The Wechsler Memory Scale, Third Edition

Chapter · December 2003

DOI: 10.1016/B978-012703570-3/50007-9

David S Tulsky
University of Delaware

Nancy Chiaravalloti
Kessler Foundation

Gordon Chelune
University of Utah

Clinical Interpretation of the WAIS–III and WMS–III

Edited by

David S. Tulsky
Kessler Medical Rehabilitation Research
and Educational Corporation
West Orange, New Jersey, USA

Donald H. Saklofske
University of Saskatchewan
Saskatoon, Canada

Gordon J. Chelune
The Cleveland Clinic Foundatión
Cleveland, Ohio, USA

Robert K. Heaton
University of California, San Diego
La Jolla, California, USA

Robert J. Ivnik
Mayo Clinic
Rochester, Minnesota, USA

Robert Bornstein
The Ohio State University
Columbus, Ohio, USA

Aurelio Prifitera
The Psychological Corporation
San Antonio, Texas, USA

Mark F. Ledbetter
Trinity Clinical Associates
San Antonio, Texas, USA

ACADEMIC PRESS
An imprint of Elsevier Science

Amsterdam Boston London New York Oxford Paris
San Diego San Francisco Singapore Sydney Tokyo

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children—Revised*. New York: The Psychological Corporation.

Wechsler, D. (1976). *Wechsler Adult Intelligence Scale: Standardization manual*. Unpublished manual. New York: The Psychological Corporation.

Wechsler, D. (1981a). *Wechsler Adult Intelligence Scale—Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1981b). The psychometric tradition: Developing the Wechsler Adult Intelligence Scale. *Contemporary Educational Psychology, 6*, 82–85.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale—Third Edition*. San Antonio, TX: The Psychological Corporation.

Weiss, L. Saklofske, D. H., & Prifitera, A. (2003). Clinical interpretation of the WISC-III factor scores. In C. R. Reynolds & R. W. Kamphaus (Eds). *Handbook of Psychological and Educational Assessment of Children: Intelligence and Achievement [2nd Edition]*. New York: Guilford Publication.

Wells, F. L., & Martin, H. A. A. (1923). A method of memory examination suitable for psychotic cases. *American Journal of Psychiatry, 3*, 243–257.

Whipple, G. M. (1908). Vocabulary and word-building tests. *Psychological Review, 15*, 94–105.

Whipple, G. M. (1909). A range of information test. *Psychological Review, 16*, 347–351.

Whipple, G. M. (1915). *Manual of mental and physical tests (in two parts). Part II: Complex Processes*. Baltimore: Warwick & York, Inc.

Whipple, G. M. (1921). The National Intelligence Tests. *Journal of Educational Research, 4*, 16–31.

Wielkiewicz, R. M. (1990). Interpreting low scores on the WISC-R third factor: It's more than distractibility. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 91–97.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *The Woodcock-Johnson III*. Itasca, IL: Riverside.

Woodworth, R. S., & Wells, F. L. (1911). Association tests: Being a part of the report of the committee of the American Psychological Association on the standardization of procedure in experiemntal tests. *Psychological Monographs, 13*, (whole no. 57).

Yerkes, R. M. (Ed.). (1921). *Memoirs of the National Academy of Sciences Volume XV: Psychological Examining in the United States Army*. Washington, DC: Government Printing Office.

Yoakum, C. S., & Yerkes, R. M. (1920). *Army mental tests*. New York: Henry Holt and Company.

CHAPTER 3

The Wechsler Memory Scale, Third Edition: A New Perspective

David S. Tulsky
Nancy D. Chiaravalloti
Kessler Medical Rehabilitation Research and Education Corporation
University of Medicine and Dentistry of New Jersey
West Orange, New Jersey

Barton W. Palmer
University of California, San Diego
La Jolla, California
and

Gordon J. Chelune
Cleveland Clinic Foundation
Cleveland, Ohio

As is true of the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III, Wechsler, 1997a; Tulsky, Saklofske, & Zhu, Chapter 2, this volume), the Wechsler Memory Scale-III (WMS-III, Wechsler, 1997b) can be simultaneously viewed as a traditional test built on early 20th century methodology and as a cutting-edge clinical scale reflecting how far memory assessment techniques had evolved as the 20th century drew to a close.

Many of the basic tasks and stimuli in the WMS-III were developed before 1925. For example, Logical Memory and Verbal Paired Associates, which together comprise the primary auditory subtests, and the Visual Reproductions subtest are modifications of tasks developed long before Wechsler first incorporated them into the Wechsler Memory Scale (WMS; Wechsler, 1945). The basic design and many of the stimuli for these subtests have seen little change since they were first normed in the late 1930s to mid-1940s. Indeed, as is true of the subtests in the Wechsler Bellevue Intelligence Scale (Wechsler, 1939) and its subsequent derivatives (i.e., the WAIS, WISC, WAIS-R, WISC-R, WAIS-III, and WISC-III), Wechsler did not, properly speaking, "invent" most of the WMS (1945) subtests. Rather, he adapted them from existing batteries such as the Binet–Simon Scale (Binet & Simon, 1905/1916a, 1908/1916b, 1911/1916c), the Wells and Martin (1923) battery, the Army testing program, and other sources. Even the Word List task, which is new to the WMS-III, can be viewed as a simple modification of Claparede's (1919) word list learning task and its subsequent derivations, (for example, Andre Rey's Auditory Verbal Learning Test, (1964); the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987); and the Hopkins Verbal Learning Test (Brandt, 1991).

As was true for the development of Wechsler's scales of intelligence, Wechsler's genius and contribution were not manifested in development of new methods of memory assessment. Rather, Wechsler appreciated a practical need in the field and brought together existing methods with normative data to guide interpretation. It is this advance that provided the basis for applying the tests in clinically useful ways. It is in this respect that the WMS-III follows the tradition set by Wechsler himself. Thus, as will be seen below, decisions about what to retain from the WMS (Wechsler, 1945) and WMS-R (Wechsler, 1987) and what to change, modify, or add in revising the WMS-III, were guided by the desire to provide a familiar and clinically useful battery, yet one that reflects contemporary neuropsychological theories and research findings while meeting contemporary standards for appropriate normative data.

The process of planning and decision making for the new revisions was guided by input from expert panels, focus groups, consultation with practicing psychologists, and senior project directors at The Psychological Corporation (TPC). Before the WAIS-III and WMS-III revision processes had even begun, a Neurogenic Focus Group met at TPC in San Antonio, Texas, in 1990, composed of both psychologists and aphasiologists. At this early meeting, the notion of developing a "family of related tests" was first introduced to TPC, as they were thinking about their future plans for test development eventually leading to the co-norming methodology that was used for the WAIS-III and WMS-III.

The "official" development of the WMS-III started in Fall 1992. Around the time the project began, advisory group meetings were convened during the annual conferences of the American Psychological Association (APA) in Washington, D.C., in 1992 and the National Academy of Neuropsychology (NAN) in Pittsburgh in 1992. The meetings were organized by Mark Ledbetter, who was the WMS-III Project Director, and moderated by Nelson Butters. Butters, a major figure in the field of neuropsychology, was instrumental in the selection of advisory board members. Although the initial panel meeting included a number of leaders in the field of neuropsychology, it was eventually narrowed to a smaller team consisting of Robert Bornstein, Robert Heaton, Robert Ivnik, and Gordon Chelune, all chosen by Butters for their leadership role in the science and practice of neuropsychology. It was during the November 1992 NAN meeting that it was officially decided to link the WAIS-III and WMS-III projects by conorming the instruments.

The current chapter will focus on this revision of the WMS within a historical context. It is divided into three sections. We begin with a review of the clinical assessment of memory and how the original WMS helped fill a need in the field. The strengths and weaknesses of the previous editions are then outlined, and we discuss how the shortcomings of the previous editions helped shape the goals and objectives of the current revision. Finally, we review the core index scores and subtests of the WMS-III and describe the origins of the current subtests. In describing the WMS-III revision in the present chapter, we attempt, wherever possible, to provide the reader with an insider's perspective on the rationale for various decisions about what was included, excluded, and modified.

HISTORICAL FACTORS LEADING TO THE DEVELOPMENT OF THE WECHSLER MEMORY SCALE

As outlined in Chapter 1, Binet had a profound impact on the assessment of intelligence and memory (Tulsky, Saklofske, & Ricker, Chapter 1, this volume). In regard to memory, some of the tests incorporated by Wechsler (1945) into the original WMS were taken directly from the Binet–Simon scales (Binet & Simon, 1905/1916a, 1908/1916b, 1911/1916c). It was also because Binet emphatically stated that intelligence and memory were separable constructs (see Binet & Simon, 1905/1916a) that the development of IQ and memory tests took separate paths. Ebbinghaus (1885/1913) began formal experimental studies of learning and memory in the late-19th century. However, few people thought about memory assessment clinically, and fewer still were using formal memory tests in clinical practice. For instance, memory disorders are not even mentioned in a comprehensive 543-page volume on the mental, nervous, and central nervous system disorders seen by the Medical Department of the United States Army during the First World War (Bailey, Williams, & Komora, 1929). Furthermore, although a proliferation of tests and research articles addressed the intelligence assessment (e.g., Boardman, 1917), tests of immediate memory were used in experimental psychopathology paradigms (e.g., Wells, 1913), and theoretical developments in the area of memory (e.g., Hebb, 1961), the clinical application of memory testing was almost nonexistent through much of the early 20th century.

By the early 1920s, only one clinical memory battery had been published (Wells & Martin, 1923). However, the Wells and Martin Memory Examination was not widely used in clinical settings (which may be partially due to an inadequate normative database to facilitate interpretation of the scores, or may simply be due to the lack of attention paid to the clinical evaluation of memory disorders). This lack of attention to the assessment and treatment of memory disorders continued for years. Even when the initial version of the Halstead Neuropsychological Test Battery, the single most widely used fixed neuropsychological test battery, was introduced in 1947, Ward Halstead did not include or mention memory assessment as an integral part of the battery (Halstead, 1947). This omission even persisted through the later expansion and popularization of the battery by Ralph Reitan (Reitan & Davison, 1974). In fact, throughout the 1940s and beyond, memory assessment was rarely performed in clinical settings and remained a secondary area of assessment for decades following, even among the pioneers and leaders in the burgeoning field of clinical neuropsychology.

World Wars I and II catalyzed a general expansion in clinical psychology. Improvements in general medical care led to a greater number of individuals who survived head injuries. Kurt Goldstein, a neurologist who had worked with World War I veterans who sustained head wounds, helped to increase the awareness of the cognitive and memory deficits that result from these injuries (Goldstein, 1942). During World War II clinical psychologists began taking a more active role in diagnosing and treating individuals with cognitive deficits. As military personnel returned from Europe and the Pacific, a new focus within psychology was on those with war-related brain injuries (Phares, 1979; Shaffer & Lazarus, 1952). As a result, the need for a clinical test to detect and evaluate memory disorders was becoming recognized. With Psychologists' increased role in treating casualities during the war and increased focus on the cognitive deficits resulting from injury, the demand for new assessment instruments grew.

Memory assessment, in the decade following World War II, began having a strong practical foundation in aiding clinical assessment (Williams, 1978), as clinicians recognized that deficits in memory often manifest themselves as initial symptoms in certain disease processes (e.g., dementia) before deficits in other areas of functioning presented themselves. Other disorders, such as Korsakoff's syndrome, were marked by severe memory dysfunction within the context of otherwise relatively normal mentation. At times, there were striking cases of memory deficits that might occur in the face of intact verbal comprehension and perceptual organization. For example, the historical case of H.M. who underwent a bilateral temporal lobectomy for the control of temporal lobe epilepsy, demonstrated intact intellectual abilities in the presence of a dense postoperative global amnesia (Scoville & Milner, 1957). A retrograde amnesia was observed, initially, with at least partial clearing over time. However, the anterograde amnesia persisted, with a complete inability to register and recall everyday events. Such cases demonstrate the clinical importance of a detailed assessment of the various components of the memory subsystems identified in the experimental literature. Moreover, given the dearth of clinically relevant memory tests in the field, the WMS represented an important advance in clinical assessment.

THE WECHSLER MEMORY SCALE

David Wechsler, by seeking to develop a standardized memory battery for clinical use in the first half of the 20th century, was well ahead of his time. He himself had become acquainted with experimental memory paradigms early in his career. For his master's thesis in 1917, Wechsler used a test battery to assess verbal memory in patients with Korsakoff's syndrome. This included an unpublished test that was a precursor to the Verbal Paired Associates subtest (Wechsler, 1917a; Wechsler, 1917b). Wechsler also spent a year working at the Boston Psychopathic Hospital with Fred Wells at approximately the time that the Wells and Martin Memory Examination was published (Wells & Martin, 1923; Wells, 1927). It is therefore likely that he became acquainted with the aforementioned Wells and Martin Memory Examination at that time.

Although inspired by the Wells and Martin Memory Examination, the actual development of the WMS did not occur for years. In the 1930s and 1940s Wechsler began development of his Wechsler Bellevue intelligence tests, and progress on the WMS appears to have taken a secondary priority. The development and publication of the WMS did not proceed uninterrupted, but rather appears to have occurred in a series of spurts over several years that were in part interrupted by focus on the two forms of the Wechsler Bellevue Intelligence Examination (Eugenia Jaros, Wechsler's secretary at Bellevue Hospital, personal communication, May 2002), as well as consultation to the military, and other miscellaneous work (e.g., Wechsler published two unrelated articles in 1944 and 1945, which indicated he was working on other projects as well; see the Cornell Selectee Index [Wechsler, Weidner, Mittleman, & Wolfe, 1944] and an experimental study of anxiety [Wechsler & Hartogs, 1945]). At the time of its initial publication, Wechsler (1945), indicated that the WMS had been in use at Bellevue Hospital by 1940 and that the publication of the scale represented the culmination of "10 years of intermittent experimentation" (Wechsler, 1945, p. 87).

Although the Wells and Martin (1923) Memory Examination did not achieve widespread clinical use, its influence on Wechsler and the structure and content of the original WMS is apparent in several respects. Foremost is the fact that Wells and Martin's test involved a variety of different types of memory-related tasks that contributed to a total score, which was then converted to an overall Memory Quotient based on an extremely small sample of individuals with psychosis (e.g., $N = 111$), whose test scores served as normative data. The influence on Wechsler can also be seen in terms of the content of specific test items. Several of the tasks in the WMS (Wechsler, 1945) have direct parallels in the Wells and Martin examination, including items relating to Personal and Current Information, counting backwards and reciting the alphabet, and repeating strings of aurally presented digits forward and backward (i.e., the Digit Span task). The Wells and Martin (1923) Memory Examination also included an associate learning task (otherwise known as Verbal Paired Associates), wherein examinees had to learn and recall a series of unfamiliar town–state associations. It is also interesting to note that the Wells and Martin Memory Examination included several items assessing what today would be labeled semantic (nonepisodic) declarative memory. Although this aspect of memory is assessed in the Wechsler intelligence tests (particularly in the Information subtest), it has never been included in the Wechsler Memory Tests (i.e., the WMS, WMS-R, and WMS-III) which are all primarily measures of episodic declarative memory, (see The Psychological Corporation, 1997).

As stated by Wechsler (1945), the WMS was designed to provide a "rapid, simple, and practical memory examination" (p. 87). The original WMS (1945) included seven subtests: (1) Personal and Current Information, (2) Orientation, (3) Mental Control, (4) Logical Memory, (5) Memory Span (i.e., Digit Span), (6) Visual Reproduction, and (7) Associate Learning (i.e., Verbal Paired Associates). There was an alternate form of the original WMS (i.e., Form II), that allowed test–retest evaluation, but the two versions were not statistically equated.

The WMS approach, like that of the Wells and Martin Memory Examination, was somewhat comparable to that of IQ testing in that a range of diverse memory related tasks were combined to compute an overall Memory Quotient (MQ), which reflected the degree to which an examinee's performance deviated from the normative mean. Note that this approach differed philosophically from an older approach in experimental memory studies, where the focus was simply on the amount of information saved (or the amount of time or additional exposure to relearn material—e.g., Ebbinghaus). Wechsler suggested adding a constant to the total score to correct for age and then provided a table for converting the sum of these scores to an MQ. Wechsler derived these MQ conversions by "plotting the mean [total WMS] scores for different ages against the weighted scores of the Wechsler Bellevue [Intelligence] Scale (age group 20–24 years) and then trying out various constants that would keep the mean MQ for any group equivalent to the mean IQ of that age group" (Wechsler, 1945, p. 90). This was an interesting approach in that, indirectly, Wechsler was asserting that interpretation of memory scores should be partially weighted by general intellectual functioning. Although Wechsler's technique only provided for age-group corrections, the approach seems to anticipate the decision to co-norm the WAIS-III and WMS-III, so that WMS-III results could in part be interpreted in light of examinee's general cognitive functioning. As there were few scales designed to measure memory functioning, the WMS quickly became the state of the art instrument when memory assessment was required, and became extremely popular in the field (see Appendix II Item 8 for initial reactions).

Despite the frequent use of the WMS in research with clinical populations and its sensitivity in detecting deficits in short-term memory, the field of experimental memory research had advanced by the 1960s and 1970s (Baddeley, 1976; Pribram & Broadbent, 1970; Tulving & Donaldson, 1972), and psychologists were growing more critical of the WMS. Perhaps the single most influential critique/response to the original WMS was that by Elbert Russell (1975). Russell was critical of the MQ as he asserted that it seemed to imply a unitary function whose various aspects were additive. He noted that the WMS blurred a potentially important distinction between verbal and figural memory. Moreover, Russell noted that research had suggested a distinction between immediate versus long-term memory, whereas the WMS measured only immediate recall. Fortunately, Russell (1975) went beyond merely criticizing Wechsler's original scale, and his proposed solution was an alternate administration technique for the WMS. The changes advanced in Russell (1975) called for (a) elimination of the

MQ; (b) systematic assessment of verbal versus visual discrepancies; and (c) systematic assessment of immediate versus delayed memory.

The Russell revision called for the comparison of scores from Logical Memory (as a measure of auditory memory) and Visual Reproductions (as a measure of figural memory). He also reintroduced the administration procedure of assessing both immediate and delayed recall for Logical Memory and Visual Reproduction that had been previously used by Rey (1941; see Boake, 2000) in his Word Lists test. The delayed recall was to be administered after a 30-minute interval following the administration of the immediate-recall condition. Although the normative data that Russell provided for these revisions was criticized by some, even Russell (1988) himself, the format changes were readily incorporated by neuropsychologists when administering the WMS, and greatly influenced the structure of the future WMS-R, as well as the future WMS-III.

Russell was not alone in critiquing the model of memory inherent in the WMS scoring. Clinically, western society's introduction to Luria's views regarding neural systems and the functional organization of the brain–behavior relationships (Luria, 1966, 1970, 1973) challenged the concept of memory as a unitary process, and thus the concept of a global MQ was questioned. Furthermore, advances in the clinical study of split brain patients who had undergone hemispherectomy or commissurotomy were influencing the way neuropsychologists viewed memory (see Kinsbourne & Smith, 1974; Nebes, 1974; Sperry, 1968), and a differentiation between verbal and visual memory, as later advocated by Russell (1975), began to develop.

Yet while Russell (1975) advocated separate consideration of verbal versus visual memory, the primary focus of his revisions was still on the detection of brain injury. As neuropsychology began to emerge from what Rourke (1982) described as its *static phase*, with an emphasis on diagnosis, lesion identification, and hemispheric localization, to its *cognitive phase*, that focused more on descriptive brain–behavior relationships, the limitations of the WMS became even more obvious. That is, there was a desire and need to describe more specific memory processes and abilities, not merely derive scores that efficiently discriminated brain-injured from nonbrain-injured patients (or those with left-hemisphere injuries from those with right-hemisphere injuries). The evolution of this new focus is seen in the emphasis on index scores in both the WAIS-III and WMS-III.

Whereas the thrust of Russell's critique had been on the construct validity of the WMS, other critics noted additional psychometric limitations in the 1945 scale. For example, Prigatano (1978) noted several salient weaknesses in the WMS, including (a) the small "standardization sample" consisting only of approximately 200 examinees, divided into two age groups, who were seen at Bellevue Hospital in New York City and who were not suspected of having memory problems,[1] (b) the composite MQ was derived from a sample of only 100 subjects; and (c) there were few reliability studies for the WMS, and those that were reported were conducted in a "varied" and "unsystematic" fashion (p. 820). Like Russell, Prigatano (1978) also expressed concern about the construct validity of the WMS, noting that factor analyses of the WMS had supported more factors than the scale was designed to measure, as purported by Wechsler, and the findings among these studies were generally not consistent. Prigatano also noted that subtests included in the WMS were largely measures of verbal memory, with only the Visual Reproduction subtest measuring visual memory. Prigatano further commented that the exact scores that could be derived from the WMS and, more importantly, exactly what constructs the WMS was measuring, remained to be determined. These latter questions raised uncertainty about the validity of the scale, calling its utility into question for many clinicians and researchers. Additionally, the WMS primarily assessed immediate recall. It did not adequately assess the process of learning (through multiple learning trials), and it contained no assessment of delayed recall and recognition.

While such critiques, and the Russell revision in particular, were highly influential in clinical practice, the first official revision of the WMS was not published until 1987, more than 40 years after Wechsler originally published the scale. Some of the considerations in this revision were similar to those that motivated Russell a decade earlier.

THE PUBLICATION OF THE WMS-R

The revision of the WMS began in the late 1970s, and David Wechsler played a significant role in designing the new subtests for inclusion (Herman, 1988). The project was delayed for a few years following Wechsler's death in 1981 and the relocation of The Psychological Corporation, first to Cleveland, Ohio, and then to San Antonio, Texas, in the mid-1980s. The Wechsler Memory Scale-Revised was published in 1987 (WMS-R, 1987) and attempted to incorporate many of the conceptual changes advocated by

[1]Although Wechsler (1945) stated that "The subjects were not hospital patients," (p. 88) it is not clear in the original article how or where they were recruited.

Russell, Prigatano, and others, while also correcting other psychometric problems. The revision featured an improved, though still limited, normative database, and included four new index scores (in addition to a General Memory Index). With the publication of the WMS-R, the comparisons between verbal and visual memory were made at the Index score level as opposed to an individual subtest level that Russell had advanced (e.g., Logical Memory vs. Visual Reproduction) because it increased reliability. The de-emphasis on a single MQ score, the focus on verbal–visual comparisons, and the focus on immediate-delayed comparisons were major improvements in the WMS-R and reflected the dominant thinking in neuropsychology at the time. Yet, despite the improved structure of the WMS-R, several areas of difficulty in the scale were noted.

Probably the most common and immediate criticisms of the WMS-R were limitations related to the normative sample (D'Elia, Satz, & Schretlen, 1989; Loring, 1989). The sample size ($N = 316$) was relatively small compared to other well-standardized pychometric scales such as the WAIS-R ($N = 1880$), which had been published a few years earlier. Additionally, Loring (1989) and D'Elia et al. (1989) expressed concerns about the lack of normative data in certain age groups. Beyond concerns with the normative sample, Loring also noted that the visual subtests lacked face validity in their assessment of visual learning and memory and lacked sensitivity to patients with known memory deficits. Moreover, the General Memory Index of the WMS-R reflected immediate recall trials and showed a slight bias toward verbal tests (most likely due to the higher correlations between the verbal subtests and the General Index Composite; see Chelune, Bornstein, & Prifitera, 1990). Additionally, the Delayed Recall Index was not constructed similar to the General Memory Index, making comparisons between immediate and delayed memory difficult.[2]

Indeed, although the sample range, size, and representativeness had increased over the WMS standardization sample, the WMS-R standardization sample consisted of only six age groups with approximately 50 individuals within each group (ages 16–17, 20–24, 35–44, 55–64, 65–69, and 70–74). This small sample size increased the probability of measurement error. Furthermore, the normative sample contained age groups for which no data were collected, and the norms for these age groups where computed through interpolation based upon linear regression from adjacent age group

[2]Unfortunately, the lack of continuity between immediate and delayed memory was not corrected in the WMS-III, but this is being rectified by Tulsky, Chelune, Price & Millis (2003) who have constructed a new Delayed Memory Index derived from the WMS-III standardization sample.

data. This becomes especially questionable if the changes of memory function across ages are not linear in nature.

The second major criticism of the WMS-R focused on the *novel* visual memory subtests that had been added in this revision. Figural Memory and Visual Paired Associates had been added to the scale with the rationale that the abstract stimuli contained within the subtests would be difficult to encode verbally. However, the field seemed to have had a backlash to such novel tasks. The subtests were not well received and were controversial from the start. Loring (1989) pointed out that the Figural Memory subtest loaded substantially on the Attention/Concentration factor in a group of mixed clinical patients, and he raised questions about what the task was measuring. Moreover, the Figural Memory subtest was a recognition task without a recall component making it different from the other tasks in the battery. The Visual Paired Associates subtest was also criticized. Loring questioned the validity of the subtest as a measure of nonverbal memory because examinees tended to verbally encode the stimuli (Loring, 1989). In addition, Loring pointed out that the validity studies for Visual Paired Associates and Figural Memory in clinical groups were inadequate. Moreover, the new subtests appeared to be limited in their ability to demonstrate differences in functioning of the right and left cerebral hemisphere. Several studies demonstrated that the verbal memory subtests were highly sensitive to left-hemisphere lesions, but performance was relatively preserved in individuals with right-sided lesions (Barr et al., 1997; Chelune & Bornstein, 1988; Naugle, Chelune, Cheek, Lüders, & Awad, 1993; Rausch & Babb, 1993; Saling, Berkovic, O'Shea, et al., 1993). However, the opposite pattern was not observed with the visual subtests, as both right- and left-lesioned groups obtained lower scores than healthy individuals with roughly the same magnitude of impairment.

Loring (1989) also pointed out that the WMS-R did not take sufficient advantage of many of the research findings on memory and learning that had accumulated between the publication of the original WMS and the 1987 revision; that is, the WMS-R reflected earlier concepts of memory rather than newer cutting-edge approaches. Specifically, independent and arguably better tests were developed between 1945 and 1987 for the detailed assessment of specific forms of memory. For example, Buschke (Buschke, 1973; Buschke & Fuld, 1974) developed the Selective Reminding Test in order to examine storage, retention, and other processes distinguished in the experimental verbal learning and cognitive neuropsychological literatures. Similarly, the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987), which is administered in a format similar to the Rey Auditory Verbal

Learning Test (Rey, 1941), was also introduced as a memory task that evaluated multiple process components of learning, delayed recall, and delayed recognition. In contrast, the WMS-R remained heavily focused on "immediate memory" with little distinction between learning/encoding, storage, retention, and retrieval.

Finally, the WMS-R (Wechsler, 1987) was also criticized for yielding inadequate recognition memory data (D'Elia et al., 1989; Fastenau, 1996). For example, D'Elia et al. (1989) noted

> There are currently *no* data available regarding recognition testing on delay for any of the original or revised WMS subtests. Without assessing delayed recognition for the material not freely recalled, one is unsure whether the patient has an *encoding* problem or a *retrieval* problem. Until such data become available, the WMS tests should be considered as providing a limited assessment of memory functioning. (p. 565)

Indeed, Fastenau (1996) criticized the WMS-R for these deficiencies and developed recognition procedures for the Logical Memory and Visual Reproduction subtests. Fastenau also added a copy condition to Visual Reproductions to sort out constructional deficits from visual memory deficits.

DEVELOPMENT OF THE WMS-III

The WMS-R had been on the market for only 10 years when the WMS-III was published. This short interval between revisions may be due in part to the opportunity to respond to developments in the field of memory research, the opportunity to co-norm the WMS-III with the WAIS-III, and the perceived problems of the WMS-R that were outlined above. Nevertheless, the time period that elapsed between the publication of the WMS-R and the WMS-III revision saw a significant shift in the popular methods for memory assessment. Four conceptual shifts to the Wechsler model of measuring memory were considered at the start of the WMS-III project.

First, it was believed that a state-of-the-art memory test would address several components of memory and learning (e.g., encoding and learning, storage retention, and retrieval), and the development of core measures that could differentiate among these cognitive processes became one of the central goals of the WMS-III revision. The WMS-III would include index scores reflecting immediate recall, delayed recall, and when possible, recognition following administration of the delayed recall measures. Provision of these index scores was intended to provide greater specificity in distinguishing between clinical disorders.

As had been noted by Fastenau (1996), as well as in the memory literature in general (e.g., Squire, 1986, 1987), there is a clinically important distinction to be made between free recall versus recognition subtests. When the WMS and WMS-R were developed, recognition memory was simply not seen as an important element. However, as conceptual models evolved, the construct of recognition memory had become an important aspect to include in the new memory test. Therefore, the WMS-III places significantly more emphasis upon delayed recognition memory than its predecessors, the WMS or WMS-R. In particular, the WMS-III delayed recognition can be directly compared to a delayed free recall index score.

The second conceptual shift was the inclusion of process scores in the WMS-III revision. In the late 1980s and early 1990s, there was a rise in the process approach to neuropsychological assessment (Kaplan, 1988; Kaplan, Fein, Morris, & Delis, 1991) and to memory assessment (Delis et al., 1987). Therefore, process scores were incorporated into the WMS-III to provide information about the nature of any underlying memory deficits (i.e., giving information about what processes were impaired, beyond the type of information provided about the level of abilities that is the focus of the core index scores). These (optional) process scores are purported to provide additional clinical utility by providing a more in-depth, comprehensive examination of cognition through examination of the process through which the individual reaches his or her response. A major premise in this approach was that the patient's final response is only a piece of the important clinical information; equally important was consideration of the *process* or method the examinee used to get to that response. However, the process scores should be viewed as somewhat more exploratory than the core index scores. Generally, these scores are not normally distributed, have poor reliability, are presented as percentile scores, and there is less clinical and research background presently by which to make firm conclusions about so-called impaired scores.

A third conceptual shift in the WMS-III revision emphasized ecological validity. In the WMS-R, the visual tasks included several figures that were thought to be difficult to verbally encode and were purely visual tasks. However, it is unclear if purely visual tasks exist. In addition, a test made from such material would have little ecological validity and would not be representative of tasks that people encounter on a daily basis in real life. Therefore, in the WMS-III there was a move away from attempting to develop abstract designs purported to be purely visual and not verbally encodable. So, in devising the stimuli for the WMS-III, one consideration was the inclusion of stimuli similar to those that people experience in everyday life—a newscast (Logical Memory Story B), faces of people

(Faces subtest), and pictures of a family engaged in activities (Family Pictures). In developing the tasks, the developers of the WMS-III focused on the mode of presentation of the information (visual, auditory), which can be manipulated experimentally, rather than making assumptions about how the brain is going to process the information. One of the new visual subtests (Family Pictures) even required verbal responses, something that would never have been considered in the WMS-R. In recognition of this focus on the mode of presentation, the verbal material was renamed "auditory" to reflect that it was presented auditorially and the visual information was presented visually.

The fourth shift in the creation of the WMS-III was the reconceptualization of the "traditional" Attention/Concentration factor, by renaming it Working Memory. This factor was thought to parallel the factor on the WAIS-III, thus building a "bridge" between the WMS-III and WAIS-III. There remained one fundamental difference—the WAIS-III Working Memory Index is composed strictly of verbal subtests, whereas the WMS-III Working Memory Index is composed of a verbal and a visual task, with Letter Number Sequencing being the link.

In addition to changes that accompanied the more major conceptual shifts, a major goal of the WMS-III revision was to obtain a larger, more representative and continuous standardization sample, across a broader age range. The sampling plan for the WMS-III was established with 100 individuals per age group and an age range of 16–89 years, which was to be stratified on the following demographic variables: age, education level, ethnicity, and sex. Furthermore, the goal was to obtain an average IQ of 100 at each age level, making it similar to Wechsler's original approach of linking memory with IQ, rather than allowing both procedures to vary independently. However, this later prerequisite proved the most difficult, and a case-weighting procedure (Cochran, 1977) was used to ensure that the sample was "reasonably" representative of the census information for U.S. sample (see Tulsky & Ledbetter, 2000, for the initial report about the case weighting procedure). While case weighting is common in test development and there are strong arguments supporting its use (see Gorsuch, 2001), the technique had not been reported in the published manual (The Psychological Corporation, 1997), an omission that might give the appearance of recruitment problems or, worse, an unrepresentative sample. In actuality, however, the weighting procedure reflects the desire to have a "perfectly balanced" sample with an average IQ of 100 for each age group. Nevertheless, the shear size (weighted $N = 1250$; unweighted $N = 1,032$ Tulsky & Ledbetter, 2000) and age range (16–89) of the WMS-III weighted standardization sample, make it one of the premier normative databases for a memory test.

DESCRIPTION OF WMS-III

The WMS-III incorporated significant changes from the previous edition. As outlined in the *WAIS-III–WMS-III Technical Manual* (The Psychological Corporation, 1997) the WMS-III includes eight primary index scores (three more than had been included in the WMS-R) as well as substantial changes at the subtest level. Notably, the Visual Index scores are composed of new subtests that are unique to a Wechsler scale (e.g., Faces and Family Pictures). These primary indexes, as well as the core subtests from which they are composed, are outlined in Figure 1.

As shown in Figure 1, the WMS-III contains three global composite scores: Immediate Memory Index, General Memory Index (delayed recall),

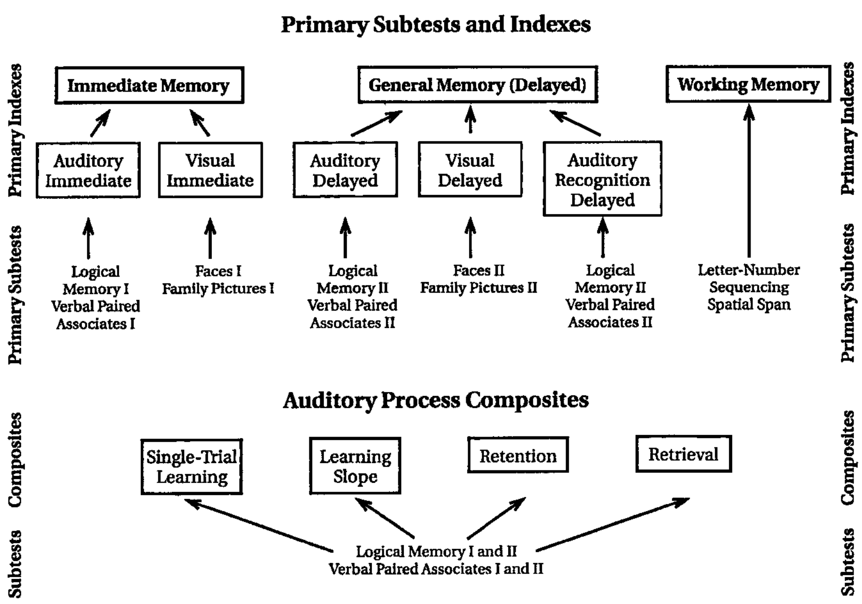

FIGURE 1. WMS-III Index and Composite Configuration. (Reproduced from Wechsler, 1997. *The WMS-III Administration and Scoring Manual.* San Antonio, TX: The Psychological Corporation. Figure copyright © 1997 The Psychological Corporation. Used with permission. All rights reserved.)

and the Working Memory Index. Within this structure, the WMS-III contains an additional five index scores allowing assessment of auditory and visual immediate memory, auditory and visual delayed memory, and auditory recognition. However, as will be discussed below, confirmatory factor analyses of the original data have not supported some of these indices (and, based upon new analyses, a new structure of cognitive functioning has been proposed by Tulsky, Ivnik, Price, & Wilkins, Chapter 4, this volume). The WMS-III also attempted to develop an Auditory Delayed Recognition Index so that differences between recall and recognition memory could be assessed directly. However, as will be reported below, such comparisons may be difficult to make due to ceiling effects among normals on the Auditory Recognition Index (particularly in terms of the Verbal Paired Associates task). An alternative method for scoring and interpreting recognition scores will be provided later in this chapter. In addition to these primary five index scores, the WMS-III contains numerous process variables, but they are outside the scope of this chapter and will not be reviewed here.

STRUCTURE OF WMS-III INDEX AND SUBTESTS SCORES

The Auditory Immediate and Delayed Index Scores

The subtests contributing to the WMS-III auditory index scores involve the auditory presentations of verbal information. The computation and interpretation of the auditory index scores are described in the *WAIS-III–WMS-III Technical Manual* as

> measures of memory functioning when stimuli are presented in the auditory modality. Low scores, relative to an individual's intellectual and attentional functioning, may suggest a verbal learning or memory problem. A low score on the delayed index, relative to that on the immediate index, may indicate a high rate of forgetting. The assessment of delayed recall should always be made in the context of the immediate condition because delayed recall (i.e., the amount of information available through recall) depends on the amount of information that was initially acquired. (The Psychological Corporation, 1997, p. 191)

The auditory index scores are based upon the sum of scaled scores of the Logical Memory I and the Verbal Paired Associates I subtests. The WMS-III also includes an optional Word Lists subtest that is not factored into the Auditory Index scores. These subtests are reviewed below.

Logical Memory

The Logical Memory subtest of the WMS-III requires the oral presentation of a narrative story to the examinee for immediate recall and subsequent delayed recall and recognition. Logical Memory I and II, on the WMS-III, parallel the traditional WMS and WMS-R subtests. For the immediate version (LM I), the examinee is asked to verbally recall each story immediately after aural presentation; for the delayed version of the stories (LM II), the examinee is asked to recall the stories after a delay interval. Immediately upon finishing the delayed recall portion, the examinee is asked a series of "yes or no" questions in which the examinee has to recognize whether or not each piece of information had been presented in the story. This latter score is a measure of auditory recognition.

The use of a narrative recall to test memory functioning can be traced back to the late 1800s when Binet and Henri (1895, 1896; also see Peterson, 1926, for a review) introduced memory for sentences as they were advancing their work on individual psychology. Their procedure entailed a teacher reading a short prose passage, after which the examinees were to reproduce it in writing. The length of the prose selections ranged from 11 to 86 words, which were divided into component ideas. At the time, Binet and Henri interpreted this test as measuring attention, characterizing the method as a *dynamometer* of attention. It was additionally noted that memory for sentences was approximately 25 times superior than memory for words (Binet & Henri, 1895; and as reviewed by Peterson, 1926).

Wissler (1901) refined this technique when he applied it to students at Columbia University. Students were read a short story and asked to recall as many of the story details as possible. Whipple (1915) later coined the term *logical memory* as the recall of "connected, meaningful material" (p. 205).

This procedure has been used widely in the assessment of intellectual abilities and memory functioning. For example, Healy-Fernald presented the Auditory Verbal Memory Test, which entailed the presentation of prose passages for immediate memory (Healy & Fernald, 1911). In Whipple's Logical Memory Test (which is the first time this term was used), "The Marble Statue" and other stories are read to the examinee, after which the examinee is required to write down everything he or she remembered (Whipple, 1915). Early logical memory stories are presented in Figure 2. Modeled after the Stanford–Binet, the Army Alpha test, which screened United States men for participation in World War I, included the reading of a written narrative, after which individuals were asked to recall as many

Three Houses Burned
(53 words, 19 ideas)
Binet & Simon, 1908

Three | Houses | Burned | Chalons-sur-Marne | September 5th | A very large fire | destroyed | last night | (Three buildings at Chalons), situated in the center of the city. | seventeen families | are without shelter (homes). | The loss will exceed 150,000 francs. | While saving | a child | in it's cradle | a barber's boy | seriously | injured | his hands.

Note: in the Binet–Simon version, the children read the stories and then repeated the passage from memory two to three seconds after they had finished reading the passage. The story was modified and adapted to a New York setting by J.E.W. Wallin, (1912).

The Marble Statue
(166 words, 67 ideas)
Reported by Whipple, 1915

A young | man | worked | years | to carve | a white | marble | statue | of a beautiful | girl. | She grew prettier | day by day. | He began to love the statue | so well that | one day | he said to it : | "I would give | everything | in the world | if you would be alive | and be my wife." | Just then | the clock struck | twelve, | and the cold | stone began to grow warm, | the cheeks red, | the hair brown, | the lips to move. | She stepped down, | and he had his wish. | They lived happily | together | for years, | and three | beautiful | children were born. | One day | he was very tired, | and grew | so angry, | without cause, | that he struck her. | She wept, | kissed | each child | and her husband, | stepped back | upon the pedestal, | and slowly | grew cold, | pale | and stiff, | closed her eyes, | and when the clock | struck | midnight, | she was a statue | of pure | white | marble | as she had been | years before, | and could not hear | the sobs | of her husband | and children.

Test XIII.
Healy & Fernald (1911 Battery)
(82 words, 12 ideas)

If a sailor | on the ocean | is shipwrecked | in a wild country, | he must first look for water to drink, | then he must find a place to sleep | where wild animals can't get at him, | and after that he can take time to look for food, | but he must be careful not to eat poisonous berries or fruit. | Next he had better hunt for other people on the land | and put up a flag | to stop ships which may be going by.

FIGURE 2. Examples of Logical Memory Stories. Three Houses Burned reprinted from Binet & Simon (1908/1916b); The Marble Statue from Whipple (1915); and Test XIII from Healy & Fernald (1911).

details as possible (Yerkes, 1921). As discussed by Tulsky, Saklofske, and Zhu, (Chapter 2, this volume), David Wechsler was very familiar with the World War I test material, and it should not be a surprise that the "Anna Thompson" story that was included in all three versions of the Wechsler Memory Scales was originally developed for the World War I test of mental abilities (Yerkes, 1921).

With the revision from the WMS-R to the WMS-III, a few modifications have been made to the Logical Memory subtest in hopes of improving its evaluative ability. Specifically, the story content was updated. Details contained in the "Anna Thompson" story have been revised to reflect more current societal norms (e.g. "police station," rather than "city hall station" in the WMS-R). Additionally, the former "truck accident" story, which was highly controversial for use with TBI survivors following a motor vehicle accident, was replaced with a more detailed story involving more emotionally neutral content. The third major change in Logical Memory involved the addition of a second presentation of the "Joe Garcia" story, which gives the clinician the opportunity to evaluate prose memory after a repeated exposure as opposed to simply relying on a one-trial presentation. This modification reflects the well-established finding that repeated exposures facilitate learning and retention (e.g., Ebbinghaus, 1895/1913). Moreover, delayed recall will be more informative if the individual has learned more of the information initially. The third major revision in administration of Logical Memory involves the addition of a recognition memory section following the long delay free recall of the stories, as proposed by Fastenau (1996) with the WMS-R Logical Memory subtest. As discussed above, this addition to Logical Memory also reflects current thoughts in the field citing the importance of assessing both recall and recognition abilities, a distinction that can be quite useful in distinguishing individuals suffering from retrieval failure from those with inefficient learning and consolidation (Fastenau, 1996; Squire, 1986).

In addition to modifications in the administration of Logical Memory, the scoring system has been modified to increase the information gained from this subtest. The traditional scoring of the number of details recalled correctly remains. However, the additional scoring of the number of more general thematic *ideas* recalled from the story has been added. This allows the clinician to quantify differences between those individuals who are unable to recall details, but recall the gist of the story, from those who simply cannot retain (or who have difficulty retaining), the information in any form. Interestingly, this thematic *idea* scoring procedure has its origins in the early 1900s, with the originator of Logical Memory. Whipple (1915) noted the "idea score" to be the most popular method for scoring this task, distinguishing between ideas that were "identical" to the narrative from those simply "equivalent" to the narrative. The revised version of the WMS-III also allows the computation of a retention score, which can be compared to retention scores in the normative sample. Once again, Whipple (1915) highlighted the importance of retention, noting that such scores could be computed by examining the percent of information lost between two recollections of the story. In its current usage, this has been applied as a percent retained from immediate recall to delayed recall.

Verbal Paired Associates

The WMS-III Verbal Paired Associates subtest requires the presentation of eight pairs of unrelated words over a series of four trials. Recall of the pairs is assessed after each individual trial and after a 25–35-minute delay. This is done by presenting the examinee with the first word in each pair and asking him or her to provide the second. Recognition memory is also assessed after delayed recall by presenting pairs of words to which examinees have to indicate whether or not they had seen the pair before.

Use of the paired associate procedure in assessing verbal memory dates back to 1917, at which time Wechsler adapted an analogy test by Woodworth & Wells (1911) for use in his master's thesis at Columbia University (Wechsler, 1917a, 1917b). In this original usage, Wechsler administered a series of "preformed associates" or pairs of related words and "new-formed" associates or pairs of unrelated words to individuals with Korsakoff's syndrome (see Figure 3). Wechsler maintained this format of presentation for his original commercially available memory scale, the Wechsler Memory Scale, and for the revised edition of the scale (WMS-R).

Verbal Paired Associates I and II in the WMS-III also parallel the WMS and WMS-R subtests on which they are based. They require examinees to learn pairs of words that are seemingly unrelated. The examinee is asked to repeat the "paired" word over four consecutive administrations to yield the immediate score for Verbal Paired Associates (VPA I). For the delayed portion, the examinee is asked to recall the word that had been paired with the initial word (Verbal Paired Associates II). Following this procedure, recognition memory is tested with the examinee saying "yes" or "no" to whether the pair of words is part of the previously learned set or not.

Although the concept of Paired Associate Learning is maintained in the WMS-III, this revision saw a significant alteration in the test items.

From Woodworth & Wells (1911)

Mixed Relations Test

I

Eye—see	Ear—
Monday—Tuesday	April—
Do—did	See—
+Bird—sings	Dog—
Hour—minute	Minute—
Straw—hat	Leather—
Cloud—rain	Sun—
Hammer—tool	Dictionary—
Uncle—aunt	Brother—
Dog—puppy	Cat—
Little—less	Much—
Wash—face	Sweep—
House—room	Book—
Sky—blue	Grass—
Swim—water	Fly—
Once—one	Twice—
Cat—fur	Bird—
Pan—tin	Table—
Buy—sell	Come—
Oyster—shell	Banana—

Mixed Relations Test

II

Good—bad	Long—
Eagle—bird	Shark—
Eat—bread	Drink—
* Fruit—orange	Vegetable—
Sit—chair	Sleep—
Double—two	Triple—
**England—London	France—
Chew—teeth	Smell—
Pen—write	Knife—
Water—wet	Fire—
He—him	She—
Boat—water	Train—
Crawl—snake	Swim—
Horse—colt	Cow—
Nose—face	Toe—
Bad—worse	Good—
Hungry—food	Thirsty—
Hat—head	Glove—
Ship—captain	Army—
Man —woman	Boy—

From Wechsler (1917b)

Performed		New
Set I	**Set II**	**Set III**
metal—iron	eagle—bird	** crush—dark
** come—go	dog—barks	school—grocery
** baby—cries	insect—fly	cabbage—doll
rose—flower	night—day	in—although
north—south	knife—sharp	jury—eagle
lead—pencil	long—short	obey—inch
up—down	pint—quart	faraway—unlikely
fruit—apple	in—out	necktie—cracker
murder—crime	cabbage—vegetable	sailor—aloud
lock—door	country—France	dig—guilty

FIGURE 3. Examples of Verbal Paired Associates. Lists reprinted from Woodworth & Wells (1911) and Wechsler, D (1917b). Note: The origins of the Verbal Paired Associates task can be seen in the Woodworth and Wells (1911), where 6–7 items were used as the basis for an early version of Verbal Paired Associates by originally published in Wechsler (1917b). These items would then appear in the Wechsler Memory Tests as nine items (not reprinted here) were included in the original WMS subtest and 10 of other items made up the list of paires for the alternate form of the WMS (see Stone, Girdner, & Albrecht, 1946).

The eight pairs of words presented are now all unrelated pairs, as these stimuli present the greatest learning challenge to the examinee due to the fact that a pre-formed relationship does not exist in the examinees' memory. In addition, the number of trial repetitions has been changed from a minimum of three trials and maximum of six trials in the WMS-R to a standard four trials in the WMS-III. Recognition memory was also added following delayed recall as the examinee is asked to "recognize" the target words, and this score is a component of the Auditory Recognition Delay Index. However, as will be discussed later in the chapter, this task is extremely easy, with most healthy examinees getting a perfect or near-perfect score.

Word Lists

A significant modification of the WMS-III included the addition of an optional list learning task. The list learning subtest of the WMS-III involves the presentation of a 12-item list of unrelated words over a series of four trials. Word List I is derived from the sum of the four trials. A second list is then presented once for immediate recall, following which the examinee is asked to again recall the first list. Free recall and recognition (yes–no format) of the initial list are later assessed after the delay interval. This procedure allows the assessment of immediate recall, delayed recall, and delayed recognition, as well as the impact of interference on learning and other process-related variables.

Auditory verbal list learning is a widely used paradigm for memory assessment, with its usage dating back to the early 20th century. Boake (2000) described the development of the Auditory Verbal Learning Test, tracing its development back to Edouard Claparede (1919), a Swiss psychologist. This test subsequently underwent modifications, becoming widely known as the Rey Auditory Verbal Learning Test (1941), and continues to be utilized widely.

Despite its very specific and specialized origins, the idea of a verbal list learning task has been conceptualized quite differently over the years, resulting in numerous such tasks available to the 21st-century psychologist. Some of the most popular list-learning tasks include the California Verbal Learning Test (Delis et al., 1987), the Hopkins Verbal Learning Test (Brandt, 1991), and the Selective Reminding Test (Buschke, 1973).

The Visual Immediate and Visual Delayed Index Scores

In the published version of the WMS-III, the Visual Immediate Index is composed of the Faces I and Family Pictures I subtests. The Visual Delayed Index is composed of the Faces II and Family Pictures II subtests. Optional visual memory subtests include Visual Reproduction I and II. However, in this book we have proposed an alternate structure of the visual memory index scores and have developed new normative information for the sums of the Visual Reproduction and Family Pictures scaled scores. For these restructured visual memory index scores, the Faces subtest is treated as optional. These norms and the rationale for replacing the Faces subtest with Visual Reproduction are discussed by Tulsky et al. (Chapter 4, this volume).

Memory for Faces

In Faces I and II on the WMS-III, the examinee is asked to view target faces and distinguish these "target" faces from foils (see Figure 4). For Faces I, the examinee is initially presented photographs of 24 target faces. The examinee is then presented photographs of 48 faces, including the 24 target faces and 24 new faces. The examinee must identify each face, as either a target face or a new one. For Faces II, the examinee is again presented the 24 target faces and 24 different faces and is again asked to identify each face as either a target face or a new one.

Face perception and memory are regarded as special types of visuospatial processing and memory (Farah, Wilson, Drain, & Tanaka, 1998), and the

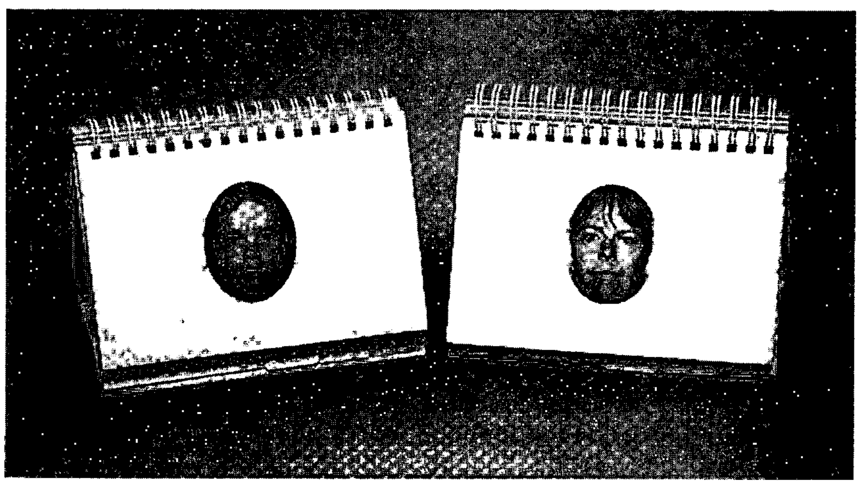

FIGURE 4. Similated items the Faces Subtest. As in the WMS-III, the faces are cropped so as to show a minimal amount of hair, and the individuals are shown with a neutral expression, not smiling. Half of the faces are of children and half are of adults. The photo has been created by the authors to appear like a WMS-III item. WMS-III faces subtest is a copyright protected by the Psychological Corporation.

ability to encode and remember new, unfamiliar faces has been well studied throughout the past century. McCarthy and Warrington (1990) concluded that "faces are special," and facial recognition appears to be "hardwired" in humans at the earliest ages (pp. 68–69). In fact, the developmental literature shows that even young babies have differential responses to faces as compared to other types of stimuli (Flin & Dziurawiec, 1989; Goren, Sarty, & Wu, 1975).

The earliest accounts in the behavioral neurology literature recognized difficulty in processing facial information (e.g., Charcot, 1883), and later the term *prosopagnosia* was derived to describe individuals who were deficient in the ability to process information about faces. This was described by Bodamer (1947) as "visual category [that is] the most profound and genetically the most primitive in our perception" (see McCarthy & Warrington, 1990, p. 56). Both the clinical and cognitive-perceptual literatures are quite clear and consistent that faces are processed differently by the brain than other visual stimuli.

There is some empirical evidence that some facial recognition tasks may show differential specificity to right-versus left-hemisphere damage (e.g., Milner, Branch, & Raomussen, 1966; Milner, Corkin, & Tenber, 1968; Warrington & James, 1967). Numerous researchers have made use of an idiosyncratic set of faces as stimuli for assessing facial memory (e.g., DeRenzi, Faglioni, & Spinnler, 1966; Milner et al., 1968; Saykin et al., 1992), some of which have been established and validated as facial memory tests, yet have not been widely utilized.

For these reasons, McCarthy and Warrington's (1990) conclusion that "faces are special" seems relevant and leads us to make the following conclusions about using faces as stimuli in a memory test:

1. Faces are encoded differently in the brain. Facial processing appears related to rather specific neuroanatomical substrates (i.e., fusiform gyri).
2. Faces represent a different type of visual stimulus. They cannot be verbally encoded as readily as other stimuli. Faces often possess greater detail (and potentially greater emotional coding—even unfamiliar faces) than nonface visual stimuli.
3. Differential perception and encoding of faces seems to be unique from a developmental perspective as well. Evolutionary theories also emphasize the differences between processing faces and all other visual stimuli.

The idea to include a facial memory subtest in the WMS-III occurred early in the revision process. Nelson Butters, who had established and led the initial advisory panel meetings, had recommended using this special type of stimuli for a visual memory task. Because memory for faces is a common task for all individuals, it was thought that the task would have ecological validity, which is generally lacking in most memory scales and had been one of the criticisms of the WMS-R visual memory subtests.

In addition, a precedent did exist for using faces as stimuli in perceptual tests (Benton, Hamsher, Varney, & Spreen 1983) and memory recognition tests (Benton et al., 1983; Warrington, 1984), and at the time the WMS-III was being developed, facial memory tasks were being included in other neuropsychological tests published by The Psychological Corporation (NEPSY, Korkman, Kirk, & Kemp, 1998; Children's Memory Test, Cohen, 1996). With these precedents in mind, a facial memory subtest seemed like an ideal addition to the WMS-III, and the faces "stimuli" were prepared.

The WMS-III Memory for Faces subtest was modeled after the Recognition Memory Test (Warrington, 1984), a widely used recognition test assessing memory for words and memory for faces. Performance on memory for faces has been shown to be disrupted following damage to the right temporal or right parietal lobes (Warrington, 1984). However, the Recognition Memory Test is thought by some to be limited in that it only requires recognition immediately following presentation. The WMS-III facial memory test in contrast, assesses recognition for the unfamiliar faces both immediately following presentation and after a delay.

Like the majority of the other subtests included on the WAIS-III and WMS-III, the subtests went through many revisions and experimental testing and appeared to be well received by both the standardization examiners and examinees. In the earliest stages of development, the items were pilot tested, cropping the pictures very tightly so that some more easily recognizable features (e.g., hair) would not be showing. In this way, examinees would not have easy cues that might facilitate recognition. Unfortunately, the effect was often frightening to the examinee because the faces did not look "human" with so many normal cues removed (Ledbetter, 1997).

Since the publication of the WMS-III, there have been some reports that questioned the inclusion of the faces subtest on the Visual Memory Index. Factor analyses by Millis, Malina, Bowers, & Ricker (1999) found that the subtest has much lower factor loadings on the WMS-III visual memory factor than Family Pictures, the other Visual subtest. This finding was confirmed by Price, Tulsky, Millis, and Weiss (2002), and replicated by Tulsky and Price (in press). Face recognition seems to be contributing unique variance to the visual memory factor. It is unclear if this result occurs because the Faces subtest is methodologically unique in that it only contains a recognition format without a recall component or if it is because Faces are "special stimuli"). Millis et al. (1999) advised that the Faces subtests should not be combined with the Family Pictures subtests in the construct of "visual memory," and this is one of the reasons why the visual memory factor was reconstituted in this joint factor project. Visual Reproductions did have a higher factor loading, along with Family Pictures, on this visual factor which also aided this decision (see Tulsky et al., Chapter 4, this volume, for a description of how the visual memory factor was reconstituted).

Family Pictures

When individuals present to a clinician with memory complaints, they rarely report that they cannot remember or draw abstract designs or geometric figures (such as in Visual Reproduction). Rather, they tend to report things like forgetting names of people they have met before; they report that they fail to remember things such as where they met someone, where they left their car keys, or where they parked their car. The Family Pictures subtest is a new subtest created for this revision of the WMS, developed in the hope of objectively assessing some of these presenting complaints.

The idea for Family Pictures arose as a result of the search for an ecologically meaningful way of assessing visual-spatial memory, without requiring complex "motor" responding (e.g., drawing). The original idea can be attributed to Robert Ivnik, who conceptualized the task and brought it to the WMS-III team during an advisory panel meeting in San Antonio in 1995. In attendence were Robert Bornstein, Gordon Chelune, Robert Heaton, Robert Ivnik, Mark Ledbetter (WMS-III Project Leader), Aurelio Prifitera, and David Tulsky. The goal was to examine spatial memory for visually presented, meaningful material while still allowing for verbal responding. In essence, most memories are verbally encoded, but often utilize the resources of spatial memory for cuing and accuracy. Even though photographs can be verbally encoded, they convey a great deal of spatial information that we learn and remember automatically. This task was designed to tap into this natural process by assessing spatial memory for visually presented material by asking people to identify spatial interrelationships of the elements in a photograph following examination. This recall is required without the examinee being told that they would have to remember the picture content or how the elements are arranged in relation to one another due to the fact that this is the typical situation in everyday memory.

The WMS-III team then developed and refined what we know today as Family Pictures for both the WMS-III and the Children's Memory Scale (Cohen, 1996). For Family Pictures I, examinees are shown a series of four scenes, all of which involve different characters from the same family. Examinees are told to remember as much as they can about the scene because they will be asked some questions later (Wechsler, 1997b). Each scene is presented for 10 seconds in a sequential fashion. After all the scenes are presented, the examinee is asked to recall which family member was in each scene, what each character was doing, and where the character was located. For Family Pictures II, the examinee recalls the same information without seeing the pictures again.

Tasks similar to family pictures do exist in other test batteries. For example, the Wide Range Assessment of Memory and Learning (WRAML; Adams & Sheslow, 1990) contains a picture memory subtest in which children examine a picture for 10 seconds, after which they see a comparable scene in which items were changed or added. They mark what is different with an "X." This procedure is repeated for four different scenes. Actually, such testing techniques have a long history. In 1911, Healy and Fernald introduced such a test as part of their battery (Healy & Fernald, 1911; Bronner, Healy, Lowe, & Shimberg, 1929), which required free recall of a picture of a butcher shop after a 10-second presentation. Free recall was followed by specific questions about the picture which are asked as "a kind of cross examination, which is calculated to bring out details of the picture which he may have forgotten, as well as his suggestibility" (Healy & Fernald, 1911, p. 22). Similarly, Fred L. Wells (1913) first reported on a test called Picture Postcards, which he later integrated into the Wells and Martin (1923) memory battery. Picture Postcards was a visual recognition test where the examinee is shown 12 "target" postcards depicting natural scenery from regions thought to be unfamiliar to the examinee. Then in a recognition format, the examinee is shown the 12 targets again mixed with 12 foils, and he or she has to indicate whether or not the picture had been seen previously.

Visual Reproduction

The Visual Reproduction subtest of the WMS-III requires individuals to study geometric figures for 10 seconds each and then draw the figures from memory. It was the only visual memory subtest in the original WMS (Wechsler, 1945). The task has a long history in intelligence and memory testing, and nearly all of the designs present in the WMS-III were taken directly from the Binet–Simon scale (Binet & Simon, 1905/1916a) or from the Army Performance Scale.

Binet and Simon (1905/1916a) developed and included their original task in the first edition of their intelligence test as a response to the need to classify school children as mentally retarded or normal. The subtest was entitled "Drawing a Design from Memory," and marked the first time that an examiner would present a stimulus card for 10 seconds and then, upon removal, ask the examinee to draw the design from memory. This administration procedure and the stimuli have been retained over the years (see Figure 5). Currently, we know the task as Visual Reproduction. In the Binet–Simon version, the examinee would be exposed to only one card with the two simple line drawings presented. These two designs still exist in the current WMS-III Visual Reproduction, although they are no longer presented together.

The drawing a design from memory task was utilized consistently throughout the history of mental testing. It was included in the Stanford Revision of the Binet Test (Terman, 1916) and then expanded upon for use in The Army Individual Examination, which tested recruits in World War I (Yerkes, 1921). For the Army tests, four "plates" containing new geographic designs, were actually created by Terman as alternative stimuli for the Stanford–Binet (Yerkes, 1921). Like the original Binet–Simon task, each design was exposed to the examinee for 10 seconds, after which it was removed and the examinee was to draw the design from memory (Yerkes, 1921).

The test that we know today as Visual Reproduction in the Wechsler Scales comes directly from these pioneering efforts, with little modification of the exact stimuli. Two of the four designs on the WMS came from the Design subtest of the Army Individual Performance Battery (Yerkes, 1921) and have continued to be included in the subsequent revisions (WMS-R and WMS-III; they are currently Designs B and C of the WMS-III). A third

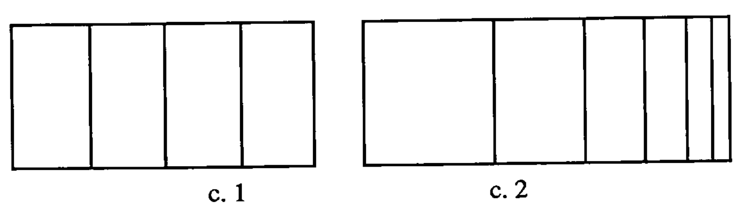

FIGURE 5. Item developed for the Army Performance Battery (Yerkes, 1921). In the public domain. Incoporated into Visual Reproduction of the Wechsler Memory Scale, Form II (see Stone, Girdner, & Albrecht, 1946; Stone & Wechsler, 1946).

design in the original WMS and in the subsequent WMS-R (4 circles) was taken directly from the Form Recognition subtest of Army Group Examination Beta (Yerkes, 1921), which was a simple matching to target visuospatial processing task.

Despite the similarity in testing procedures, however, there are significant distinctions between the tests. Notably, the WMS-III has expanded scoring criteria, with quite detailed criteria (Wechsler, 1997). In 1921, the scoring of this task was simple: full credit, half credit, or failure for each item (Yerkes, 1921). The administration procedures have also been expanded since the original WMS was published. Although the WMS included only an immediate recall procedure, and the WMS-R included immediate and delayed recall procedures, the WMS-III has added a series of optional tasks to allow examiners to test recognition memory, constructional praxis, and visual discrimination. Such tasks can be useful in determining the source of deficit observed on the immediate and delayed procedure, particularly for neurological patients who may have significant perceptual and/or visuoconstructional deficits. Finally, changes were made to the subtest items themselves. Specifically, the floor and ceiling levels in Visual Reproduction have been modified through the replacement of the more difficult WMS-R Design B (four circles) with a new item, (the WMS-III Design A), which is a simple line drawing of two "flags" facing opposite directions. The WMS-III subtest includes two new items—Design D, which is made up of two figures each composed of simple geometric shapes that are arranged side by side, and Design E which had been developed and used in the Binet-Simon test battery (Binet & Simon, 1905/1916a) and in the original WMS subtest (Wechsler, 1945).

An additional change in the revision of the WMS-R to the WMS-III is that this subtest is no longer a required component of the standard Visual Memory Index score. In the published form of the WMS-III, it is considered an optional subtest because of the length and multifaceted nature of this subtest. However, as will be presented in Chapter 4 (Tulsky, et al., this volume), this decision has been challenged on psychometric grounds, and the Visual Memory factor has been reconstituted to include Visual Reproduction in the six factor model.

Many other memory scales include Visual Reproduction-like tasks, but include multiple learning trials to help distinguish initial learning difficulties from true memory difficulties, (e.g., Heaton Figure Memory Test (Heaton et al., 1991) and the Brief Visual Spatial Memory Test-Revised (Benedict, Schretlen, Groninger, Dobraski, Shpritz, 1996)). The omission of such learning trials from the WMS-III Visual Reproductions subtests is

In later years, this test was again adapted by Corsi who used it in his doctoral dissertation (Corsi, 1972; Milner, 1971). The Spatial Sequential Learning Task (SSLT: Saykin, et al., 1992) was later developed as yet another adaptation to the task used by Corsi. The SSLT consists of a nine-block board similar to the apparatus presented by Corsi, on which the examinee is asked to reproduce a nine block tapping sequence demonstrated by the examiner (Saykin et al., 1992). The same tapping sequence is repeated until the participant correctly duplicates the entire series on two consecutive trials, or until 15 trails are completed. The number of trials to reach the criterion is the measure of the individual's ability to learn new spatial information. This adaptation of the original Knox (1914) task is designed to assess long-term memory capabilities, as opposed assessing spatial working memory, as in the WMS-III.

The Visual Spatial Span test was initially added to the WMS-R as a visual analog to the Digit Span test. A significant change from the WMS-R to the WMS-III is in the stimuli construction. The WMS-R Spatial Span stimuli consisted of squares printed on a two-dimensional card, whereas those on the WMS-III Spatial Span consist of a three-dimensional array of blocks (see Figure 7) with the block number to be visible only to the examiner. This change was made to facilitate administration and scoring.

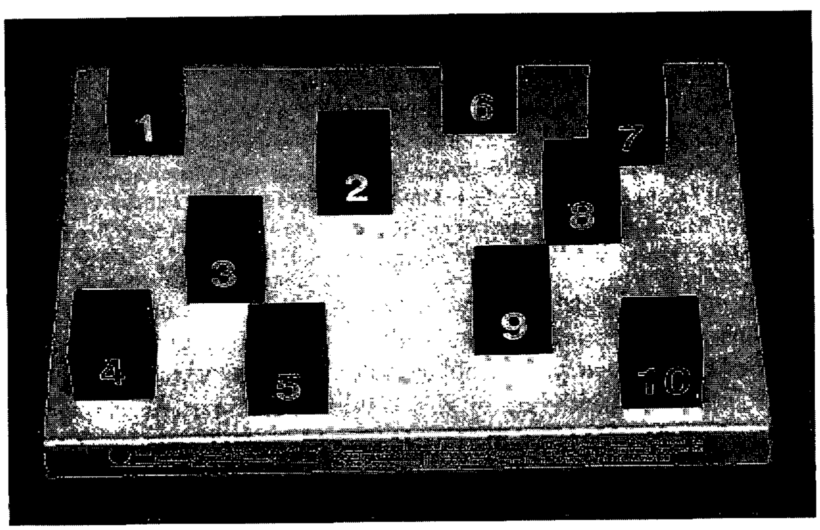

FIGURE 7. Photo of the Spatial Span board in the WMS-III. (Material copyright © 1991, 1997 The Psychological Corporation. Used with permission. All Rights reserved.)

Note: Photo displays the examiners perspective. The numbers are not visible to the examinee.

Optional Tasks of Basic Processes

Two of the optional WMS-III subtests (Information and Orientation and Mental Control) are primarily of value in providing a context for understanding poor performance on the primary memory indices.

Information and Orientation

The Information and Orientation subtest is an optional component of the WMS-III, and the scores are not incorporated into any of the Index scores. However, difficulties with this subtest would be an important factor to consider in interpreting poor performance on any of the index scores. The subtest represents a merging and expansion of two subtests from the original WMS: that is, in the original WMS Personal and Current Information was a separate subtest from Orientation. Regarding the former, Wechsler (1945) stated that the test discriminated very little (or not at all) between normal or near normal subjects. Instead, the subtest was useful in examining individuals with special defects (e.g., people with aphasia and people with dementia) who tend to score significantly lower on the test (p. 87). He indicated a similar rationale for inclusion of the Orientation subtest. In short, the WMS-III Information and Orientation subtest can be useful in determining whether a patient is oriented to a person, place, time, situation, and the like.

Mental Control

Mental Control is also an optional subtest in the WMS-III, but has appeared in some form on each of the earlier WMS versions. According to the WMS-III manual (Wechsler, 1997b), Mental Control measures the examinee's ability to retrieve overlearned information and to mentally manipulate that information. As formulated in the original WMS (as well as the WMS-R), it consisted of three tasks: (a) counting backwards from 20 to 1, (b) saying the letters of the alphabet, and (c) counting by threes. The latter task was dropped from the WMS-III version, but several additional tasks were added, including counting forwards from 1 to 20, stating the days of the week forward and backward, and stating the months of the year forward and backward. These tasks were added in order to assess an individual's ability to recite and manipulate overlearned information. Wechsler (1945) asserted that the utility of the original Mental Control Subtest was "primarily in cases of organic brain disease that are not too far gone but show deficits which would not be made evident by simple rote memory items" (p. 87–88).

When a patient has poor performance on Mental Control, it suggests that he or she may lack sufficient attentional abilities for adequate performance of the memory tests. That is, whereas impaired memory in the context of intact attention might be interpreted as possibly reflecting impairment in structures or brain systems underlying consolidation of episodic memory (such as the left or right hippocampus), when that same impairment is present in the context of significantly impaired attention, one is much more conservative in their interpretation of the memory findings. Thus, the documented poor performance on memory testing could be attributed to deficient information acquisition due to compromised attentional capacities. In sum, the interpretation of WMS-III performance may be altered significantly depending on the patient's basic attentional skills.

The Auditory Recognition Delayed Index

One of the primary goals of the WMS-III development, when possible, was the administration of a recognition procedure following administration of the delayed recall for Logical Memory, Verbal Paired Associates, Word List, and Visual Reproduction subtests. In the published test, an Auditory Recognition Delayed Index (ARDI) is computed by summing the *raw* recognition scores of Logical Memory and Verbal Paired Associates and then converting this raw total score to a standardized index score. The recognition index is based solely on verbal tasks. The Family Pictures subtest does not lend itself to a recognition procedure, and the Faces subtest is only based upon recognition (there is not a recall procedure for this subtest). Due to these differences in procedures, they were not included in a Recognition index. Moreover, since the Word Lists and Visual Reproduction subtests are optional, they are also not included in the Auditory Recognition Delayed Index, despite having a recognition procedure.

Since the publication of the WMS-III, it has become apparent that most normal individuals perform very well on this index, demonstrating a ceiling effect. Reports have suggested that even patients, some of whom have severe cognitive deficits, often score near perfectly on the measure. To better understand why this occurs, descriptive statistics on the Logical Memory Delayed Recognition Total Score, the Verbal Paired Associates Recognition Total Score, World List Recognition Total Score, and the Auditory Recognition Delayed Total Score have been computed (see Table 1). The distributions of all of these scores are highly skewed in a negative direction with the vast majority of the standardization sample obtaining a perfect score on

TABLE 1 Descriptive Statistics for Auditory Delayed Recognition Total Scores[a]

	Verbal Paired Associates Delayed Recognition total score	Logical Memory Delayed Recognition total score	Word List Delayed Recognition total score	Auditory Delayed Recognition total score
N	Weighted $n = 1250$ (unweighted $n = 1,032$)	Weighted $n = 1250$ (unweighted $n = 1,032$)	Weighted $n = 1250$ (unweighted $n = 1,032$)	Weighted $n = 1250$ (unweighted $n = 1,032$)
Minimum	12	13	0	26
Maximum	24	30	24	54
Mean	23.41	24.49	21.93	47.89
Median	24	25	23	49
Mode	24	27	24	51
Std. Deviation	1.5	3.5	2.6	4.3

these recognition variables. This results in a restriction in the range of the scores, making it very difficult to scale this index. In fact, in a normal sample, the Verbal Paired Associate Recognition component contributes very little additional variance to the recognition index, as 75% of the WMS-III standardization sample (or 958 out of the weighted *N* of 1250) obtained a perfect score of 24 points. This indicates that the entire variance of the Auditory Recognition Delayed Index in the standardization sample is made up of the Logical Memory Delayed Recognition score, which is also negatively skewed. However, the Logical Memory Delayed Recognition has more variance than the Verbal Paired Associates Recognition score and validates the argument put forth by clinicians; namely, the subtest and items were too easy and a ceiling effect was present.

Since the recognition scores are so skewed, traditional scaling techniques probably should not have been used to create an index score with a mean of 100 and standard deviation of 15. When nearly all individuals obtain a perfect or near perfect score, conversion to a scaled score metric is deceiving. A more appropriate scaling procedure for the recognition variables would have been a presentation of the percentile data similar to that reported in the WAIS-III Administration and Scoring Manual for the Digit Symbol optional procedures (Table A.11; Wechsler, 1997b). As pointed out in chapter 1 (Tulsky et al., Chapter 2, this volume), these scores can be extremely valuable clinically when deficits are noted. However, because the raw scores were associated with the lowest percentiles, the examiner is clearly using the recognition scores as deficit indices.

Table 2 presents an alternate way to determine if an individual scores as poorly as the lowest performers in the WMS-III standardization sample: Cumulative Percentages for the WMS-III recognition variables. This table shows cumulative percentages for the raw score values and is presented as an option in place of the Auditory Delayed Recognition Index that is described in the WMS-III manual. Even though the majority of normal individuals score quite high on this task, the recognition variables may, in fact, be clinically useful in separating groups. This remains an empirical question. With such skewed distributions, it is rare for someone to miss items. However, when this does occur, it becomes an unusual event, one that should be explored further to determine if the lower score is indicative of a cognitive deficit, or a suboptimal effort (e.g., malingering). Given the skewness of the data, high or average scores are not particularly meaningful. Using the cumulative percentages in Table 2, the recognition variables serve more as "pathonomic" indicators that "flag" potential deficits in recognition that should be explored further. For these reasons, we recommend that the

TABLE 2 Memory Recognition Scores: Cumulative Percentages Associated with Raw Scores[a]

	Cumulative Percentage	16–17	18–19	20–24	25–29	30–34	35–44	45–54	55–64	65–69	70–74	75–79	80–84	85–89
Logical Memory Delayed Recognition	1	20	19	17	17	18	14	19	17	16	16	14	13	14
	2	20	21	17	17	19	15	20	17	17	18	14	14	14
	5	20	21	19	18	20	16	20	18	18	18	15	16	15
	10	22	22	21	19	23	20	21	20	20	19	16	16	16
	25	24	24	24	24	24	24	24	22	22	21	20	19	18
	50	26	26	26	26	26	27	26	25	25	24	23	21	22
Verbal Paired Associates Delayed Recognition	1	23	22	21	18	20	18	20	19	20	18	12	12	12
	2	23	22	22	18	20	19	20	19	20	19	12	13	12
	5	23	23	23	23	22	21	21	21	21	20	15	17	14
	10	24	24	23	24	23	23	22	22	22	21	21	19	18
	25	24	24	24	24	24	24	23	23	23	23	23	22	21
	50	24	24	24	24	24	24	24	24	24	24	24	24	23
Word List Delayed Recognition	1	18	0	17	13	10	15	14	15	11	14	13	10	14
	2	19	18	17	13	16	16	14	17	16	16	13	11	14
	5	21	20	20	18	19	17	17	17	17	16	15	13	15
	10	22	21	20	20	20	18	19	18	18	17	16	15	16
	25	23	23	23	22	22	22	20	21	19	20	20	19	18
	50	24	24	24	23	23	23	23	23	21	22	22	21	20
Auditory Delayed Recognition	1	43	43	40	37	41	36	41	37	37	38	27	26	28
	2	43	45	40	37	42	36	41	40	40	40	27	31	29
	5	44	45	42	41	44	40	43	41	40	40	36	35	33
	10	46	46	45	43	45	44	44	43	42	42	39	37	34
	25	48	48	47	48	48	47	47	45	46	44	43	40	39
	50	50	50	50	50	50	50	49	48	49	47	47	45	45

recognition variables no longer be used in the same manner as the other index scores that are core to the WMS-III.

General and Immediate Memory

The publication of the WMS-III yields two global memory scores: an Immediate Memory Index (IMI) and the General Memory Index (GMI). The technical manual (The Psychological Corporation, 1997) describes the GMI of the WMS-III as "the best overall measure of the types of abilities that are critical to effective memory in day-to-day tasks" (p. 193). To give the most ecologically valid index of "memory," emphasis was placed on delayed memory rather than immediate, and in the WMS-III, unlike the WMS and WMS-R, the GMI is composed of the sum of delayed subtest scaled scores and auditory recognition delayed scores, rather than the immediate subtest scaled scores. Hence, the GMI is similar to the old Delayed Recall Index from the WMS-R.

Continuity between the WMS-R and WMS-III was believed to be preserved by including the Immediate Memory Index, a score that is relatively analogous in composition to the WMS-R General Memory Index (The Psychological Corporation, 1997), and the technical manual reports that the IMI is considered the best global indicator of immediate memory functioning. The technical manual also points out that at least on the surface, there are significant differences at the subtest level that preclude direct comparison between the WMS-III IMI and WMS-R GMI. However, continuity is preserved to the extent that the subtests on the WMS-III assess similar constructs as the Index scores on the WMS-R. In fact, the correlation between the WMS-III IMI and the WMS-R GMI is moderately high ($r = 0.62$; The Psychological Corporation, 1997), which suggests that there is some similarity between the measures, despite their differences.

More relevant to clinicians using the WMS-III is that the IMI and the GMI are NOT directly comparable. The WMS-III IMI is composed of the sum of the scaled scores on the four subtests that compose both the Auditory Immediate Index and the Visual Immediate Index scores. Like the IMI, the WMS-III GMI is composed of the sum of the scaled scores on the four subtests that compose both the Auditory Delayed Index and the Visual Delayed Index. However, the Auditory Recognition Delayed Index score is also included in the WMS-III GMI, which makes direct comparison of immediate and delayed memory less precise. Thus, the GMI is confounded due to the fact that both recall and recognition are included in the score, which is only made more complicated by the previously noted ceiling effects in the Auditory Recognition Delayed Index. To assist examiners with this comparison, Tulsky, Chelune, Price, & Millis (2003) have created a new *Delayed* Memory Index derived from the WMS-III standardization sample using only the auditory and visual delayed recall subtests. This new index includes only the delayed components of the same four subtests found in the IMI.

SUPPORT FOR THE WMS-III STRUCTURE

The initial support for the new WMS-III structure of eight primary index scores was provided in the *WAIS-III–WMS-III Technical Manual* (The Psychological Corporation, 1997). Confirmatory factor analyses (CFA) were used to support the construct validity of this structure by pitting six alternative models against each other and then evaluating the goodness-of-fit statistics. The original report claimed support for a five-factor model consisting of attention/concentration (or working memory), immediate verbal memory, immediate visual memory, delayed verbal memory, and delayed visual memory (see The Psychological Corporation, 1997). Shortly after publication of the WMS-III, Millis et al. (1999) attempted to replicate the factor structure using the covariance matrices that were reported in the *WAIS-III–WMS-III Technical Manual* and were unable to achieve convergence of the five-factor model. Instead, the analyses indicated an inadmissible solution due to a nonpositive definite covariance matrix, and the authors speculated that poor model fit was responsible for the inadmissible solution. Following up on this article, Price, Tulsky, Millis, and Weiss (2002) attempted to replicate the factor structure of the WMS-III using the standardization data set with the goals of (1) re-creating the results that had been obtained and (2) verifying that an error had been made in the original report (The Psychological Corporation, 1977). A warning message on the statistical output that had been accidentally overlooked when the results were first published. Price et al. then replicated their initial results in a restrictive cross-validation study using a sample of extra cases from the standardization sample that had been used in previous analyses by Tulsky and Price (in press). These replication studies confirm that the distinction between immediate and delayed memory cannot be supported, and the construct validity for this model of memory in this test battery has not yet been supported. Based upon these findings, the impact of these results and recommended changes that are advanced by the lead author of this chapter, as well as several of the editors of this volume, are described in Tulsky, et al., Chapter 4, this volume.

CONCLUSIONS

Although the WMS-III includes several new components, some that are unfamiliar and others that might have weaknesses in their design, this revision of the WMS was significant in its attempt to represent the more recent research in the field of memory assessment. Also the WMS-III has one of the largest and most representative standardization databases from which to assess memory and to make optimal clinical recommendations. Furthermore, the value of the powerful database that has been obtained through the co-norming procedure of the WMS-III with the WAIS-III cannot be overstated.

Kate Levine Kogan (1949), in writing the initial review of the Wechsler Memory Scale in the *Third Mental Measurement Yearbook* (reprinted as Item 8 in Appendix 2, this volume), stated that the most important use of the Wechsler Memory Scale is

> in conjunction with the Wechsler-Bellevue intelligence scale, since the memory quotient is designed to be directly comparable to the intelligence quotient. Use of the test permits intra-individual comparison of the patient's memory impairment with his loss in other intellectual functions rather than only a comparison with a general average or norm. (pp. 398–399)

Although the emphasis has moved away from IQ scores to factor scores (see Tulsky et al. chapter 4, this volume) and away from MQ scores to specific memory processes (as described earlier in this chapter), the most significant and meaningful advance in the production of the third editions of the WAIS-III and WMS-III was the use of a conorming methodology. However, during the development phase, the tests were developed in a relatively independent fashion utilizing two distinct project teams that were working on each of the revisions. Although there were attempts to integrate the use of the two tests (e.g., through base rates of discrepancies between these measures), they were still published as unique measures. Some more adventurous goals of combining subtests across the batteries to develop true joint factors of cognitive functioning were not realized in the initial publications, and, in this sense, the continuity of the two separate tests was maintained.

As with every assessment tool, work remains to be accomplished to improve our ability to most effectively apply the tools within the WMS-

III and to interpret the data to its maximal capacity. The WMS-III should therefore be viewed as a work in progress, with ongoing research in the field facilitating improved methodology and means of comparison.

In fact, Chapter 4 (Tulsky et al., this volume) outlines a new structure of subtests that can be given to assess cognitive functioning (a construct that includes both traditional intellectual measures, memory measures, working memory, and processing speed). Moreover, the follow-up chapters discuss new norms and methods of utilizing these tools in a more informed manner.

REFERENCES

Adams, W., & Scheslow, D. (1990). *Wide Range Assessment of Memory and Learning Manual.* Wilmington, DE: Jastak Associates.

Baddeley, A. D. (1976). *The psychology of memory.* New York: Harper & Row, Open University Set Books.

Bailey, P., Williams, F. E., & Komora, P. O. (1929). *The medical department of the United States Army in the World War: Volume X: Neuropsychiatry.* Washington: U.S. Government Printing Office.

Barr, W. B., Chelune, G. J., Hermann, B. P., Loring, D., Perrine, K., Strauss, E., Trenerry, M. R., & Westerneld, M. (1997). The use of figural reproduction tests as measures of nonverbal memory in epilepsy surgery candidates. *Journal of the International Neuropsychological Society, 3*, 435–443.

Benedict, R. H. B., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996). Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychological Assessment, 8*, 145–153.

Benton, A. L., Hamsher, K. S., Varney, N. K., & Spreen, O. (1983). *Contribution to neuropsychological assessment.* New York: Oxford University Press.

Benton, A., & Van Allen, M. (1968). Impairment in facial recognition in patients with cerebral disease. *Cortex, 4*, 344–358.

Binet, A., & Henri, V. (1895). La mémoire des phrases, [Sentence memory]. *L'Année Psychologique, 1*, 24–59.

Binet, A., & Henri, V. (1896). La psychologie individuelle, [Individual Psychology]. *L'Année Psychologique, 2*, 411–465.

Binet, A., & Simon, T. (1905/1916a). New methods for the diagnosis of the intellectual level of subnormals. In *The development of intelligence in children* (E. S. Kite, Trans.). Baltimore: Williams & Wilkins Company. (Original work published 1905)

Binet, A., & Simon, T. (1908/1916b). The development of intelligence in the child. In *The development of intelligence in children.* (E. S. Kite, Trans.). Baltimore: Williams & Wilkins Company. (Original work published 1908)

Binet, A., & Simon, T. (1911/1916c). New investigation upon the measure of the intellectual level among school children. In *The development of intelligence in children.* (E. S. Kite, Trans.). Baltimore: Williams & Wilkins Company. (Original work published 1911)

Black, F. W. (1973). Memory and paired associate learning of patients with unilateral brain lesions. *Psychological Reports, 33*, 919–922.

Black, F. W. (1974). The utility of the memory for designs test with patients with penetrating wounds of the brain. *Journal of clinical psychology, 30*, 75–77.

Boake, C. (2000). Historical note: Edouard Claparede and the Auditory Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology, 22*(2), 286–292.

Boardman, H. (1917). Psychological tests: A bibliography. *Bureau of Educational Experiments, 6*.

Bodamer, J. (1947). Die Prosopagnosia [Prosopagnosia]. *Archiv für Psychiatrie und Zeitschrift für Neurologie, 179*, 6–54.

Bornstein, R., & Chelune, G. (1988). Factor structure of the Wechsler memory scale revised. *The Clinical Neuropsychologist, 2*, 107–115.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist, 5(2)*, 125–142.

Bronner, A. F., Healy, W., Lowe, G. M., & Shimberg, M. E. (1929). *A manual of individual mental tests and testing*. Boston: Little, Brown, and Company.

Buschke, H. (1973). Selective reminding for the analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior, 12*, 543–550.

Buschke, H., & Fuld, P. A. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning *Neurology, 24*, 1019–1025.

Charcot, J. M. (1883). Un cas de suppression brusque et isolée de la vision mentale de signes et des objets (formes et couleurs). *Progrés Medicale, 11*, 568–571.

Chelune, G. J., & Bornstein, R. A. (1988). WMS-R patterns among patients with unilateral brain lesions. *The Clinical Neuropsychologist, 2*, 121–132.

Chelune, G. J., Bornstein, R. A., & Prifitera, A. (1990). The Wechsler Memory Scale-Revised: Current status and applications. In P. McReynolds & J. C. Rosen, et al. (Eds), *Advances in Psychological Assessment Vol.* 7. (pp. 65–99). New York: Plenum Press.

Claparede, E. (1919). Percentilage de quelques tests d'aptitude. [Percentiles of some aptitude tests] *Archives de Psychologie, 17*, 313–324.

Cochran, W. G. (1977). Sampling techniques (3rd ed.). New York: Wiley.

Cohen, M. J. (1996). *Children's Memory Scale*. San Antonio, TX: The Psychological Corporation.

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. Unpublished doctoral dissertation, McGill University, Montreal, Canada.

D'Elia, L. F., Satz, P., & Schretlen, D. (1989). Wechsler Memory Scale: A critical appraisal of the normative studies. *Journal of Clinical and experimental Neuropsychology, 4*, 551–568.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test—Adult Version*. San Antonio, TX: The Psychological Corporation.

DeLuca, J., Barbieri-Berger, S., Johnson, S. K. (1994). The nature of memory impairments in multiple sclerosis: Acquisition versus retrieval. *Journal of Clinical & Experimental Neuropsychology 16(2)*, 183–189.

DeLuca, J., Gaudino, E. A., Diamond, B. J., Christodoulou, C., & Engel, R. A. (1998). Acquisition and storage deficits in multiple sclerosis. *Journal of Clinical & Experimental Neuropsychology, 20(3)*, 376–90.

DeLuca, J., Schultheis, M. T., Madigan, N., Christodoulou, C., & Averill, A. (2000). Acquisition versus retrieval deficits in traumatic brain injury: Implications for memory rehabilitation. *Archives of Physical Medicine and Rehabilitation, 81*, 1327–1333.

De Renzi, E., Faglioni, P., & Spinnler, H. (1966). The performance of patients with unilateral brain damage on face recognition tasks. *Cortex, 4*, 17–34.

Ebbinghaus, H. 1. (1913). *Memory: A contribution to experimental psychology. (H. A. Ruger & C. E. Bussenius, Trans.)* New York: Teachers College, Columbia University, Bureau of Publications. (Original Work published 1895)

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.

Fastenau, P. S. (1996). An elaborated administration of the Wechsler Memory Scale—Revised. *The Clinical Neuropsychologist, Nov 10(4)*, 425–434.

Flin, R., & Dziurawiec, S. (1989). Developmental factors in face processing. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing*. Amsterdam: North Holland Publishing.

Goldstein, F. (1942). *After-effects of brain injuries in war.* New York: Greene and Stratton.

Goren, C. C., Sarty, M., & Wu, R. W. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*, 544–549.

Gorsuch, R. (2001). Chapter 3. *WAIS-III Canadian Technical Manual.* San Antonio, TX: The Psychological Corporation.

Haaland, K. Y., Price, L., & LaRue, A. (2003). What does the WMS-III tell us about memory changes with normal aging? *Journal of the International Neuropsychological Society*, 9, 89–96.

Halstead, W. C. (1947). *Brain intelligence: A quantitative study of the frontal lobes.* Chicago, IL: University of Chicago Press.

Healy, W., & Fernald, GM (1911). Tests for practical mental classification. *Psychological Monographs, 13 (54).*

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead-Reitan Battery: Demographic corrections, research findings, and clinical applications.* Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Paulsen, J. S., McAdams, L. A., Kuck, J., Zisook, S., Braff, D., Harris, M. J., & Jeste, D. V. (1994).Neuropsychological deficits in schizophrenia: Relationship to age, chronicity, and dementia. *Archives of General Psychiatry, 51*, 469–476.

Hebb, D. O. (1961). *The organization of behavior: A neuropsychological theory.* New York: Science Editions.

Herman, D. O. (1988) Development of the Wechsler Memory Scale—Revised. *The Clinical Neuropsychologist, 2*, 102–106.

Kaplan, E. (1988). A process approach to neuropsychological assessment. In T. J. Boll & B. K. Bryant (Eds.): Clinical neuropsychology and brain function. Washington: American Psychological Association.

Kaplan, E., Fein, D., Morris, R., & Delis, D. C. (1991). *WAIS—R as a Neuropsychological Instrument.* San Antonio, TX: The Psychological Corporation.

Kinsbourne, M., & Smith, W. L. (Eds.) (1974). *Hemispheric disconnection and cerebral function.* Springfield, IL: Charles C. Thomas.

Knox, H. A. (1914). A scale based on the work at Ellis Island for estimating mental defect. *Journal of the American Medical Association, 62*, 741–747.

Kogan, K. L. (1949). Review of the Wechsler Memory Scale. In O. K. Buros (Ed.), *The third mental measurement yearbook.* (pp. 398–399). New Brunswick: Rutgers University Press. (Reprinted in Appendix 2, this Volume).

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment manual.* San Antonio: The Psychological Corporation.

Ledbetter, M. (1997). *An Introduction to the Wechsler Memory Scale, Third Edition.* Video workshop from the Distance Learning Network. San Antonio, TX: The Psychological Corporation.

Loring, D. W. (1989). The Wechsler Memory Scale-Revised, or the Wechsler Memory Scale-Revisited. *The Clinical Neuropsychologist, 3,* 59–69.

Luria, A. R. (1966). *Higher cortical functions in man.* Oxford, England: Basic Books.

Luria, A. R. (1970). The functional organization of the brain. *Scientific American, 222,* 66–78.

Luria, A. R. (1973). *The working brain.* (B. Haigh, Trans.). New York: Basic Books.

McCarthy, R. A., & Warrington, E. K. (1990). *Cognitive neuropsychology.* San Diego: Academic Press.

Millis, S. R., Malina, A.C., Bowers, D. A., & Ricker, J. H. (1999). Confirmatory factor analysis of the Wechsler Memory Scale—III. *Journal of Clinical and Experimental Neuropsychology, 21*(1), 87–93.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27,* 272–277.

Milner, B., Branch, C., & Rasmussen, T. (1966). Evidence for bilateral speech representation in some nonrighthanders. *Trans Am Neurol Assoc., 91,* 306–308.

Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14 year follow-up study of H. M. *Neuropsychologia, 6,* 215

Naugle, R. L., Chelune, G. J., Cheek, R., Lüders, H., & Awad, I. (1993). Detection of changes in material specific memory following temporal lobectomy using the Wechsler Memory Scale – Revised. *Archives of Clinical Neuropsychology,* 7, 381–395.

Nebes, R. D. (1974). Hemispheric specialization in commissurotomized man. *Psychological Bulletin, 81,* 1–14.

Peterson, J. (1926). *Early conceptions and tests of intelligence.* Yonkers-on-Hudson, NY: World Book Company.

Phares, E. J. (1979). *Clinical psychology: Concepts, methods, and profession.* Homewood, IL: Dorsey Press.

Pribram, K. H., & Broadbent, D. E. (1970). *Biology of memory.* New York: Academic Press.

Price, L. R., Tulsky, D. S., Millis, S. R., & Weiss, L. (2002). Redefining the factor structure of the Wechsler Memory Scale-III: Confirmatory factor analysis with cross-validation. *Journal of Clinical and Experimental Neuropsychology, 24,* 574–585.

Prigatano, G. (1978). Wechsler Memory Scale: A selective review of the literature. *Journal of Clinical Psychology, 34,* 816–832.

The Psychological Corporation. (1997). *WAIS-III–WMS-III Technical Manual.* San Antonio, TX: Author.

Rausch, R., & Babb, T. L. (1993). Hippocampal neuron loss and memory scores before and after temporal lobe surgery for epilepsy. *Archives of Neurology, 50,* 812–817.

Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: Current status and applications.* New York: V. H. Winston & Sons.

Rey, A. (1941). L'examen psychologique en les cas d'encephalopathie traumatique. [Clinical teats in traumatic encepholopathy]. *Archives de Psychologie, 28,* 286–340.

Rey, A. (1964). *L'examen clinique en psychologie.* [Clinical exams in psychology]. Paris: Presses Universitaries de France.

Rourke, B. P. (1982). Central processing deficiencies in children: Toward a developmental neuropsychological model. *Journal of Clinical Neuropsychology, 4,* 1–18.

Russell, E. (1975). A multiple scoring method for the assessment of complex memory functions. *Journal of Consulting and Clinical Psychology, 43,* 800–809.

Russell, E. (1988). Renorming Russell's version of the Wechsler Memory Scale. *Journal of Clinical and Experimental Neuropsychology, 1,* 235–249.

Saling, M. M., Berkovic, S. F., O'Shea, M. F., & et. al. (1993). Lateralication of verbal memory and unilateral hippocampal sclerosis: Evidence of task specific effects. *Journal of Clinical and Experimental Neuropsychology, 15, 608–618.*

Saykin, A. J., Robinson, L. J., Stafaniak, P., et al. (1992). Neuropsychological changes after anterior temporal lobectomy: Acute effects on memory, language, and music. In T. L. Bennett (ed.), *The neuropsychology of epilepsy.* New York: Plenum Press, 263–290.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20,* 11–21.

Shaffer, G. W., & Lazarus, R. S. (1952). *Fundamental concepts in clinical psychology.* New York: McGraw-Hill.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23,* 723–733.

Squire, L. R. (1986). Mechanisms of memory. *Science, 232,* 1612–1619.

Squire, L. R. (1987). *Memory and brain* New York: Oxford University Press.

Stone, C. P., Girdner, J., Albrecht, R. (1946). An alternate form of the Wechsler Memory Scale. *Journal of Psychology, 22,* 199–206.

Stone, C. P., & Wechsler, D. (1946). Wechsler Memory Scale form II. New York: The Psychological Corporation.

Terman, L. M. (1916). *The measurement of intelligence.* Boston: Houghton Mifflin.

Tulsky, D. S., Chelune, G. J., Price, L. R., & Millis, S. R. (2003). Development of a new delayed memory index for the WMS-III. Manuscript in preparation.

Tulsky, D. S., & Ledbetter, M. F. (2000). Updating to the WAIS-III and WMS-III: Considerations for research and clinical practice. *Psychological Assessment, 12*(3), 253–262.

Tulsky, D. S., Ivnik, R. J., Price, L. R., & Wilkins, C. (2003). Assessment of cognitive functioning with the WAIS-III and WMS-III: Development of a 6-Factor model. In D. S. Tulsky et al. (Eds). *Clinical Interpretation of the WAIS-III and WMS-III.* (pp. 147–179). San Diego: Academic Press.

Tulsky, D. S., & Price, L. R. (in press). The Joint WAIS-III and WMS-III factor structure: Development and cross-validation of a 6-factor model of cognitive functioning. *Psychological Assessment.*

Tulsky, D. S., Saklofske, D. H., & Ricker, J. H. (2003). Historical overview of intelligence and memory: Factors influencing the Wechsler scales. In D. S. Tulsky et al. (Eds). *Clinical Interpretation of the WAIS-III and WMS-III.* (pp. 7–41). San Diego: Academic Press.

Tulving, E., & Donaldson, W. (1972). *Organization of memory.* New York: Academic Press.

Wallin, J. E. W. (1912). *Experimental Studies of Mental Defectives: A Critique of the Binet-Simon Tests and a Contribution to the Psychology of Epilepsy.* Baltimore: Warwick & York, Inc.

Warrington, E. K. (1984). *Recognition Memory Test Manual.* Windsor: NFER-Nelson.

Warrington, E. K., & James, M. (1967). An experimental investigation of facial recognition in patients with unilateral cerebral lesions. *Cortex, 3,* 317–326.

Wechsler, D. (1945). A standard memory scale for clinical use. *Journal of psychology, 19,* 87–95, 820.

Wechsler, D. (1917a). A study of retention in Korsakoff psychosis. *Psychiatric Bulletin,* 2, 403–451.

Wechsler, D. (1917b). *A study of retention in Korsakoff psychosis.* Unpublished Master's thesis. Columbia University.

Wechsler, D. (1939). *Measurement of Adult Intelligence.* Baltimore: Williams and Wilkins, Co.

Wechsler, D. (1944). The psychologist in the psychiatric hospital. *Journal of consulting psychology, 8,* 281–285.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology, 19,* 87–95.

Wechsler, D. (1987). *Wechsler Memory Scale—Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *WAIS-III administration and scoring manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *WMS-III administration and scoring manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. & Hartogs, R. (1945). The clinical measurement of anxiety. *Psychiatric Quarterly, 19,* 618–635.

Wells, F. L (1913). Experimental pathology of the higher mental processes. *Psychological Bulletin, 10(6),* 213–224.

Wells, F. L. (1927). *Mental tests in clinical practice.* Yonkers-on-Hudson, NY: World Book Company .

Wells, F. L., & Martin, H. A. A. (1923). A method of memory evaluation suitable for psychotic cases. *American Journal of Psychiatry, 3,* 243–258.

Whipple, G. M. (1915). *Manual of Mental and physical tests (in two parts) Part II: Complex Processes.* Second Edition. Baltimore: Warwick & York, Inc.

Williams, M. A. (1978). Clinical memory assessment. In P. McReynolds (Ed.), *Advances in psychological assessment, Vol 4.* San Francisco: Jossey-Bass.

Williams, M. A., Rich, M. A., Reed, L. K., Jackson, W. T., LaMarche, J. A., & Boll, T. J. (1998). Visual reproduction subtest of the Wechsler Memory Scale-Revised: Analysis of construct validity. *Journal of Clinical Psychology, 54(7),* 963–971.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review, 3 (Monograph Suppl. 16)*

Woodworth, R. S. & Wells, F. L. (1911). Association Tests. Being a part of the report of the committee of the American Psyhcological Association on the Standardization of procedure in experiemntal tests. *Psychological Monographs, 13* (whole no. 57).

Yerkes, R. M. (1921). *Memoirs of the National Academy of Sciences Volume XV: Psychological Examining in the United States Army.* Washington, D.C.: Government Printing Office.

Zielinski, J. J. (1993). A comparison of the Wechsler Memory Scale-Revised and the memory assessment scales, administrative, clinical, and interpretive issues. *Professional Psychology: Research and Practice, 24,* 353–359.