrieval weakness	2.a. WJ III Fluid reasoning cluster 2.b. WJ III Numbers Reversed and Auditory Working Memory 2.c. See 1.c.	2.c. Observe retention of material in classroom	2.c. Question teacher about long-term retention	2.c. Check early childhood history for long-term retention problems
3. Difficulty with written expression	3.a. Executive WM weakness 3.b. Processing speed weakness 3.c. Language development delay	3.a. See 2.b. 3.b. WISC-IV Processing Speed 3.c. Refer to speech/language therapist	3.b. Observe for processing speed in classroom and during testing	3.c. Question parent about language development	3.c. Check early childhood history for language delay

Directions: Under Memory Factors or Subtests, specify the name of the battery, factor, and/or subtest that will be used. Under Observations, specify when and where. Under Interviews, specify with whom.

an informal procedure, such as an observation, will provide enough data to test one or more hypotheses. For other hypotheses, standardized testing will be necessary. In most situations, one instrument, such as a comprehensive memory scale, can be used to test several hypotheses. For many hypotheses, a single subtest will provide an adequate test of one hypothesis; however, the use of two subtests provides a more reliable sampling of the component or process being investigated. Selection of a specific subtest should be based on the extent to which the subtest measures the specific underlying process thought to account for the problem. To determine subtest classifications based on the Integrated Model of Working Memory, see Chapters 7 and 8 and consult Appendices A and E. The hypothesis-driven method meshes well with a cross-battery, selective testing approach (discussed later in this chapter) because selective testing limits standardized testing to the subtests that address the hypotheses, avoiding administration of entire batteries just because they contain the subtests of interest. For example, in Joey's case (see Table 6.2), only one subtest from the Woodcock-Johnson III (WJ III) was selected to test the phonological short-term memory weakness hypothesis, and two WJ III subtests were selected to test his long-term storage and retrieval. Effective use of the hypothesis-testing approach depends on delaying the selection of standardized measures until after the hypotheses have been identified. In conclusion, memory and processing assessment should be hypothesis driven, not battery driven. Selecting the instrument(s) in advance of choosing hypotheses, or using the same instruments with every assessment, is not only inefficient but poor practice as well.

After hypotheses that account for the problems have been selected, it may be necessary to expand the assessment further. One drawback to limiting testing to the areas of concern is that potential strengths are ignored. Assessing strengths creates a balanced view of the individual and allows the identification of assets that can be utilized during interventions. There are also instances when not enough is known about the individual's difficulties, as in the case of a young child, or the individual is experiencing a broad range of difficulties related to working memory. In such cases, a more comprehensive assessment that goes beyond the derived hypotheses is warranted. Nonetheless, "standard" assessment batteries and redundant testing should still be avoided. In instances where the information is so limited that it is difficult to select any plausible hypotheses, it may be best to begin with screening. In such cases, a broad-based but short measure with high predictive validity will serve the purpose. For instance, a digit-span task is ideal, as digits forward measures phonological short-term memory and digits backward measures executive working memory. If performance on either is subaverage, then these memory components and verbal working memory should be tested in depth.

From a scientific perspective, the main purpose of generating hypotheses before collecting and analyzing data is to increase the objectivity of the investigation and the interpretation of results. When using a hypothesis-testing approach to assessment, it is critical that evaluators keep an open mind and avoid hypothesis confirmation bias

(Hale & Fiorello, 2004). Confirmation bias leads to ignoring data that do not support a hypothesis while focusing on pieces of data that do support it. The data for and against a hypothesis should be carefully weighed before reaching a conclusion. The best way to counteract confirmation bias is to assume that the hypothesis is false unless there is considerable convergent evidence supporting it.

When hypotheses are selected prior to assessment, they are known as *a priori hypotheses*. As data are collected and analyzed, new insights often arise and more hypotheses are added. Also, when the results are inconsistent or are not what was predicted from the hypotheses, it is often necessary to generate new hypotheses that account for the unexpected findings. These new hypotheses are referred to as *a posteriori hypotheses*. When additional hypotheses are considered, the examiner should cycle back to an earlier step in the planning and assessment process. Follow-up testing may be necessary. In some instances, a posteriori hypotheses can be evaluated by reexamining already existing assessment data.

Basing memory evaluations on hypothesis testing serves several functions and has several advantages. First, explicitly generating and selecting hypotheses forces the examiner or multidisciplinary evaluation team to carefully think about and consider the referral concerns and how best to assess them. Each hypothesis should be one possible explanation for *why* the individual is experiencing a specific problem. Second, the hypothesis approach can increase the understanding of the learner even before assessment begins. Third, following a hypothesis-testing approach truly individualizes the assessment, forcing the evaluator to abandon a standard battery approach and adapt to the unique concerns of each case. Fourth, following the hypothesis-testing method, coupled with selective testing, results in an efficient, time-saving assessment that avoids redundancies while measuring all of the processes that need to be assessed. In conclusion, using hypotheses as the basis for assessment planning results in an individualized, comprehensive, and efficient assessment. Hypothesis-driven assessment should result in an in-depth understanding of the examinee. It also provides direction for subsequent interventions.

As this point, the reader may be wondering why we should bother with generating hypotheses—why not complete a comprehensive assessment of working memory in all cases? The main reason is that it is not necessary in every case. Some specific reasons include:

1. Not all memory components have a strong relationship with every problem. For example, if the only referral concern is mathematics reasoning (and mathematics calculation skills are above average), it is highly unlikely that phonological short-term memory has much to do with the problem.
2. Some hypotheses are determined by the age of the examinee. For example, if a 4-year-old is having speech and language development problems, a hypothesis involving executive working memory would be premature because executive working memory has hardly begun to develop.

3. The functioning level in some memory components may have already been evaluated during previous testing and found to be quite normal.
4. The functioning level in a memory component may already be a clear strength. For example, a child with superb spelling, vocabulary, and phonetic decoding skills is highly unlikely to possess a weakness in phonological short-term memory.

Multimethod Assessment of Working Memory

A test score alone should never be used to define an individual. Psychological tests and the scores they produce are fraught with inherent weaknesses, among them measurement error and limited construct validity. When cognitive abilities, such as working memory processes, are assessed, it is crucial that data collected through other methods are considered along with the test scores. Informal assessment methods include observations, interviews, rating scales, and records review. These informal procedures also have serious drawbacks, such as inconsistent application and being easily influenced by biases. Moreover, it is difficult to analyze, interpret, and draw suitable inferences from informally gathered data. Clearly, standardized testing is necessary during a memory assessment, despite the shortcomings of standardized batteries. Standardized measures are particularly necessary when a learning disability diagnosis is being considered or when the goal is to determine whether the examinee has strengths, weaknesses, or deficits in working memory. Furthermore, contemporary measures of working memory capacity do reliably predict real-world performance in academic learning and daily functioning (see Chapter 5). Thus, standardized testing is recommended whenever working memory weaknesses are hypothesized; however, informal procedures should always be conducted along with testing. Generally, the validity of assessment results improves as the number of data-collection methods increases. When different sources of data provide convergent evidence, evaluators can be more confident of test results. Diagnostic and programming decisions should never be based on one source of data alone. Best practices in assessment of working memory and related cognitive processes require multidimensional assessment.

Reviewing Records for History

Unless there has been an acquired brain injury, subnormal capacity in short-term and working memory has usually been present since birth. That is why gathering the individual's history is an important aspect of working memory assessment. Although much of the history can be obtained through interviews, reviewing available psychological and educational records can provide documentation of previous performance and ongoing problems. When reviewing educational records, one must often play detective in looking for clues indicative of ongoing problems because records seldom contain explicit comments about memory, especially working memory. When the records are anecdotal, the reviewer should look for reports of behaviors that are associated with working memory difficulties (see Table 6.6 for a listing of working memory related behaviors). For instance, a teacher's comment that a fourth-grade student is still using his or her fingers during arithmetic calculation is a sign of working memory deficiencies. When searching for evidence of ongoing memory and processing weaknesses, the reviewer should pay particular attention to the first report of the suspected weakness and how frequently and consistently it has been reported. Isolated reports should not be accepted as adequate documentation but rather as red flags that need to be investigated further. Some reported difficulties may be situation-specific behaviors that serve a function in a particular environment but are not actual memory or processing deficiencies. Moreover, deficits in basic processes, such as short-term and working memory, should be evident as soon as a child enters school. However, deficits in the higher level processes such as fluid reasoning and executive working memory may not become apparent until the later childhood years when the frontal lobes become fully developed.

In addition to educational records, a review of medical, neurological, and psychological reports is pertinent whenever there are concerns about memory and cognitive processing. When test records are available, it is important for the informed reviewer to reexamine, reanalyze, and reinterpret the test scores instead of relying on the analysis and interpretation found in the existing report(s). The primary reason for this is that previous examiners may have been uninterested or unaware of working memory measures embedded in the scales that were used. For example, a child may have obtained an extremely low Digit Span subtest score, but it was interpreted as a measure of distractibility, instead of short-term and working memory. To reanalyze previous test scores from a contemporary working memory perspective, practitioners should classify subtests according to the tables in Appendices A and E, use the *Working-Memory Clinical Analysis Worksheet* found in Appendix C, and follow the interpretative guidelines discussed later in this chapter and subsequent chapters.

Interviews

Teacher, parent, and student interviews are an essential component of any comprehensive psychoeducational assessment. Information garnered during preliminary interviews can be used to generate hypotheses that will guide remaining assessment and testing procedures. Subsequent interviewing allows more in-depth investigation of variables thought to underlie the referral concerns and the test results. Unfortunately, structured and semistructured interview formats seldom include items specifically related to working memory. Therefore, the interviewer must make a special effort to develop questions that allow assessment of possible working memory impairments. Information obtained from interviewing should never be used alone to confirm hypotheses or reach diagnostic decisions. Rather, interview data need to be corroborated by other assessment data before their accuracy is accepted.

Teacher Interviews

With teachers, the most efficient and focused strategy is to first clarify the academic learning concerns. Then, proceed with specific short-term and working memory items that are known to be highly related with the academic area of concern (see Table 6.3). For instance, when a third-grader is struggling significantly with mathematics learning and performance, ask about behaviors, such as finger counting, that are known to be associated with working memory weaknesses. Also, ask questions intended to identify which working memory components may be implicated. For example, when older elementary students are referred for math problems, a likely working memory hypothesis is that there is a deficit in executive working memory. After questioning the teacher about the student's ability to stay focused on relevant information, the interviewer might proceed with asking about a related cognitive process, such as fluid reasoning. In general, an effective interviewing technique is to begin with open-ended questions and progress to closed questions. When the

Table 6.3 Suggested Teacher Interview Items for Working Memory

1. What types of learning activities are most difficult for the student?
2. How well does the student remember information?
3. Does the student have difficulties memorizing information?
4. How well does the student retain information during multistep operations, such as when completing a complex arithmetic problem without paper and pencil?
5. How much repetition does the student require before learning new information?
6. How often does the student ask you to repeat directions?
7. Does the student have any difficulties with listening comprehension?
8. How well does the student stay focused on the task at hand?
9. How well can the student do two things simultaneously, such as listen and take notes?
10. Is the student slow to recall information that he or she knows?
11. Does the student have difficulty expressing ideas orally or in writing?
12. Does the student seem to be stronger in either visual or auditory learning?
13. What memory strategies, if any, have you observed the student using?
14. To what do you attribute the student's learning problems?
15. What signs have you observed that indicate the student might have problems with short-term memory?

interviewee seems unfamiliar with the behavior in question, the interviewer should provide examples.

Near the end of the initial interview, it is important to elicit the teacher's hypotheses regarding the student's learning problems. The expression of these hypotheses may illuminate other cognitive difficulties that could account for the concerns. For example, it may be the teacher's hypothesis that the student is struggling because of long-term retrieval problems. One way to encourage the expression of processing hypotheses is to ask the teacher *why* she or he thinks the student is experiencing each specific learning problem. When teachers seem uncertain about possible memory and processing hypotheses, the interviewer may need to provide more structure by asking questions, such as ''Do you think the student is having difficulty with math because he can't memorize the facts?''

Parent Interviews

Whether the evaluation is conducted in a school or a clinic, parents can be a valuable source of information regarding a student's working memory functioning. Because many parents of struggling students attempt to support their child's learning, parents are often more aware of the child's learning and memory processes than might be expected. Consequently, many of the interview items suggested for teachers in Table 6.3 can easily be adapted for parent interviews. To make the interview more relevant, the interviewer should also create items that are specific to the child's functioning in the home environment. In particular, the interviewer should ask how often the child forgets to complete daily activities. Also, the interviewer might ask how well the child remembers to complete all the steps in a multistep task. Similar to teacher interviews, parents should also be asked questions about strategy use and their hypotheses regarding the child's learning difficulties. A unique feature of parent interviews are questions about early development; in particular, interview items about early speech and language development are especially relevant to working memory. This author remembers a case where a parent volunteered that she knew her child had a memory problem when the child could not remember nursery rhymes after repeated readings. Finally, interviewers should be prepared to offer some consultation, even during the initial interview. After questioning parents about their child's memory functioning, it is often appropriate to explain to them how working memory functions relate to academic learning.

Student Interviews

Although students vary widely in their ability to self-appraise, student interviews can also be a valuable source of assessment data. Certainly, students who are middle school age or older should be directly questioned about behaviors related to working memory, and age-appropriate interview items also should be attempted with elementary students. Given the lack of structured interviews that include any items related to memory, it will be necessary for the interviewer to construct items. Of course, the

Table 6.4 Suggested Student Interview Items for Working Memory

1. Do you ever forget to do something? Can you give me an example? How often does this happen?
2. Do you sometimes forget what the teacher just talked about?
3. Do you sometimes ask the teacher to repeat directions?
4. Do you sometimes raise your hand in class and then forget what you were going to say?
5. When you study something, do you have difficulty remembering it the next day?
6. Do you plan things but then forget to do them?
7. When you are writing, do you leave out letters or words without knowing that you did?
8. Is it hard for you to listen and take notes at the same time?
9. Do you have difficulty memorizing facts?
10. Do you sometimes lose your place when reading?
11. Is it hard for you to do arithmetic in your head?
12. What do you do when you want to remember information that is difficult to remember?

main challenge with student interviews is using age-appropriate terminology and limiting the items to behaviors of which the student is aware. Questions that ask the student to make inferences about his or her memory are usually inappropriate. It is important to include items related to strategy use, especially when working memory interventions are under consideration. See Table 6.4 for some suggested student interview items.

Observations

Observation of behavior is a fundamental assessment method that should be included in every comprehensive assessment of working memory. To increase the validity of observations regarding working memory, the observer needs to become familiar with the intricacies of working memory and the behaviors that are indicative of limitations or dysfunctions. Knowledge of the relationships between academic functioning and working-memory processes (see Chapter 5) will also be particularly beneficial when conducting observations. In addition to an in-depth understanding of working memory functioning and associated behaviors, the observer should analyze the processing demands of the task the examinee is engaged in (Hale & Fiorello, 2004). Even an observer with expertise in working memory needs to be cautious about making inferences from observed behaviors, mainly due to the lack of one-to-one correspondence between behaviors and processes. That is, any one of several cognitive processes may underlie a specific observable behavior; for example, observers should not assume that failure to follow directions can be attributed to a working-memory shortcoming. Nevertheless, observations serve an indispensable purpose; they provide data that can be used to corroborate the results of testing.

Observation During Testing

Unstructured observations during the administration of standardized tests can provide valuable clinical information about the examinee's working memory strengths and weaknesses. Insights into an examinee's working memory functioning can be gained from observations during any type of testing, not just during working memory subtests. Familiarity with observable indicators of working memory processes will increase the examiner's awareness of behaviors that are noteworthy. Many of the behaviors suggested for classroom observation (see Table 6.6) should also be observable during testing. Because much of cognitive and achievement testing requires complex processing, there should be ample opportunities to observe how the examinee behaves when heavy demands are made of working memory. See Table 6.5 for testing behaviors that are indicative of an overloaded working memory.

When observing during testing, it is important to note when the behaviors occurred; do not just make general observations after all of the testing is finished. A convenient method of tracking behaviors is to record the observation alongside the item or subtest associated with it. After completing the administration, review the recorded observations and consider the demands of the task at the time each one occurred. Task analysis can provide insights into which working memory components, processes, or strategies may be deficient. Keep in mind that observations gathered during testing are informal data that should be considered, along with other data, when weighing the evidence regarding possible working memory weaknesses.

Table 6.5 Testing Behaviors Related to Working Memory

—Asking for directions to be repeated.
—Requesting supplemental materials, such as paper when the task is mental arithmetic.
—Inability to work quickly.
—Increasing frustration as the complexity of the task increases.
—Difficulty elaborating upon a response when requested to do so.
—Difficulty retrieving simple information on demand.
—Difficulty staying focused on the task at hand.
Also, observe for these indications of strategy use:
—Whispering or lip movements (subvocal rehearsal).
—Grouping or clustering information (chunking).
—Thinking aloud.

Classroom Observations

Much can be learned about a referred student's working memory characteristics through careful observation of the student in the classroom. At the very least, observations can lead to hypotheses that can be tested through later formal assessment. It is most productive to observe the referred pupil when she or he is engaged in academic areas that are challenging, because the student is less likely to display working memory related behaviors during simple, routine activities. Until the observer becomes versed in grade-appropriate working memory expectations, it is helpful to compare the subject's behavior with that of one or more random peers. For example, when a second-grader uses finger counting to complete arithmetic calculation, observe peers for the same behavior and compare frequency of occurrence. Similar to observing for other learning and behavior concerns, not much weight should be awarded to isolated behaviors. Behaviors that are indicative of underlying working memory problems should be reoccurring and similar behaviors should also be evident. For example, finger counting alone may be insufficient evidence. A stronger case for a working memory problem can be made if the learner is also having difficulties with learning math facts, retrieving known math facts, remembering partial solutions, and confusing known math facts.

The behaviors suggested in Table 6.6 are derived from observations reported in the research literature on working memory deficits and learning disabilities. The reader is encouraged to use the items in Table 6.6 with the caveat that these items have not been piloted or used in research. Therefore, these items should be used cautiously and it should be assumed that the items lack reliability and validity. For instance, many of the behaviors are also characteristic of other cognitive deficits. The suggested observations are intended for informal, clinical use only. They should mainly be used to generate hypotheses for further investigation. The information gathered during observations using these items can be used as collateral evidence for diagnostic and intervention decisions; however, the items should be given no more weight than any other informal data.

Cross-Battery and Selective Testing

Hypothesis-driven assessment frequently leads to cross-battery and selective testing. Even when a practitioner has access to a comprehensive memory scale, there are times when additional subtests should be drawn from other scales. For instance, when the primary scale measures phonological short-term memory with digits only, the practitioner may decide to replace the digits subtest with one that uses words. However, the more likely scenario for cross-battery testing is when the examiner is required to or decides to first administer a comprehensive cognitive battery. Within that cognitive battery are subtests that measure some but not all aspects of short-term and working

Table 6.6 Suggested Items for Classroom Observation of Working Memory

General Working Memory

—Classroom performance is poorer than would be predicted from standardized achievement test scores.
—Has difficulty staying focused during cognitively demanding activities but attends well when cognitive demands are minimal.
—Prefers to simplify tasks whenever possible.
—Fails to complete complex activities.
—Has difficulty keeping track of place during challenging activities.
—Has difficulty retrieving information when engaged in another processing task.
—Has difficulty associating current situation with past experience.
—Has difficulty integrating new information with prior knowledge.
—Rarely contributes to class discussions.
—Make comments such as, "I forget everything."
—Has difficulty organizing information during written expression.
—Has difficulty retaining partial solutions during mental arithmetic.
—Has difficulty memorizing and retaining facts.
—Is very slow at arithmetic computation.
—Is slow to retrieve known facts.
—Confuses known facts.

Phonological Short-Term Memory

—Has difficulty remembering multistep oral directions.
—Has difficulty restating instructions.
—Has more difficulty remembering digits than words (indicative of mathematics disability).
—Makes many counting errors.
—Has difficulty blending phonemes into words when reading.
—Has difficulty with phonetic decoding of text.
—Has difficulty with phonetic recoding (spelling).
—Has difficulty learning new vocabulary.
—Has difficulty producing multiword utterances.

Visuospatial Working Memory

—Does not notice the signs (e.g., "+") during arithmetic calculation.
—Has episodic memory lapses for the relatively recent past.
—Loses place when reading.

Verbal Working Memory

—Requires frequent reminders.
—When called on, forgets what was planning to say.
—Forgets the content of instruction.
—Has difficulty paraphrasing spoken information.
—Has difficulty comprehending syntactically complex sentences.

(*Continued*)

Table 6.6 (Continued)

—Has difficulty taking meaningful notes.
—In third grade and above, continues to finger count during arithmetic calculation.
—Rereads text when there has not been a decoding problem.
—Has difficulty remembering the first part of the sentence or paragraph when reading.
—Has difficulty detecting targets in spoken or written language, such as identifying the rhyming words in a paragraph.
—Produces only short sentences during written expression.
—Has frequent subject-verb agreement errors in written expression.
—Omits some of the content when writing a sentence.
—Repeats words when writing a sentence.

Executive Working Memory

—Answers to oral comprehension questions are off-topic or irrelevant (has difficulty inhibiting irrelevant information).
—Has difficulty switching between operations (e.g., from addition to subtraction problems).
—Has difficulty taking notes and listening at the same time.
—Inaccurately estimates memory performance before, during, or after a task.
—Does not use learning strategies or does not use them on a consistent basis.
—Prefers to use simple instead of complex learning strategies.
—Does not use the most basic strategies, such as subvocal rehearsal.
—Selects inefficient strategies during problem solving.

memory. To complete the assessment of working memory, the evaluator need only select subtests that tap the working memory components that remain untested. Informal cross-battery assessment is not new; many practitioners have mixed tests and batteries when conducting psychoeducational or neuropsychological evaluations. However, a systematic method of cross-battery cognitive assessment has only recently been advocated by Flanagan and Ortiz (2001, 2007). The approach to analyzing working memory test scores recommended in this text (see Appendix C) is an adaptation of the cross-battery model proposed by Flanagan and Ortiz (2001, 2007).

The cross-battery method involves administering a compilation of subtests from different batteries in order to systematically measure all of the areas selected for assessment. It can actually be time saving and efficient even though more than one scale is utilized. From the cross-battery perspective, evaluators should not administer an entire battery just because it contains some desired subtests. Rather, they should administer only those subtests that measure the memory components and cognitive processes selected during assessment planning. Redundant testing should be avoided; there is no need to test the same process twice. For example, if the entire WISC-IV has been administered and then is supplemented with another scale, there is no need to measure executive working memory again as it has already been adequately sampled (see Appendix A). Administering only selected factors and subtests is acceptable, unless a test's authors specifically state that the entire scale or a certain subset must be administered for the results to be valid. A battery's subtests each have their own scaled score that can be used independently to determine an individual's functioning in the primary process measured by that subtest. Nevertheless, cross-battery testing does not give practitioners permission to abandon testing standards. Also, the number of batteries involved should be restricted to three at the most. Restricting the amount of battery crossing serves to maintain the reliability and validity of the resulting scores, as well as keeping the interpretation manageable. Furthermore, there should be a psychometric or empirical basis to the selection of subtests. Appendices A and E are intended to provide guidance for subtest selection.

Cross-Battery Assessment Concerns

Although the cross-battery method is well suited for assessment of memory components and processes, a cautious interpretation of cross-battery results is necessary because of the inherent weaknesses of the method. The lack of cross-battery norms is the main concern. There are no norms for any of the numerous cross-battery "scales" that can be created. The composite, factor, and subtest scores obtained from different batteries are based on standardization samples, distributions, and norms unique to each test. Caution is particularly urged when a specific memory subprocess, such as phonological short-term memory, is assessed with subtests from more than one scale. This source of error can be reduced somewhat by using tests that were normed about the same time. When conducting an intraindividual, or *ipsative*, analysis, subtest and "clinical" factor scores representing memory components are usually compared to a cross-battery mean computed by averaging the scores of the subtests involved. This cross-battery mean has no norms and there are no statistical tables for determining significant discrepancies between it and individual factor or subtest scores. Flanagan and Ortiz (2001) recommend using a one standard deviation discrepancy as the criterion for significance, but the number of points necessary for a statistically significant difference will vary. Despite these concerns, a structured, systematic procedure to cross-battery assessment and interpretation is preferable to a completely clinical approach. For a further discussion of strengths and weaknesses of cross-battery assessment and interpretation see Flanagan and Ortiz (2001, 2007).

Selective Testing

Selective testing is associated with cross-battery assessment, as well as with the hypothesis-testing approach to assessment. The principle behind selective testing is that evaluators should select and administer only those subtests that measure what needs to be assessed. When the primary battery does not supply all of the necessary measures, portions of other batteries should be included in the assessment. Informed judgment is necessary when selecting subtests that measure specific memory components and processes. The names assigned to memory factors and subtests can be misleading. Before making a decision, examiners should examine the task and consider how it compares to the usual working memory paradigms described at the beginning of this chapter. If the task changes during the subtest, it is likely that more than one working memory component is being measured. Further guidance in subtest selection is provided in Appendices A and E.

Testing Related Processes

Cross-battery selective testing for short-term and working memory concerns should also include assessment of related cognitive processes. Some examinees who appear to be struggling because of working memory deficiencies may actually have impairments in related processes, instead of working memory itself. Processes that influence performance on working memory tasks include attention, phonological processing, processing speed, long-term retrieval, long-term storage, general executive processing, fluid reasoning, visual processing, and auditory processing (see Chapter 4 for details). The selection of cognitive processes for testing should be based on the generation and selection of hypotheses. Testing of related cognitive processes can also be conducted on a selective and cross-battery basis. For a breakdown of processing subtests by cognitive scale, see Appendix G and see Dehn (2006) for additional details.

Assessment Recommendations for Specific Disabilities

Reading Disabilities

When referral concerns include the possibility of a reading disability, the primary memory components to assess are phonological short-term memory, verbal working memory, and rapid automatic naming. In children older than 6 years of age, executive working memory should also be examined. When executive working memory is included it is advisable to use subtests that allow the discrimination of verbal working memory from executive working memory. For example, memory for sentences could be contrasted with a last word task (see Appendix A). Using reading span tests to assess verbal working memory can confound results, even in older children; listening span is preferable. The testing of related cognitive processes should include phonological processing. For children who are able to decode words, a test of reading fluency is essential, as it reveals the level of automaticity and has implications for reading comprehension.

Mathematics Disabilities

When testing for the possibility of a mathematics disability, practitioners should include the administration of short-term and working memory subtests that do not include numbers. As research has documented (Andersson & Lyxell, 2007), students with mathematics learning difficulties often have a storage-specific deficit for numerical information. Although children with mathematics disabilities may possess normal overall storage capacity in phonological short-term memory and verbal working memory, they typically have shorter spans for numbers. Their phonological and verbal spans that do not involve numbers may be normal. Also, those with mathematics disorders are known to have difficulty quickly retrieving arithmetic facts. To determine if their long-term retrieval difficulty is specific to numerical information, examiners might administer a rapid automatic naming task that does not involve numbers. Finally, the visuospatial short-term and visuospatial working memory components should be assessed, especially in younger children, when mathematics performance and visuospatial abilities are closely related.

Written Language Disabilities

First and foremost, executive working memory should be scrutinized with two or more subtests. Phonological short-term memory and verbal working memory should also be assessed, regardless of age. When there are written expression problems, several related cognitive processes should be tested, including: general executive functions, planning, and processing speed. Furthermore, an assessment of oral language development should be considered if not previously conducted.

ADHD

For children with ADHD or suspected ADHD, executive working memory should be examined in depth. Subtests that introduce interference or require dual processing are particularly relevant but they should be contrasted with executive working memory subtests that involve simple transformation (e.g. digits backward). Contrasting executive working memory with verbal working memory subtests that do not introduce interference is also helpful. The idea is to factor out the influence of attentional control in an effort to determine if poor working memory performance is a result of attentional problems or if there is also a deficiency in working memory itself.

General Guidelines for Interpreting Test Scores

Interpretation of working memory test results presents some unique challenges. These challenges arise whenever there is cross-battery testing, computation of clinical factor scores, realignment of subtests, and intrasubtest analysis. Realignment of subtests occurs when subtests are used to measure a different memory component than the test's author(s) intended. These unique procedures require more reliance on clinical judgment than is usually encouraged by experts in psychometrics. Concerns about clinical judgment are justified. Open-ended clinical interpretation is analogous to implementing interventions that are not research based. On the other hand, actuarial-based, statistical analysis also has its limitations. Ultimately, the meaning of test scores, discrepancies, and other statistical findings must be determined by the clinician (examiner). Thus, clinical judgment needs to be balanced with the actuarial-based, statistical procedures recommended by experts in psychometrics. In addition to informal analysis of test scores, clinical interpretation includes an examination of all the data, including: history, observations of the examinee, information gathered through interviews, and data collected through other informal methods.

There are several factors that enhance the quality and meaningfulness of interpretation. First and foremost, interpretation should be theory based. Regarding working memory, the practitioner has several theories to choose from, including: a general information processing theory, Baddeley's theory, Cowan's theory, or the integrated theory suggested in this text. A practitioner might even develop his or her own theory that combines elements of various research-supported theories. The benefits of theory-based interpretation are consistency across cases, a structure for integrating information, research-based support, and terminology other professionals understand. Interpretation is also facilitated when data result from planned and organized testing that addresses all of the referral concerns. The hypothesis approach to planning an assessment discussed earlier in this chapter provides a structure for interpretation. After testing is complete, the data regarding each hypothesis are considered and conclusions are drawn.

Meaningful and beneficial interpretation is focused on the individual, not the test scores. Because the goal of assessment is to better understand the individual and why that individual is experiencing problems, simply reporting scores, reviewing data, and documenting symptoms is insufficient. The meaning of assessment data will vary, depending on the individual being evaluated. Identical test scores do not have the same meaning and implications across individuals. Test scores take on meaning that is dependent on the characteristics of the examinee. For example, an average working memory score of 91 might be an individual strength for one person but an individual weakness for another. The implications of a specific memory component score depend on how that score relates to the rest of the individual's memory and cognitive profiles.

Profile Analysis

The traditional interpretative approach of analyzing subtest scores and identifying patterns of strengths and weaknesses within the individual is known as *profile analysis*. Profile analysis is conducted by computing the mean of the subtest scores involved, and then using the mean to determine which subtest scores are significantly discrepant. Generally, the .05 or .01 level of significance is used. When statistical tables are unavailable for making this determination, a one standard deviation discrepancy is usually considered significant. Subtests with significantly lower scores are interpreted as intraindividual weaknesses, whereas significantly higher subtests are individual strengths. Thus, profile analysis is an intraindividual analysis.

There is a longstanding controversy surrounding subtest profile analysis. Critics (Glutting, McDermott, & Konold, 1997) contend that profile analysis is unreliable and that the resulting profiles have no diagnostic validity. Their criticism stems mainly from the low reliability of subtests and the fact that each subtest measures more than one ability. Nevertheless, valuable information about the examinee's memory strengths and weaknesses and how these relate to overall cognitive processing will be lost if a profile analysis is not conducted. Short-term and working memory components and processes are mostly measured by subtests; there are very few factor or composite scores that are truly representative of these memory constituents. Therefore, complete analysis of separate memory elements depends on profile analysis. With profile analysis, clinicians can better understand the interactive functioning of the examinee's memory systems.

Despite the criticism of subtest profile analysis, the practice is defensible. First, the reliability and specificity of many subtests have increased as batteries have been revised. More importantly, empirical investigations and factor-analytic studies have identified the specific subprocesses measured by subtests. Similarly, brain research has documented the separate memory components that the subtests purport to measure. Moreover, profile analysis can be justified when the interpretation is grounded in theory and when the test results are corroborated by other data. When theory and research support the conclusions drawn from profile analysis, those interpretations have increased validity. Finally, profile analysis is more credible when it is conducted hand-in-hand with a normative analysis that compares the individual's performance to the distribution of test scores for his or her age group.

Weakness versus Deficit

Depending on how *weakness* is defined, nearly everyone can lay claim to at least one weakness in a cognitive process or memory component. In psychological test interpretation, it is important to adhere to the usual ground rules so that the term is not used loosely. When examining an individual's test scores there are two types of weakness: normative and ipsative. Any test score that falls below the average range (below a standard score of 90 when the mean of the distribution is 100 and the standard deviation is 15) represents a *normative weakness*. An *ipsative weakness* occurs when a subtest or factor score is significantly lower than the individual's mean for the broad domain, with a discrepancy of approximately one standard deviation an acceptable criterion for statistical significance. For example, when an individual's verbal working memory score is an 84 and the FSIQ is 99, the discrepancy is considered a significant ipsative, or intraindividual, weakness. In this text, the term *deficit* refers to a score that is both a normative *and* an ipsative weakness. Defined as such, a deficit is rare and it is indicative of an impairment because it represents poor performance relative to peers, as well as in comparison to the individual's overall abilities. When an examinee has a deficit in one or more short-term or working memory components, the deficit is most likely causing significant problems in learning or daily functioning. Although a deficit is clearly cause for concern, an ipsative or normative weakness alone may also severely interfere with the acquisition of academic skills or some aspects of daily life. Thus, all ipsative and normative weaknesses should be examined closely, with pertinent informal assessment data taken into account. Although a normative weakness may account for the difficulty an individual is experiencing, practitioners tend to ignore a specific normative weakness when all other cognitive scores are also subnormal. In such instances, it makes sense not to single out one score as worthy of attention; on the other hand, practitioners should not deny that performance associated with that ability is going to be difficult. The same guidelines can be applied to strengths. When a strength is both normative and ipsative, it is referred to as an *asset*.

Unitary versus Nonunitary Factors

When two or more subtests are aggregated and represented by one score, that score is referred to as a *factor* score. Depending on the extent of testing or the battery used, functioning of some memory components will be represented by factor scores. In such instances, it is important to examine whether the performance on the subtests comprising that factor is consistent. Inconsistent performance will produce divergent scores. When the difference between the highest and lowest subtest scores within a factor is extreme, the factor is said to be *nonunitary*. The criterion for this determination is an approximate discrepancy of 1.5 standard deviations. Accordingly, when the standard score (from a distribution with a mean of 100 and a standard deviation of 15) difference between two subtests is more than 22 points, the factor should be considered nonunitary. When a factor is nonunitary, the factor score may not represent a unitary ability. Nonunitary factors most likely occur because the underlying subprocesses measured by the subtests are not equivalent for that individual. When a factor is nonunitary, it is not appropriate to interpret the factor score as truly representing the component or process it is suppose to represent.

Nonunitary factors and the subtests they comprise should still be included in the computation of a mean cognitive processing or mean memory score. However, they should not be used in pairwise comparisons. Nonunitary factors should be interpreted, especially when they represent a weakness or a deficit, but the interpretation should be cautious. When a nonunitary factor occurs, the clinician should examine the subtests involved and generate hypotheses that account for the discrepancy, taking testing behaviors and other evaluation data into account. For instance, an examinee may perform poorly on one short-term memory subtest but not the other. Examining the content of the two subtests and the narrow abilities that they measure may reveal an explanation. For example, the low-score subtest might have involved the repetition of numbers while the other involved the repetition of words. One hypothesis might be that the examinee lacks number facility. Such hypotheses should be investigated through further testing and assessment. Another approach to dealing with nonunitary factors is to administer an additional subtest that taps another narrow ability within the broad factor. This approach will provide a broader sampling of the memory factor.

Base Rates

Base rate is another standardized test statistic usually considered when conducting profile analysis and interpreting test scores. Test manuals that provide critical values for significance testing also provide base rate tables that report the frequency of intraindividual differences in the standardization sample. The idea of using base rate information is to determine just how infrequent or rare a given discrepancy is. For example, a difference of 12 points between working memory and the mean processing score might be statistically significant but yet occur in approximately 25% of the population. Given that this discrepancy is not that uncommon, the practical significance of the finding is questionable. Many test authors and experts in psychometrics recommend interpreting significant ipsative strengths and weaknesses only when the base rate is lower than 15%. However, strictly following such an actuarial rule and not using clinical judgment may result in not identifying a weakness that is actually causing impairment. Just because a given ipsative weakness occurs in 15% or more of the population does not mean that it is not causing a serious memory problem for the individual. A relatively common significant weakness may still be an impairment for some individuals. Especially in cases where an actual *deficit* has been identified, base rates should be ignored. This is because *all* deficits can be considered unusual and infrequent (Naglieri, 1999). When cross-battery analysis is conducted, base rates for ipsative discrepancies are generally unavailable. Under such circumstances, clinicians should carefully consider all potential weaknesses that meet the criterion of one standard deviation below the mean or composite score used to represent the mean. When other assessment data support the existence of a weakness that is impairing functioning, the weakness should be regarded as being important and having practical significance. In general, each profile should be carefully examined and all assessment data taken into account. Rules regarding base rates should not take precedence.

Hypothesis Testing

After profile analysis is complete, the evaluator should weigh all of the evidence for and against each of the a priori hypotheses pertaining to memory components and related processes. The evidence consists of test scores and relevant data collected through other methods. The *Working-Memory Interpretative Summary* found in Appendix D is offered as a structure for summarizing the data and reaching a conclusion about each hypothesis. Even when the data clearly support a hypothesis, oral and written conclusions and generalizations should be stated cautiously. In many cases, the data will be inconclusive and the evaluator will not be able to reach a decision about the hypothesis. In instances where results are inconclusive or unexpected and answers are still sought, the evaluator should generate a posteriori hypotheses and investigate these through further assessment. As discussed earlier in this chapter, the clinician should be on guard against hypothesis confirmation bias.

Processing Strengths

If only suspected areas of weakness are tested, a profile analysis is unlikely to reveal any strengths or weaknesses because the individual's mean score is based mainly on weaknesses. Because assessments originate with referral concerns, areas of potential strengths and assets are often not tested, especially when the assessment is based on hypotheses that account for the referral concerns. Fortunately, not all hypothesized weaknesses turn out to be such; many memory components hypothesized to be deficient are discovered to be individual strengths in the normal range of functioning. Such findings restore some balance to the profile analysis mean. Another option of comparing memory component scores to a more valid estimate of overall cognitive ability is to use a FSIQ or similar composite score (see the next section). When interpreting test results, clinicians should emphasize strengths and assets as much as weaknesses and deficits because incorporating strengths into interventions may increase the probability of success.

Analysis of Working Memory Test Scores

The goal of conducting a profile analysis of an individual's working memory test scores is to gain a better understanding of the individual's memory strengths and weaknesses. Determination of strengths and weaknesses should take statistical guidelines into account but should not be bound by those guidelines. Scores and discrepancies between scores only provide indications of levels of functioning and the relative degree of strengths or weaknesses. Deciding that an individual has a deficit that is impairing functioning must ultimately be based on clinical judgment. To rule out the possibility of a deficit because the discrepancy pointing toward a deficit is barely significant or has a relatively high base rate in the population is poor clinical judgment. All relevant assessment data and information should be taken into account before such decisions are made because the same set of scores and the same degrees of discrepancy mean different things for different individuals. Consequently, the profile analysis procedure suggested in this section combines an actuarial and a clinical approach, with clinical judgment taking precedence.

The *Working Memory Analysis Worksheet* found in Appendix C provides consistent structure and guidelines for analyzing nearly every set of test scores. When a cross-battery assessment has been completed, the worksheet is ideal for combining scores into one comprehensive analysis, as opposed to a test-by-test analysis that leaves the evaluator and the clients wondering what it all means. The versatile analysis worksheet is also ideal for analyzing scores resulting from only one battery. Because the memory subtests and factors of many batteries do not align well with theoretical models, there is often a need to restructure the subtests and compute clinical factor scores. The primary purpose of the worksheet is to identify and compare the six main memory components of the Integrated Model of Working Memory, but additional related processes, such as processing speed, can also be incorporated. The procedures for completing the *Working Memory Analysis Worksheet* are described in the following, using Joey's completed worksheet in Table 6.7 as an example (an abbreviated set of instructions is also found on the blank worksheet in Appendix C).

1. *In the lower cells of the first column, include any related cognitive processes that were tested.* The six components of the Integrated Model of Working Memory are preprinted in the first column but the worksheet can still be used if one or more of the memory components is not tested. In the case of Joey, a priori hypotheses led to the inclusion of phonemic awareness, fluid reasoning, and processing speed (see Table 6.7).
2. *Write the name of the test battery in the second column.* This is the name of the scale or battery from which the subtests measuring that component are drawn. In instances where subtests from different batteries are used to measure the same component, put both batteries and related information in the same row. For an example of this, see the analysis of scores in Table 8.2.
3. *In the third column write the name(s) of the subtest(s) used to measure that memory component or process.* Classification of the subtests is found in Appendices A and E and in subsequent chapters. In instances where a factor score can be used as is, enter the battery's name for the factor.
4. *For each subtest or factor, calculate standard scores and enter these in the fourth column.* Scaled and standard scores found in the test's manual should be used. However, all subtest scores with a mean of 10 and standard deviation of 3 will first need to be transformed to scores that have a mean of 100 and a standard deviation of 15. Use the table in Appendix F to transform the subtest scores.
5. *Compute the mean of the subtest scores for each memory component and cognitive process and enter it in the fifth column.* For components that are assessed with only one subtest, simply carry over that subtest score from the fourth column. When two or more subtests are used, compute the mean, rounding to the nearest whole number. The mean of the subtests can be considered a *clinical factor* score that represents the broad functioning of that component or process. The computation of clinical factors is necessary whenever subtests from two different batteries are used to measure the same memory component or when subtests from the same scales are aligned differently than that test's structure. For example, the WJ III COG has a Long-Term Retrieval cluster but a clinical long-term retrieval cluster has been computed for Joey because Rapid Picture Naming is not one of the subtests that comprise the WJ III's Long-Term Retrieval factor. In Joey's case, the two subtests used to

Table 6.7 Working Memory Analysis Worksheet–Completed Example

Examinee's Name: Case Study of "Joey" DOB: ——— Age: 9 Grade: 4 Dates of Testing: ———

Memory Component	Battery Name	Subtest/ Factor Name	Subtest Score	Component Mean	Composite or Mean	Difference	Normative S or W	Ipsative S or W	Deficit or Asset
Phonological STM	WJ III COG	Memory for Words	106	106	WJ III GIA–106	0	Avg.	—	—
Visuospatial STM	WISC-IV Integrated	Spatial Span Forward	85	85	106	−21	W	W	Deficit
Verbal WM	WJ III COG	Auditory Working Memory	101	101	106	−5	Avg.	—	—
Visuospatial WM	WISC-IV Integrated	Spatial Span Backward	95	95	106	−11	Avg.	—	—
Executive WM	WJ III COG	Numbers Reversed Auditory Working Memory	77 101	89	106	−17	W	W	Deficit
Long-Term Retrieval	WJ III COG	Rapid Picture Naming Retrieval Fluency	95 87	91	106	−15	Avg.	W	—
Phonemic Awareness	WJ III COG	Phonemic Awareness	128	128	106	+22	S	S	Asset
Fluid Reasoning	WJ III COG	Fluid Reasoning	95	95	106	−11	Avg.	—	—
Processing Speed	WJ III COG	Processing Speed	107	107	106	+1	Avg.	—	—

(Continued)

Table 6.7 Working Memory Analysis Worksheet–Completed Example (Continued)

Subtest or Clinical Factor Score	Subtest or Clinical Factor Score	Discrepancy	Significant: Y/N
Phonological STM (106)	Phonemic Awareness (128)	22	Y
Phonological STM (106)	Visuospatial STM (85)	21	Y
Verbal WM (104)	Executive WM (89)	15	Y
Executive WM (89)	Long-Term Retrieval (91)	2	N
Long-Term Retrieval (91)	Processing Speed (107)	16	Y

Directions: (1) Convert all subtest scores to standard scores with a mean of 100 and an SD of 15. (2) For each component, compute the mean of the subtest scores and round to the nearest whole number. (3) Enter a cognitive composite, such as a FSIQ, or compute the mean of all available memory components. (4) Subtract the composite or mean from each component mean and enter amount in Difference column. (5) Indicate whether the component mean is a normative weakness or strength (90–109 is average). (6) Using a criterion of 12 points, determine intraindividual strengths and weaknesses. (7) Determine deficits and assets. A deficit is both a normative and intraindividual weakness; an asset is both a normative and intraindividual strength. (8) Determine which factors are nonunitary. Factors are nonunitary when the two subtests involved are significantly different or when the range between the highest and lowest subtest scores exceeds 1.5 standard deviations. Nonunitary factors should be interpreted cautiously and should not be used in pairwise comparisons. (9) Compare logical pairs of components, using a 15-point difference as an indication of a significant discrepancy.

evaluate long-term retrieval average out at 91. Not just any subtests should be paired to produce clinical factors. It should be apparent that the tasks involved are measuring that memory component to a significant degree. The classification of subtests in this text is based on which research paradigms have consistently been used to measure specific memory components, and it is based on the functions of the components in the Integrated Model (see Chapter 31). Classifications of subtests from numerous batteries are found in Chapters 7 and 8, as well as in Appendices A and E. Clinical factor scores need to be interpreted more cautiously than the regular factor scores provided by batteries.

6. *Enter a cognitive composite or mean in the sixth column.* In general, this score should represent overall cognitive processing ability, overall memory functioning, or a combination of memory functioning and related cognitive processes. The first option is to use an IQ score or similar composite from an intellectual or cognitive scale. In Joey's case, the General Intellectual Ability (GIA) score is used. An IQ or cognitive composite is appropriate for memory profile analysis because of the high correlations between IQ and working memory. In the absence of an IQ or cognitive composite score, the next option is to compute an overall processing mean from the set of memory and other cognitive processes that are being analyzed on the Worksheet. To arrive at this mean, average the scores found in the "Component Mean" column, rounding to the nearest whole number. When assessment has been restricted to suspected areas of weakness or when all of the memory scores are low, the mean of the scores will also be low, resulting in fewer discrepancies. In such instances, it is better to use an IQ or other global cognitive score, even if that score was obtained during a previous evaluation. Another option is to use a global memory score obtained from a memory battery or compute a memory mean based only on memory component scores. Each cell in the sixth column will have the same value. When reporting results later on, the practitioner should always identify the source of the value entered in this column so that colleagues and clients know what the memory components were compared with.
7. *Calculate and enter difference scores in the "Difference" column*. Subtract the composite or mean from each component score and enter the difference with a + or −. For example, Joey's composite of 106 was subtracted from his visuospatial short-term memory score of 85, resulting in a difference score of −21.
8. *Determine normative strengths and weaknesses*. In the "Normative S or W" column, enter an *S* (strength) for component means that are above 109. Enter a *W* (weakness) for component means that are below 90. For scores in the average range (90 to 109), simply put an *A* for average. For the sake of consistency, classify all scores on this basis, even scores from tests that describe the average range as 85–115.

9. *Determine intraindividual strengths and weaknesses and enter an S or W in the appropriate cells of the "Ipsative S or W" column*. Enter a *W* for weakness when the memory component mean is 12 or more points lower than the individual's cognitive composite or mean, and enter an *S* for strength when the memory component mean is 12 or more points higher than the individual's cognitive composite or mean. For example, Joey has an ipsative weakness in executive working memory because he has a negative discrepancy of 17 points. Whenever clinical factors have been computed or only part of a test has been administered, tables for determining significance and base rates are unavailable. In these instances, a discrepancy of 12 points or more (see the "Difference" column) is indicative of a significant difference. In a cross-battery analysis with clinical factor scores and a cognitive composite drawn from another scale, it is highly unlikely that a 12-point discrepancy is actually statistically significant at the .05 level. A 15-point discrepancy, roughly a difference of one standard deviation or more, is more likely to achieve a satisfactory level of significance. However, very few relative strengths or weaknesses will be identified if a criterion of 15 points is applied. Working memory weaknesses and deficits that are actually causing impairments in the individual may be missed if a 15-point discrepancy is required. When 12 points is used as the cutoff, it is very important that other assessment data corroborate the existence of the specific intraindividual weakness.
10. *Determine processing deficits and assets and enter in the last column*. Enter "Deficit" for components that have both a normative and ipsative weakness. Enter "Asset" for components that have both a normative and ipsative strength.
11. *Determine whether each component is unitary*. Using standard scores that have a mean of 100 and a standard deviation of 15, compare the lowest and highest subtest scores within each component or cognitive process. When the difference is greater than 22 points (greater than 1.5 standard deviations), consider the factor to be nonunitary. Nonunitary factors should be included in the profile analysis, but they should be interpreted cautiously. Practitioners should generate hypotheses to account for the disparity within the nonunitary memory component. Additional testing will often clarify the reasons for the discrepancy. Nonunitary factors should not be included in pairwise comparisons.
12. *Conduct pairwise comparisons in the lower portion of the worksheet*. Compare the scores of logically related components and processes, such as phonological short-term memory versus visuospatial short-term memory or phonological short-term memory versus phonological processing (see Tables 6.8 and 6.9). For example, Joey has a significant 16-point discrepancy between his Long-Term Retrieval score of 91 and his Processing Speed score of 107. When both components are represented by single subtest scores from the same

Table 6.8 Pairwise Comparisons of Short-Term and Working-Memory Components

—Phonological short-term memory versus visuospatial short-term memory
—Phonological short-term memory versus verbal working memory
—Phonological short-term memory versus visuospatial working memory
—Phonological short-term memory versus executive working memory
—Phonological short-term memory versus long-term retrieval
—Visuospatial short-term memory versus verbal working memory
—Visuospatial short-term memory versus visuospatial working memory
—Visuospatial short-term memory versus executive working memory
—Visuospatial short-term memory versus long-term retrieval
—Verbal working memory versus visuospatial working memory
—Verbal working memory versus executive working memory
—Verbal working memory versus long-term retrieval
—Visuospatial working memory versus executive working memory
—Visuospatial working memory versus long-term retrieval
—Executive working memory versus long-term retrieval

Table 6.9 Suggested Pairwise Comparisons of Memory Components and Related Processing Scores

—Executive working memory versus attention
—Executive working memory versus executive processing
—Executive working memory versus fluid reasoning
—Executive working memory versus general intelligence
—Executive working memory versus processing speed
—Executive working memory versus rapid automatic naming
—Executive working memory versus retrieval fluency
—Verbal working memory versus verbal ability
—Visuospatial working memory versus visual processing
—Phonological short-term memory versus auditory processing
—Phonological short-term memory versus phonological processing
—Phonological short-term memory versus processing speed
—Phonological short-term memory versus successive processing
—Long-term retrieval versus verbal ability

battery, check the battery's manual for statistical tables for determining pairwise discrepancies. When such tables are unavailable or when clinical factors have been calculated, use a 15-point discrepancy as indicative of significance and infrequency. (With pairs, a higher critical value is necessary than when comparing scores to a mean.) Discrepancies of 20 points or greater are very likely to be significant. Discrepancies in the 15- to 19-point range need more support from other evaluation data before they are interpreted.

Interpretation of Working Memory Assessment Results

The meaning of test scores and other data is more easily discerned when the practitioner has expertise in the domains assessed. Most professionals who are assessing working memory have extensive experience and strong background knowledge in cognitive development, learning disabilities, standardized testing, and related domains. Nonetheless, they are encouraged to pursue more reading on current research in working memory. Expertise is also acquired through experience. Once practitioners begin to test working memory on a regular basis, they will begin to recognize common cognitive profiles associated with working memory difficulties and associated learning challenges. The remainder of this chapter offers suggestions intended to facilitate interpretation of working memory weaknesses and deficits. (See Chapter 10 for advice on how to present results orally and how to explain them in evaluation reports.)

The general strategy for interpreting working memory functioning is to pull it apart and examine the components and subprocesses. Current testing technology, as well as theory and research, gives us the ability and rationale for doing so. Restricting interpretation to global working memory leaves many questions unanswered and can result in a misunderstanding of what is actually happening. Most individuals are likely to possess strengths and weaknesses within working memory. Identifying these will increase the evaluator's understanding of the individual (and hopefully the individual's self-awareness as well), resulting in the ability to select interventions, accommodations, and strategies most likely to improve the individual's working memory performance. For example, a blanket interpretative statement, such as "Joey has a deficit in working memory" (see Table 6.7), may mask the fact that Joey's phonological short-term and verbal working memory are just fine. His problems lie mainly with executive working memory. In other situations, impairments erroneously attributed to working memory may actually be due to phonological short-term memory. Furthermore, each aspect of working memory has unique relationships with other processes and functions. Of course, the risk of pulling cognitive processes apart and examining the subprocess is that an understanding of the system is lost. Thus, evaluators need to ultimately put the pieces back together, integrating the information and explaining how the components and processes relate to one another. The evaluator may very well conclude with the global statement that "the examinee has a working memory deficit"; but by the end of the interpretation, all listeners or readers should understand which memory components are involved.

Other general strategies for interpreting working-memory assessment results include: (a) checking for similarities by domain (e.g., if visuospatial working memory is high, visual processing is also likely to be high); (b) determining areas of expertise (e.g., a student with exceptionally high mathematics skills may perform exceptionally well on any task involving numbers); (c) contrasting performance with related cognitive functions, such as processing speed (see Table 6.9); (d) taking into account related influences (see Table 6.10); (e) checking for consistency between memory components and closely related academics (see Table 5.2); (f) considering the extent of strategy knowledge and usage—in particular, the automaticity of strategies; and (g) contrasting performance with overall cognitive ability, such as an IQ score.

Pairwise Comparisons

The most direct and primary method of examining memory components and disentangling influences is to contrast test scores that differentiate between different components and processes. For comprehensive lists of suggested pairings, see Tables 6.8 and 6.9. Some prominent pairings and the explanations and implications of discrepancies are discussed in the following. (There should be at least a one standard deviation discrepancy between paired components before the difference is considered significant.)

Table 6.10 Memory Components and Related Cognitive Processes and Influences

Phonological short-term memory: speech development, speech rate, phonological processing, phonemic awareness, sequential processing, processing speed, rote learning, auditory processing, and auditory discrimination. For tasks involving numbers—arithmetic skills and number facility.

Verbal working memory: verbal abilities, language development, reading fluency, long-term semantic memory, long-term retrieval speed, and fluid reasoning.

Visuospatial short-term memory and visuospatial working memory: visual processing, spatial ability, simultaneous processing, attention, verbal recoding, and executive processing.

Executive working memory: general executive processes, attention, metacognition, metamemory, planning, and fluid reasoning.

Short-Term Memory Components versus Working Memory Components

When short-term memory scores are higher, the examinee tends to do well with simple, rote tasks but struggles with more complex cognitive and learning activities. Short-term memory can function adequately without normal working memory but the converse is less likely because impairments in short-term memory place more demands on working memory as it tries to compensate for impaired short-term functions. Also, when the load on working memory increases, short-term span tends to decline as much as 30% (Duff & Logie, 2001). Thus, it is normal for the digits backward span to be lower than the digits forward span. When individuals do perform significantly better on more complex working memory tasks than they do on simple-span tasks, it may be that they find the tasks more engaging and thus focus their attention better than they do on the simple-span activities.

Short-Term Memory or Working Memory versus Long-Term Memory

When long-term storage is weaker than short-term and/or working-memory components, the initial and common hypothesis is that there is a problem with long-term retention of information. However, the problem may be at the encoding level. Just because short-term retention is stronger does not mean that working memory successfully encoded the information into long-term memory. Support for such a hypothesis is indicated when verbal and/or executive working memory is weaker than short-term memory. For example, effective semantic encoding requires associating new information with existing long-term schemas. When the converse—impaired short-term or working memory but normal long-term memory—occurs, it seems paradoxical at first. The unexpected profile might be attributable to many opportunities to learn under "low load" conditions (see Chapter 9), or the deficit might be limited to one or two specific short-term or working memory processes that have less of an impact on long-term encoding.

Phonological Short-Term Memory versus Visuospatial Short-Term Memory

Discrepancies between these two components, and similarly between verbal working memory and visuospatial working memory, primarily reflect relative differences in the development and strength of the two modalities: visuospatial and auditory/verbal. The difference between these two components may change with age. In particular, visuospatial performance may improve as verbal recoding develops and as executive working memory is able to provide more coordination between visuospatial and verbal processes.

Phonological Short-Term Memory versus Verbal Working Memory

This comparison is crucial whenever an examinee is experiencing academic learning problems. When phonological short-term memory is lower, it is frequently because delayed phonological processing or slow speech rate is influencing auditory short-term retention, and there is little that strategies can do to improve performance. Conversely, verbal working memory is more amenable to strategies, is enhanced by well-developed fluid reasoning, and is directly supported by long-term memory retrieval. It could be that the individual has diminished short-term memory capacity, but in such cases it is unlikely that verbal working memory would be much higher, as short-term and working memory resources are interdependent. Conversely, when phonological short-term memory is higher than verbal working memory, it implicates shortcomings in the higher level processes associated with verbal working memory, whereas the processes associated with short-term memory are intact. Such individuals may do well in simple activities, such as spelling, but struggle with complex tasks like written expression.

Verbal Working Memory versus Executive Working Memory

A discrepancy in favor of verbal working memory implicates poor executive management of working memory resources relative to well-developed verbal abilities. Other hypotheses that might account for this difference include: (a) there is high capacity for storing and encoding verbal information but not when it needs to be transformed; (b) strong long-term memory structures provide more support for verbal working memory than executive working memory; and (c) the individual has difficulty dealing with interference but otherwise has good ability to manipulate and store verbal information. This profile is likely to be observed in an intelligent child who has ADHD. The reverse discrepancy indicates well-developed executive coordination of memory tasks and a strong ability to inhibit irrelevant information in contrast to weak verbal abilities, poor verbal encoding, and/or weak long-term semantic memory structures.

Educational Implications and Assessment Recommendations for Specific Deficits

Because of the complex interrelationships various working memory components and processes have with other cognitive processes and academic learning, it is imperative to consider how specific short-term and working memory processes may be interacting with other cognitive processes and how they may be impacting aspects of academic learning (see Table 5.2). It is not enough to thoroughly explain working memory performance; the evaluator must also accept responsibility for discussing the implications for learning and daily functioning. When short-term and working memory weaknesses and deficits have been identified, the evaluator needs to closely examine performance in related domains (see Table 6.10). In some instances, further assessment may be necessary. Regarding the assessment recommendations listed in the following, it is assumed that all other short-term and working memory components have already been tested. Identification of the deficits discussed in this section results from a profile analysis procedure (see Appendix C), not from pairwise comparisons. See Chapter 9 for intervention recommendations for each area of deficiency.

Deficit in Phonological Short-Term Memory

First, do not base this determination on a digits-only subtest; a word or nonwords subtest should be included. The primary related areas to examine are phonological processing and basic reading decoding skills. When digits are low, counting and arithmetic skills should also be scrutinized. In addition to the acquisition of basic reading and arithmetic skills, deficient phonological short-term memory capacity is also likely to impact: rote learning, spelling, vocabulary development, and speech development. It is essential that additional testing include: phonological processing or phonemic awareness; sequential processing; processing speed; auditory processing; and verbal abilities. It is also important to informally evaluate the examinee's development and use of rehearsal and other basic strategies.

Deficit in Visuospatial Short-Term Memory

Visuospatial memory has its strongest relationship with visual processing; thus, poor performance in this domain necessitates assessment and examination of visual processing ability. In general, visual processing has low correlations with all types of academic learning, and deficits in visuospatial memory do not have the consistent relationships with learning domains that phonological and verbal deficits have. Nonetheless, in young elementary school children, arithmetic skills should be assessed. Also, it would be helpful to determine whether the child has developed the strategy of recoding visuospatial stimuli into verbal information.

Deficit in Verbal Working Memory

A specific deficit in verbal working memory is indicated by poor performance recalling sentences or stories. The size of this sentence-based working memory is related to reading comprehension (Daneman & Carpenter, 1980), as the retention of sentences that have already been read is a critical feature of reading comprehension. Other activities that are likely to be difficult are: taking notes while listening; remembering multistep directions; relating new information to prior knowledge; verbal fluid reasoning; written expression; oral language comprehension; and oral expression, especially paraphrasing and summarizing. Areas that should be assessed include: reading comprehension; reading fluency; long-term retention of information; long-term retrieval; fluid reasoning; and language development.

Deficit in Visuospatial Working Memory

Similar to a visuospatial short-term deficit, general visual processing ability is the primary consideration. As visuospatial working memory depends on the manipulation of mental images, a weakness or deficit indicates that visual mnemonics may not be beneficial and that verbal recoding of visuospatial information will be important. Also, visuospatial working memory depends heavily on executive working memory; thus, an executive deficiency may impact visuospatial performance.

Deficit in Executive Working Memory

In general, these individuals will have more difficulty with complex activities that require the coordination of memory systems, such as a task that requires both visual and auditory processing. The more severe the deficit, the broader the range of difficulties the individual will experience; for example, more than one area of academics may be impacted. Upon discovering such a deficit, it is incumbent on the examiner to use all available means (unfortunately, there are no standardized tests for this purpose) to assess the extent of strategy use, as executive working memory efficiency is partly a function of strategy development and usage. Inefficient executive working memory processing will have a detrimental impact on other working memory functions but not necessarily on short-term components. When executive working memory performance is poor, additional testing should be conducted in: attention, fluid reasoning, broad executive processes, reading comprehension, written expression, and mathematics reasoning.

The Use of Nonstandardized Working Memory Measures

In applied settings, there are times when informally constructed working memory tests are appropriate or even necessary. Such instances include group or individual screenings, progress monitoring during interventions, or times when standardized working memory tests are unavailable. In such instances, informally constructed tests can provide useful data for decision-making purposes. To prepare an informal measure, first review the frequently used paradigms described near the beginning of this chapter and select one or more appropriate methods. Once the content and procedure have been selected, construction of progressively more difficult items is relatively easy. In constructing span tasks, simply add one more item to be remembered at each level, and allow about three trials at that level of difficulty. When testing, stop after the examinee fails all three trials at one level and record the longest span of items the examinee responded to correctly. Keep in mind that a task will not measure the full extent of working memory capacity unless it presents enough of a challenge. Overly easy tasks or activities for which the individual has attained automated processing are not challenging measures of working memory. For example, to assess executive working memory, construct a dual-task procedure in which the examinee is required to remember items sequentially while a brief processing task is interspersed between items. Complex tasks that require additional processing will result in shorter spans than simple-span tasks. After completing a set of test items, pilot the informal measure by sampling a few students. Piloting will provide feedback on clarity of directions and ease of administration, as well as assess the difficulty level of the test. If broad use of the test is intended, collected data can even be used to establish local norms.

Preschool Screening

Preschool screening is one situation in particular in which informally constructed tests, adapted standardized tests, or local norms may be beneficial. According to the work of Gathercole and Pickering (2000b) in the United Kingdom, screening the working memory of all children entering school can provide valuable information about who is at risk for learning problems. Gathercole and Pickering modified a backwards digit recall task to make it age-appropriate for 4 year olds. Similar procedures that can be efficiently administered and scored might also be created from the paradigms described in this chapter.

Progress Monitoring

Informally constructed measures of working memory may also be appropriate for measuring progress during interventions involving working memory. Similar to measuring academic skills progress with formal tests, standardized scores from working memory tests are often poor indicators of change. Furthermore, the repeated administration of standardized tests will create practice effects that will invalidate the results. Informally constructed working memory measures are similar to curriculum-based measurement procedures in that the content should relate to the goals for the intervention. In such applications, only raw data are necessary and it is left to the practitioner to assess the clinical or practical value of the improvement, based on his or her knowledge of working memory capacity and development. Also, standardized scores are unnecessary for intervention purposes, as progress is usually determined by comparing the individual's baseline data to subsequent data. For example, when trying to improve the complex working memory span of an 8-year-old, the trainer might use an originally created counting span as the dependent measure. During baseline testing, it is determined that the 8-year-old has a counting span of three. As the intervention proceeds, the learner's counting span might increase to four. The relevant question at that point is the extent to which the data represent an improvement that will translate into better academic performance or improved academic skills. Perhaps, in the future, helpful *benchmarks* for working memory will be created to assist educators and related professionals in making judgments about the development of children's working memory abilities.

Key Points

1. Processing deficits, particularly deficits in working memory, are one of the main reasons students with disabilities are unable to respond successfully to regular education interventions.
2. When short-term and working memory components are being assessed, closely related cognitive processes should also be tested.

3. Short-term and working memory subtests evolved out of experimental research paradigms, not all of which are valid measures of real-world working memory functioning.
4. The memory subtests found in most cognitive and memory scales are often incorrectly classified. For example, a subtest purported to measure phonological short-term memory may be primarily measuring executive working memory.
5. Simple-span activities measure short-term memory and complex-span tasks measure working memory. Simple-span tasks require little more than passive retention of sequential items. Complex-span tasks introduce a secondary processing task that interferes with short-term retention.
6. Working memory assessment should begin with the generation and selection of hypotheses that account for the referral concerns. Hypotheses point toward memory components and cognitive processes that are likely to be deficient. Subtests that allow the testing of these hypotheses are then selected.
7. The cross-battery approach can be an efficient method of organizing and conducting working memory assessment. However, because cross-battery analysis is lacking the usual psychometric statistics, interpretation of results should be cautious.
8. Although clinical judgment is necessary, it needs to be balanced with acceptable actuarial-based, statistical procedures.
9. Profile analysis is a justifiable method of determining an individual's memory strengths and weaknesses.
10. A normative weakness is indicated when a score is below a standard score of 90, and an ipsative weakness is indicated when there is at least a 12-point discrepancy between the score and a relevant broad domain or IQ score. A deficit exists when both a normative and an ipsative weakness are present. Deficits are rare and are indicative of an impairment. Whenever scores meet the criteria for a deficit, the usual rules regarding base rates and nonunitary factors can be suspended.
11. The *Working Memory Analysis Worksheet* (Appendix C) can be used to analyze scores from one battery or scores obtained during a cross-battery assessment.
12. When interpreting working memory results, clinicians should first analyze the separate memory components and processes. Then, integrate the information so that the functioning of the overall working memory system can be understood.