Chapter 6

WECHSLER MEMORY SCALES

The Wechsler memory scales are one individually administered, composite batteries designed to better understand various components of a patient's memory. Now in its third edition (WMS-III), it has been co-normed with the WAIS-III. Another major feature is that it provides a full range of memory functioning and has been carefully designed according to current theories of memory. As a result of these features, it is typically considered to be a core component of any thorough cognitive assessment, which is reflected in its being ranked as the ninth most frequently used test by clinical psychologists (and third by neuropsychologists; Camara et al., 2000).

Memory complaints are extremely prevalent among client populations. They are associated with depression, anxiety, schizophrenia, head injuries, stroke, learning disabilities, and neurotoxic exposure. For example, the impact of alcohol and other drugs on memory might need to be carefully evaluated. Occupational contexts might need to similarly evaluate the impact of workers who have been exposed to industrial agents (lead, mercury, organic solvents) that can potentially result in impaired memory function. The increasingly aging population means that distinguishing normal memory loss from the early expression of dementia will become progressively more important. One crucial differential diagnosis is to distinguish between pseudodementia resulting from depression and Alzheimer's disease. As various drugs are developed for treating cognitive difficulties, it will also become increasingly important to monitor client improvement with a particular emphasis on memory functions. This array of symptoms suggests a developmental perspective in that children are most likely to experience memory complaints related to learning disabilities, adults typically experience difficulties because of neurotoxic exposure or head injuries, and older populations have memory problems related to dementing conditions.

Many of the early conceptualizations of memory considered it a unitary process. From a practical assessment perspective, it was not necessary to have a composite battery that assessed various components of memory. In contrast, more recent conceptualizations consider memory to have various components. One major distinction is between short-term and long-term memory (sometimes described as primary and secondary memory storage, respectively). For memory to be effectively stored, there also needs to be some active engagement on the part of the person. Thus, "working memory" was conceptualized as containing an executive component that initiated, monitored, and evaluated information. It also included an attentional component that had a limited capacity. A further well-supported distinction is between memory that is conscious and reflected in verbal reports of facts, events, and experiences (*declarative, explicit,* or *episodic* memory) versus memory that is more unconscious and measured implicitly by changes in performance (*procedural,* or *implicit* memory). Finally, memory can involve various sensory components, particularly visual and auditory modes of processing.

HISTORY AND DEVELOPMENT

In some ways, the development of the Wechsler memory scales have paralleled the development of knowledge on memory. Each of the three editions has increasingly incorporated advances in the theoretical understanding of memory. The original Wechsler Memory Scale (WMS; Wechsler, 1945) reflected the earlier nonspecific conceptualizations of memory. It was composed of brief procedures on memory for number sequences, text, simple visual designs, and paired words. The advantage of using a variety of procedures was that a client might have intact memory for visual information but not auditory information or vice versa. Despite the fact that the early WMS procedures could be *logically* divided into visuospatial versus auditory tasks, the overall scoring was a composite Memory Quotient that, similar to the Wechsler intelligence scale IQs, had a mean of 100 and a standard deviation of 15. This was extremely valuable information for clinicians because they could easily compare a client's Intelligence Quotient with their Memory Quotient. Any large discrepancy could be investigated further to understand the underlying reasons for such a discrepancy. The WMS was also quite popular as it was a relatively brief procedure, typically taking about 15 minutes to complete. Because retesting a client would be likely to result in practice effects, it had the further advantage of having a parallel form. As a result of these advantages, it became an ubiquitous procedure among clinicians.

The WMS had surprising longevity given a formal new version did not become available until 1987 (a 42-year interval). The WMS was limited, however, because it included unsophisticated methods of scoring the various procedures. In addition, the algorithms to determine the Memory Quotient were overly simple because they did not consider a sufficient number of client variables. The norms were derived from a small sample of 200 patients between ages 25 and 50 at Bellevue Hospital. Scores for either older or younger persons were extrapolated from this sample but were not based on actual participants. In addition, the alternate form was rarely used, and the research supporting it was quite limited. Finally, it did not reflect advances in knowledge related to memory processes.

One early attempt to correct for the deficiencies of the WMS was Russell's (1975, 1988) adaptation in which he administered two of the subtests (Logical Memory and Visual Reproduction) in an immediate format combined with a delay of 30 minutes. This allowed comparisons to be made between short-term and long-term memory. Research on Russell's WMS supported the predicted difference between left (relatively lowered auditory recall based on Logical Memory) and right (relatively lowered visual reproduction based on Visual Reproduction) hemisphere lesions. Despite these advantages, the psychometrics were weak and it was poorly standardized. Unfortunately, it was titled *Wechsler Memory Scale-Revised* (WMS-R). This potentially could create confusion because The Psychological Corporation developed a full revision of the WMS that was also titled *Wechsler Memory Scale-Revised.* Subsequent publications have attempted to clarify the two versions by referring to them as either *Russell's WMS-R* or the *WMS-R.*

The 1987 revision (Wechsler Memory Scales-Revised or WMS-R) was a significant improvement over the WMS. It had age-related norms for nine different age groups ranging between 16 and 17 years for the youngest group and 70 to 74 years for the oldest group. However, three of the age groups (18 to 19 years, 25 to 34 years, and 45 to 54 years) were not based on actual subjects but rather were estimated based on the gradual monotonic decrease in performance. The standardization sample was composed of 316 persons, who had characteristics that closely approximated 1980 census data. There were approximately 50 subjects in each of the age groups. Whereas the WMS had only one composite Memory Quotient, the WMS-R had twelve subtests from which the following five composite scores could be derived: General Memory, Attention-Concentration, Verbal Memory, Visual Memory, and Delayed Recall. Each of the index scores has a mean of 100 and a standard deviation of 15. This division into index scores is consistent with theories that have divided memory into short term and long term (note the Delayed Recall used to assess long-term memory) and verbal/auditory versus visual (note the Verbal Memory and Visual Memory indexes).

Reliability of the WMS-R has been generally low to adequate. Test-retest reliabilities over a four- to six-week interval were moderate (i.e., Mental Control $r = .51$; Visual Paired Associates $r = .58$; Verbal Paired Associates $r = .60$ for initial and .41 for delayed recall). Internal consistencies ranged from a low of .44 for Figural Memory to a high of .88 for Digit Span (Wechsler, 1987). The standard error of measure ranged from a high of 8.47 for the Visual Memory Index to a low of 4.86 for the Attention-Concentration Index (Wechsler, 1987).

Similar to studies on reliability, the validity of the WMS-R has been good to adequate. A number of factor analytic studies have generally found that the different subtests can be organized into two major factors described as a General Memory/Learning factor and an Attention-Concentration factor (Bornstein & Chelune, 1988; Roid, Prifitera, & Ledbetter, 1988; Wechsler, 1987). There was also some evidence for a three-factor solution composed of Verbal Memory, Nonverbal Memory, and Attention (Bornstein & Chelune, 1988). A wide range of studies supports the ability of the WMS-R to distinguish between normal and clinical groups (A. Hawkins, Sullivan, & Choi, 1997; Reid & Kelly, 1993; Wechsler, 1987), distinguishes the relative severity of deficits based on subjective complaints (Gass & Apple, 1997), provides an index that relates to client ratings of level of everyday memory (Reid & Kelly, 1993), and predicts the degree of brain atrophy (Gale, Johnson, Bigler, & Blatter, 1995). In addition, the Attention-Concentration Index was found to be one of the most sensitive measures in identifying cognitive impairment (M. Schmidt, Trueblood, Merwin, & Durham, 1994). Despite a conceptual basis for believing that visual and verbal memory would relate to laterality of deficits, research on this has produced inconsistent results (Chelune & Bornstein, 1988; Loring, Lee, Martin, & Meador, 1989). Therefore, interpretations related to laterality should be made with caution. For example, having an impaired Visual Memory Index but good Verbal Memory Index does not necessarily mean that a patient has unilateral damage to the right hemisphere.

The WMS-R had clear advantages over the WMS because it had a far better normative base, was validated on diverse populations, had quite extensive studies performed on it, and divided memory into various indexes, thereby allowing the possibility for measuring various aspects of memory. It was, however, its weaknesses that resulted in its revision within a relatively short period. One of the most serious limitations of the WMS-R has been the relatively low reliabilities of the subtests and indexes (Elwood, 1991). This is likely to significantly reduce the accuracy of measurements. In addition, the different indexes are probably not very good measures of specific components of memory. This is not to say they are not sensitive to both *general* cognitive impairment and the *degree* of that impairment. However, the specific nature of the impairment cannot be accurately determined by referring to the specific indexes despite the fact that the names of the indexes suggest that this differentiation can be made. Finally, current theories of memory were not used in the design of the WMS-R (Lichtenberger, Kaufman, & Lai, 2002).

The Wechsler Memory Scale-III was published just ten years after the release of the WMS-R. The new revision was designed not merely as a facelift of the WMS-R, but rather a "state of the art assessment instrument that comprehensively addresses the complexity of brain/behavior relationships involved in learning and memory" (Edith Kaplan in the forward to the WMS-III manual, p. iii). To accomplish this goal, new subtests were added, scoring procedures were made more sophisticated, stimulus materials were changed, and new index configurations were developed. This resulted in six primary and five optional subtests. Eight index scores could then be developed (see Table 6.1). Whereas the manual states that it is possible to administer the six primary

Table 6.1 WMS-III indexes, primary subtests, and optional subtests

Indexes	Subtests used to Calculate Indexes
Auditory Immediate	Logical Memory I, Verbal Paired Associates I
Visual Immediate	Faces I, Family Pictures I
Immediate Memory	Logical Memory I, Verbal Paired Associates I, Faces I, Family Pictures I
Auditory Delayed	Logical Memory II, Verbal Paired Associates II
Visual Delayed	Faces II, Family Pictures II
Auditory Recognition	Logical Memory Recognition, Verbal Paired Associates Recognition
General Memory	Logical Memory II, Verbal Paired Associates II, Faces II, Family Pictures II, Auditory Recognition
Working Memory	Letter-Number Sequencing, Spatial Span
Optional Subtests and Procedures:	Information and Orientation
	Word Lists I and II
	Visual Reproduction I and II
	Mental Control
	Digit Span

Source: Adapted from "The Wechsler Memory Scales," by Franzen and Iverson, 2000. In G. Groth-Marnat (Ed.), *Neuropsychological assessment in clinical practice: A guide to test interpretation and integration.* New York: Wiley.

subtests in 30 to 35 minutes, research with a clinical population indicated that it took 42 minutes to administer the eleven primary subtests (see Axelrod, 2001).

One of the most important aspects of the WMS-III is that it was developed simultaneously with the WAIS-III. This has enabled the two tests to not only share two subtests, but also to be co-normed. The normative sample consisted of 1,250 adults ranging between 16 and 89 years. Instead of 9 groups as in the WMS-R, the WMS-III had 13 different groups. These groups not only had more subjects (50 in each group for the WMS-R versus 100 for the first 11 groups of the WMS-III), but also extended to a far higher age range (74 for the WMS-R versus 89 for the WMS-III). This is appropriate because one of the more important functions of memory assessment is to evaluate older clients.

RELIABILITY AND VALIDITY

The WMS-III has better reliability than its predecessor, the WMS-R. The *WAIS-III/WMS-III Technical Manual* indicates that internal consistency for the primary subtest scores ranges between .74 and .93 for all age groups. As would be expected, the primary indexes have even better internal consistencies of .82 or higher. The one exception is the somewhat lower reliability of .74 for Auditory Recognition Delayed. Test-retest reliabilities for all age groups over a two- to twelve-week interval mostly ranged between .62 and .82 for the individual subtests and between .75 and .88 for the indexes. Again, Auditory Recognition Delayed had a somewhat lower reliability of .70. The technical manual states that even those subtests requiring the most judgment (Logical Memory I and II, Family Pictures I and II, Visual Reproduction I and II) had interscorer reliabilities above .90. However, scorers (on the WMS-R) have been found to make an average of four errors per protocol, indicating that extra care should be taken to ensure that scores are congruent with the criteria in the manual (Sullivan, 2000).

There is ample evidence that the WMS-III can effectively differentiate between clinical and normal populations. Various clinical groups (Alzheimer's disease, Huntington's disease, Parkinson's disease, multiple sclerosis, chronic alcohol abuse, temporal lobe epilepsy, schizophrenia) consistently score lower than the standardization sample (D. Fisher et al., 2000; K. A. Hawkins, 1998; The Psychological Corporation, 1997). For example, patients with mild Alzheimer's disease scored in the 60 to 71 range for most of the primary indexes except for a mean score of 80 for Working Memory (Psychological Corporation, 1997). Similarly, Fisher et al. found that patients with moderate-to-severe traumatic brain injury scored low on all indexes. WMS-III Visual Delayed and Visual Immediate (and WAIS-III Processing Speed) were particularly sensitive to the severity of the injury. Finally, the WMS-III has been found to correspond to clinician ratings of the severity of brain injury (Makatura, Lam, Leahy, Castillo, & Kalpakjian, 1999).

Although differentiating between normal and clinical groups is essential, it is also a relatively easy criterion to achieve. What is particularly crucial for the practicing clinician is to determine whether the individual indexes can accurately measure subcomponents of memory. Factor analytic studies and determining whether patterns of scores match theories of memory (i.e., visual/verbal in relation to laterality) are particularly important. The technical manual reported a factor analysis using the standardization sample and concluded that, for ages between 16 and 29, a three-factor model composed of working memory, visual memory, and auditory memory most closely fit the data. In contrast, a five-factor model composed of working memory, auditory immediate memory, visual immediate memory, auditory delayed memory, and visual delayed memory fit the age groups from 30 to 64 and 65 to 89. For ages 30 to 89, this closely corresponds to five of the eight index scores. The change in factor structure between the younger and older age groups is also consistent with findings that the components of memory become more clearly distinguishable ("dissociated") with age (Dolman, Roy, Dimeck, & Hall, 2000). Thus, the index scores might become more meaningful with older populations. An additional factor analysis also using the standardization sample supported the three-factor model reported in the technical manual composed of working memory, visual memory, and auditory memory (Millis, Malina, Bowers, & Ricker, 1999). Although the researchers did not find support for the five-factor model, it should be noted that they did not separate their groups into different ages. It is interesting to note that the previous WMS-R consistently produced factors that supported a distinction between immediate memory and delayed recall (Bowden et al., 1997; Hunkin et al., 2000). This suggests that the WMS-III may have succeeded in emphasizing more visual tasks (as well as more clearly defined factors) but, when compared with the WMS-R, may have lost differentiation between immediate and delayed memory (K. Hawkins, 1998; Millis et al., 1999). In addition, the Immediate Memory and General Memory indexes may be redundant because they have been found to correlate .98 in a variety of clinical populations (K. Hawkins, 1998; Weiss & Price, 2002).

There is some evidence that various types of clinical populations perform differently on various indexes. As would be predicted given knowledge about brain laterality, patients with right lobectomies performed considerably lower on the Visual Immediate Index than on the Auditory Immediate Index. Conversely, left lobectomy patients performed worse on the Auditory Immediate Index when compared with their Visual Immediate Indexes (K. Hawkins, 1998). However, both groups of patients performed poorly on the Visual Immediate Index. Both the Visual Immediate and Visual Delayed Indexes have also been found to be relatively sensitive to alcohol abuse, multiple sclerosis, and traumatic brain injury (D. Fisher et al., 2000; K. Hawkins, 1998). Thus, visual memory may be a particularly sensitive indicator of brain injury in general. Somewhat similarly, traumatic brain-injured patients with mild injuries showed lower than expected scores on Auditory Immediate, Visual Immediate, Visual Delayed, and General Memory (D. Fisher et al., 2000). With more severe injury, Visual Delayed and Visual Immediate were particularly affected. The index least susceptible to lowering was the Auditory Recognition Delayed.

The technical manual reports a number of performances for clinical populations. For example, Korsakoff's syndrome is characterized by severe difficulties with encoding and storing new information but the patient's attention and working memory are normal. This is reflected on the WMS-III index performances wherein Working Memory was in the normal range but all other index scores were in the impaired range (Psychological Corporation, 1997). The previous sampling of studies indicates that many of the predicted theoretical and clinical patterns of performance have occurred on the various WMS-III indexes.

ASSETS AND LIMITATIONS

The WMS-III is generally an excellent instrument capable of measuring a wide range of memory functioning. It has been based on theoretical research into the processes of memory, it has excellent standardization, and most research indicates solid empirical support. It is clearly an improvement over both the original WMS and the WMS-R. Perhaps its strongest advantage is its co-norming with the WAIS-III. This allows practitioners to make realistic comparisons between performance on the two instruments.

An important unanswered question with the WMS-III is the extent it can actually measure the various components of memory. Its divisions (and corresponding indexes) into working, visual, and auditory memories are quite well supported. However, the distinction between immediate and delayed memory may be questionable. In addition, the Immediate and General Memory indexes may be redundant. Thus, the number and titles of the indexes may promise more specificity than can actually be delivered. A related and important issue is that the various components of memory (and corresponding indexes) are likely to perform differently across various clinical populations and age groups. A final unanswered question in need of further exploration is the extent to which the WMS-III relates to aspects of everyday memory. Given the considerable research that resulted from the WMS-R, these, and many additional questions, will be answered over the next few years.

The original WMS had the advantage of taking only 15 minutes to administer. The WMS-R and now the WMS-III have increased the administration times to an average of 42 minutes, but it may actually take up to 100 minutes for some clinical populations (Lichtenberger et al., 2002). When the WMS-R was released, many clinicians either continued to use the WMS, or used only selected portions of the WMS-R (Butler, Retzlaff, & Vanderploeg, 1991; Sullivan & Bowden, 1997). Given the present co-norming with the WAIS-III and vastly superior standardization, it is difficult to justify the continued use of the original WMS. It is, however, quite likely that practitioners concerned with time efficiency will use only portions of the WMS-III. For example, they might give only those subtests that are necessary to develop an Immediate Memory index and then compare this with the WAIS-III IQs to notice discrepancies. A further option might be to give only those subtests that seem to be most sensitive to cognitive impairment (Visual Immediate and Visual Delayed) or to use empirically based short forms to extrapolate various index scores. For example, a three-subtest short form consisting of Logical Memory, Verbal Paired Associates, and either Faces or Family Pictures correlated at a .97 level with General Memory (and Immediate Memory; Axelrod, Ryan, & Woodward, 2001). A two-subtest short form composed of Logical Memory and Verbal Paired Associates had a quite similar correlation of .96 with General Memory (and Immediate Memory). These two short forms account for 95% to 97% and 87% of the variance in General Memory and Immediate Memory, respectively (Axelrod & Woodward, 2000). However, this use of nonstandardized administrations raises the possibility of introducing unknown error.

The scoring and administration of the WMS-III is mostly clearly described in the manual. However, Logical Memory does not present guidelines regarding the speed at which the stories should be read. It also does not have guidelines for intonations, pauses, or inflections. Examiner variation in each of these areas may, therefore, result in the potential for error. Lichtenberger et al. (2002) suggested that an audiotaped administration might be introduced by the test developers. A further issue with both Logical Memory I and II is its high degree of cultural loading; therefore, persons whose first language is not English may be disadvantaged.

In many ways, the complexity of the WMS-III is an advantage because it allows for the possibility of assessing a wide range of memory functions. It should be noted that other comprehensive batteries have similar levels of complexity. However, the fairly complex procedures may discourage some clinicians from learning and using it. The relatively long administration and detailed scoring procedures may also introduce the possibility of scoring and computational errors.

INTERPRETATION PROCEDURE

The WMS-III measures a wide range of different functions. As a result, interpretation can be complex. The following sequence is a general strategy that is conceptually quite similar to the interpretive procedure for the WAIS-III. The initial interpretations are for global measures, followed by progressively more specific considerations. In addition, far more interpretive strategies might be used. For example, scoring for four Auditory Process Composites are provided in the *WMS-III Administration and Scoring Manual.* These indicate a patient's level of initial learning (Single-Trial Learning), rate which learning improves over multiple trials (Learning Slope), degree information is retained after a delay (Retention columns), and extent cueing increases a person's retrieval of information (Retrieval Composite). In addition, interpretation of subtests has not been included because neither the psychometrics nor the research literature warrants such procedures. Far more in-depth strategies can be found in Lichtenberger et al. (2002). The information included next is considered both manageable and psychometrically sound and thus provides a solid, clear introduction to WMS-III interpretation.

1. Interpret the General Memory Index

The original Wechsler Memory Scale resulted in a single memory quotient that clinicians found quite useful because they could compare it to a person's Full Scale IQ to note whether there were any large discrepancies. This information could then be used to infer a relative strength or weakness in global memory compared to a person's other cognitive abilities. The General Memory Index can be used in the same manner. Indeed, an IQ score that is much larger than a global memory score (IQ > Memory) has often been used to "red flag" the possible presence of brain dysfunction. This is based in part on the clinical observation that one of the most frequent complaints by patients with brain damage is that they have memory difficulties.

One item that may not be immediately clear is the title "General Memory," which does not clearly indicate that it measures delayed memory. However, a brief review of the subtests comprising the General Memory Index indicates that they are all delayed tasks. Part of the rationale for referring to it as "general" memory is that the types of tasks assessed by this index (delayed tasks) relate more clearly to everyday types of memory functions. As a result, a more accurate title might have been the "Global Delayed Memory" index (Lichtenberger et al., 2002). Conceptually, this can still be considered an adequate global measure of memory given that memory refers to the *persistence* (delay) of learning, whereas more immediate measures of "memory" are greatly influenced by attention. Thus, because each of the subtests used to calculate the General Memory Index involves delay/persistence, any lowering of performance because of poor attention will result in a corresponding lowering in performance on the General Memory Index.

Despite the potential usefulness of the IQ > Memory distinction, caution should be used for two major reasons. First, large fluctuations occur among normal populations. Differences of 13 points occurred in 15% of the standardization sample and 15 points in 10% of the sample. An abnormal discrepancy (occurring in only 5% of the standardization sample) was a difference of 20 points or more. Table 6.2 summarizes this information and also describes the extent of unusual differences for the other index scores. Second, an IQ > Memory difference has been found as only a weak indicator of dysfunction. This is because, with brain dysfunction, there is often a corresponding decline in other nonmemory abilities as well. This results in a reduction in not only IQ, but also in measures of memory (General Memory Index), which thereby results in little difference between the two measures.

Because quite large differences between IQ and Memory are a fairly common occurrence even among normal populations, a more sensitive indicator might be to compare the relatively stable WAIS-III Verbal Comprehension Index with the highly sensitive WMS-III Visual Immediate Memory Index (K. Hawkins, 1998). While the frequency of large differences is not provided in the manual, differences of 18 or more should be investigated further. It also suggests that further research may find this a particularly good indicator of brain dysfunction.

2. Interpret Patterns of High and Low Index Scores

The purpose of interpreting patterns of index scores is to better understand a person's relative memory-related strengths and weaknesses. Initially, this might be done by noting the absolute values of the index scores. For example, a relatively low score on Visual Memory might indicate a relative weakness in this modality. In contrast, a low

Table 6.2 Reliable, unusual, and abnormal differences between FSIQ and Index Scores averaged for all ages

		Unusual		Abnormal
	Reliable	≤15%	≤10%	≤5%
Auditory Immediate	8.8	14	17	23
Visual Immediate	13.1	18	22	29
Immediate Memory	9.7	15	17	23
Auditory Delayed	11.7	14	17	23
Visual Delayed	12.9	17	21	26
Auditory Recognition Delayed	15.7	16	20	25
General Memory	9.9	13	16	22
Working Memory	11.9	13	15	20

Derived from Table C.1 and C.4; pp. 288 and 291 in the Technical Manual.

score on Working Memory might suggest attentional difficulties and possibly problems with organizing information in sequential order. However, the same caveats for interpreting index/subtest fluctuations that were discussed for the Wechsler intelligence scales are also relevant for the WMS-III. Specifically, base rates for discrepancies need to be considered so that relatively frequently occurring differences are not overinterpreted to indicate pathology. In addition, the indexes may lack sufficient specificity. Clinicians should also be aware that fluctuations could occur for a number of different reasons. It is up to each clinician to carefully evaluate these various possibilities by carefully integrating additional relevant information. Therefore, the following possible interpretations should be considered tentative.

The level of significance between the various patterns of indexes should be determined first. This can be accomplished by subtracting one index scale score from another and consulting Table F.1 in the *WMS-III Administration and Scoring Manual.* For example, a difference of 14.5 points between Auditory Immediate and Visual Immediate was significant at the .05 level (across all age ranges). Table 1.2 (also in the administration and scoring manual), however, indicates that although this is statistically significant, a 14.5-point difference still occurred among 25% of the standardization sample. When the level of significance and frequency of this difference is established, clinicians can consider possible interpretations.

The following clusters of profile interpretations are organized according to those distinctions that are both most clinically useful and have received empirical and theoretical support. One of the major concerns is to know whether there are differences between immediate (short-term) and long-term (delayed) memory. A further pattern that sometimes emerges is the difference in the relative strength of visual or auditory modalities. It is also often relevant to know if there are differences between a person's ability to retrieve (recall) information or the less difficult task of recognizing this material when presented with relevant stimuli. The final distinction is between complex attentional processes (working memory) that involve manipulating newly presented information as opposed to simpler encoding and acquisition. Knowledge related to each of these components of memory has relevance for diagnosis, treatment planning, as well as for understanding normal levels of strengths and weaknesses.

Immediate/Delayed

Immediate Memory/General Memory (Delayed) As noted previously, the General Memory index is most appropriately considered a measure of general delayed memory (or Global Delayed Memory; Lichtenberger et al., 2002). Thus, it can be used as the comparison score to contrast a person's immediate memory with his or her delayed memory. This is an important distinction that concerns practicing clinicians. As a result, it may even be one of the referral questions. If delayed memory is considerably lower than immediate memory (12 points or more for a .05 level of significance), it suggests that the person can initially learn material but then the information decays over a period of time. It should be stressed in this regard that performance on immediate memory becomes the benchmark for how much information has been lost. In other words, unless a person has at least learned something initially, there is nothing to lose. The exception might be that a person has acquired information but then may not be able to recall it (poor retrieval). However, recognizing the information is generally a much easier task. This means that the person might be able to accurately recognize information he or she may not have otherwise been able to recall/retrieve without the prompting (see interpretations related to Auditory Delayed/Auditory Recognition Delayed).

A further issue is that factor analysis of the immediate/delayed distinction on the WMS-III may not be as strong as would be optimal (see Psychological Corporation, 1997 versus K. Hawkins, 1998; Millis et al., 1999). It is likely that the two indexes (Immediate Memory and General Memory) are redundant, as they had a .98 correlation for a variety of clinical populations (K. Hawkins, 1998; Weiss & Price, 2002). Thus, WMS-III interpretations related to immediate versus delayed memory should be tentative.

Auditory Immediate/Auditory Delayed This discrepancy occurs in the same modality (auditory) but evaluates whether there are differences between short-term and long-term abilities. Low scores on both of these indexes relative to a person's IQ suggest difficulties in verbal learning and memory. In contrast, if a person's Auditory Immediate index is significantly higher than his or her Auditory Delayed (13 points for a .05 level), he or she might be experiencing a high rate of forgetting. For example, the person might be told about a meeting time and place or given a set of instructions but would have difficulties retaining this information. However, this inference needs to always take into account how much he or she originally learned based on the height of Auditory Memory as this is an indication of how much information was originally acquired. In other words, the person can forget information only in relation to how much was originally learned.

Visual Immediate/Visual Delayed Visual Immediate and Visual Delayed are within the same modality (visual), but the difference is short-term versus longer term differences in ability within this modality. Low scores on both indexes relative to a person's intelligence would indicate an overall memory difficulty with this modality. However, if immediate memory is significantly higher (17 points for a .05 level), there is likely to be important losses of visual information over a period of time. For example, the person might have studied a map or been to a meeting; but, after a period of time, he or she may not be able to recall relevant directions or remember who had attended the meeting. Keep in mind that the Visual Immediate score is always the benchmark for comparing Visual Delay because the Visual Immediate index is dependent on how much the person originally learned.

Modalities (Auditory/Visual)

Auditory Immediate/Visual Immediate One of the basic distinctions supported by WMS-III factor analysis is between auditory and visual memory. The difference between these modalities (and the indexes that measure them) can thus be used to hypothesize relative auditory versus visual strengths and weaknesses. Thus, a significant difference between Auditory Immediate and Visual Immediate (11 points or more for a .05 level of significance) can indicate either lifelong patterns related to differences in abilities or acquired deficits in these modalities. Laterality differences have been previously noted (K. Hawkins, 1998) in that patients with unilateral left hemisphere damage have been found to do more poorly for verbal-auditory information than for visual information. For example, they would be expected to have particular difficulty when given verbal directions. In contrast, they might perform far better when shown a visual map of how to get from one place to the next. In contrast, patients with unilateral right hemisphere damage would be expected to do more poorly on visual immediate tasks. Thus, they would be expected to benefit most from auditory-verbal directions compared to directions that were visually presented. However, visual memory performance was found to be the most sensitive to any type of brain damage, and patients with both unilateral right and left hemisphere damage performed poorly on visual memory types of tasks. If one modality was found to be relatively stronger than another, this stronger modality might be used to maximize learning. For example, if a person's auditory learning was poor, he or she might use learning strategies that capitalized on visual modes (or vice versa).

Auditory Delayed/Visual Delayed The same interpretive considerations as noted previously (between Auditory Immediate/Visual Immediate) also apply, except that the extent that memory is lost over a period of time is also measured. Thus, a significant difference between the two indexes (12 or more points) may indicate that there is more "decay" of memory in one modality than in another. Again, this may have practical implications in terms of developing learning strategies because either the auditory or visual mode might be used to compensate for a relative weakness.

Recall (Retrieval) versus Recognition

Auditory Delayed/Auditory Recognition Delayed Distinguishing whether a person once knew something or never knew it can be important clinical information. For example, patients with early dementing conditions frequently complain of difficulty retrieving information. Relevant behavioral observations might be statements such as "I know the answer but I just can't think of it." Qualitative approaches to the WAIS-R and WISC-III use multiple-choice formats on the Wechsler intelligence scale Information or Vocabulary subtests to try to determine this (see E. Kaplan et al., 1999; Milberg et al., 1996). The WMS-III uses this strategy by presenting recognition items on recall of stories (Logical Memory II) and pairs of words (Paired Associate Learning II). Significantly higher scores on recognition compared with delay (16 points or more for a .05 significance) suggest that the client has retrieval difficulties. He or she might experience this as frustration over not being able to find the correct word or difficulty recalling relevant and commonly known facts. Friends or work colleagues might have commented that he or she seems to have difficulty remembering information that other people are certain he or she once knew.

Complex Attention versus Acquisition/Encoding

Working Memory/Immediate Memory The WMS-III Working Memory Index is similar to the WAIS-III index of the same name in that they share Letter-Number Sequencing. However, the WAIS-III is composed of two auditory tasks (Digit Span and Letter-Number Sequencing), whereas the WMS-III has one auditory task (Letter-Number Sequencing) and one visually presented task (Spatial Span). This similarity is reflected in that they are highly correlated (.82), and they measure similar functions (see Working Memory interpretations in Chapter 5). Specifically, Working Memory measures a person's ability to attend to stimuli while at the same time performing other mental tasks. It also requires the person to effectively sequence information and make mental shifts. Because a person must monitor his or her performance and respond flexibility, Working Memory is also related to executive functioning. Thus, it involves complex attentional abilities. In contrast, the tasks of Immediate Memory (recalling story events, recognizing details from pictures) do not require as high a level of attention.

If Working Memory is significantly lower than Immediate Memory (14 points for the .05 level of significance), it suggests that the person can adequately work with simple encoding and acquisition, but may have more difficulty with more complex attentional tasks. For example, a brain-injured person might be able to learn relatively simple information in a quiet environment, but when confronted with distractions ("multitasking"), he or she might be unable to cope effectively. If both Working Memory and Immediate Memory are low in relation to intelligence, it suggests that poor attention is affecting the person's ability to learn new information.

Working Memory/General Memory A Working Memory Index that is significantly lower (10 points or more for a .05 level of significance) than the General Memory Index indicates the person is likely to experience difficulties with tasks requiring complex attention. In contrast, he or she is likely to more effectively work with tasks requiring simple encoding and acquisition. Thus, some of the same interpretations that apply for significant differences between Working Memory/Immediate Memory also apply for Working Memory versus General Memory because General Memory comprises the same tasks as Immediate Memory, but there is a 25- to 35-minute delay. In contrast, a General Memory that is significantly lower than Working Memory suggests that complex attentional skills are a relative strength. It may also suggest that there has been some decay of memory between the immediate tasks and delayed tasks (check the difference between Immediate Memory and General Memory).

3. Evaluate More Specific Functions Derived from Individual Subtests

Research and clinical lore on interpreting the WMS-III subtests are minimal when compared with the Wechsler intelligence scales. The advantage for WMS-III interpreters is that it helps professional psychologists stay more clearly focused on the much better validated index scores. However, the various combinations of subtests do measure a number of functions that are not necessarily extracted by the index scores. These might include visual construction, degree to which the patient is oriented, and visual perception. A listing of some of these functions, along with additional interpretive material and relevant WMS-III subtests, follows (derived from Franzen & Iverson, 2000 and Lichtenberger et al., 2002):

- *Orientation:* The degree to which the person is oriented to person, place, date, and general information. This is core information for any mental status evaluation (see optional Information and Orientation section).

- *Simple Attention:* Ability to comprehend and repeat simple auditory and visual information. This is assessed on the WMS-III by procedures that request the patient to repeat a series of numbers and repeat a series of sequential taps or a form board (see Digit Span Forward and Spatial Span Forward).
- *Complex Attention:* Ability to concentrate on, hold, organize, and manipulate complex information. WMS-III-related tasks include repeating digits backward, repeating a sequence of taps on a form board backwards, mentally reorganizing a series of numbers mixed in with letters (see Digit Span Backward, Spatial Span Backward, and Letter-Number Sequencing).
- *Learning:* Ability to combine new information and later recall it. Specific tasks include learning pairs of words that belong together and recalling a list of words that has been read by the examiner (see Verbal Paired Associates and Word Lists).
- *Visual Perception:* Ability to visually distinguish between and match geometric designs (see Visual Reproduction Discrimination).
- *Visual Construction:* Ability to accurately perceive and copy designs (see Visual Reproduction Copy).
- *Malingering:* The Logical Memory Delayed Recognition task requests clients to state whether (yes or no) an item was included in one of the previously read stories. Because random guessing would produce a score of 50%, scores of less than this suggest that the client is malingering (see Killgore & Dellapietra, 2000). Malingering may also be suggested if recognition doesn't improve in comparison to recall (e.g., negative Retrieval Total Score) because recognition tasks are easier than free recall tasks. Malingering may also be indicated if a patient has a lower score on the Logical Memory I Thematic Score than on the Logical Memory I Total Score. The patient should be able to have a higher performance on the far easier Thematic Score (recalling the underlying themes of stories) than the more difficult Total score (that requires him or her to recall quite specific units of information). A final quite general indicator is dramatic differences between a person's day-to-day functioning (based on evidence from corroborating sources) and performance on WMS-III measures.

In addition to the previous listing of nonindex functions, a number of observations have been made regarding behaviors and performance on the individual subtests. A listing of some of these follows. This is part of what is, no doubt, an accumulating body of qualitative observations.

- *Information and Orientation:* It is quite rare for patient groups (5% or fewer) to not know the current U.S. president's name. It suggests the possibility of pronounced adaptational and cognitive deficits (J. Ryan & Paul, 1999).
- *Verbal Paired Associates:* Although the WMS-III norms do not take into account sex differences, females typically perform better than males on Verbal Paired Associates (i.e., $M = 10.58$ for females versus 8.46 for males on total recall scales scores; Basso, Harrington, Matson, & Lowery, 2000). This effect is moderately strong (approximately 3 subscale points) and should, therefore, be considered when making interpretations.

- *Logical Memory I and II:* Excessive embellishment of stories may be a maneuver to compensate for or cover up difficulty remembering accurate information. This may result in coherent elaboration or more illogical confabulations. A further behavioral observation is to note whether a client remembers primarily the global gist of the story as opposed to quite specific linear details. This may suggest either a global, holistic mode of processing as opposed to a more linear approach.
- *Letter-Number Sequencing:* Because letter-number sequencing is quite a demanding task, this subtest is the most likely to produce stimulus overload for the patient. The person might look frustrated or say that the task is too difficult ("You're expecting too much of me"). This suggests that he or she might experience similar frustration in everyday situations that similarly require complex reasoning (multitasking).

4. Evaluate Whether Change Has Occurred (relevant only for repeat testing)

Sometimes Wechsler memory scale scores are used to document deterioration or to monitor improvement. It is tempting to peruse pretest and posttest scores and quickly infer that some sort of actual change has occurred in the patient's level of functioning. For example, a client might have had a WMS General Memory Index score of 80 directly after a head injury and, three months later, achieved a score of 85. It might, therefore, be inferred that the patient's memory has improved. However, this does not take into consideration factors such as practice effects, regression to the mean, or the relative reliability of the measure. The improvement between the pretest of 80 and the posttest of 85 might simply be the result of the patient's practicing the tasks three months previously, or the difference might simply be measurement error (reflected in its test-retest reliability). To provide a more reliable measure of change, Iverson (1999) has calculated the following change indexes:

Auditory Immediate Index	11
Visual Immediate Index	10
Immediate Memory Index	10
Auditory Delayed Index	13
Visual Delayed Index	10
Auditory Recognition Delayed Index	15
General Memory Index	8
Working Memory Index	9

To be at least 80% certain that clinically meaningful change has occurred, a patient must have a difference of equal to (or greater than) the values indicated in the right column. For example, a patient should have gone from a General Memory Index score of 80 to at least 88 (an increase of 8 points) to be 80% certain that actual change had occurred. However, it should be stressed that these values were derived from patients with traumatic brain injury; therefore, these values may not necessarily transfer to other patient groups. The values have also been rounded to the nearest whole number. Finally, the previous statistical evaluation of change accounts for the unreliability of the instrument, which does not necessarily prove that the personal or social significance of the change in scores has been demonstrated (see Beutler & Moleiro, 2001). Determining the personal and clinical meaning of changed scores requires clinicians to integrate information from a wider variety of sources.

RECOMMENDED READING

Franzen, M. D., & Iverson, G. L. (2000). The Wechsler Memory Scales. In G. Groth-Marnat (Ed.), *Neuropsychological assessment in clinical practice: A guide to test interpretation and integration.* New York: Wiley.

Lichtenberger, E. O., Kaufman, A. S., & Lai, Z. C. (2002). *Essentials of WMS-III assessment.* New York: Wiley.