Pearson Clinical Assessment

19500 Bulverde Rd
San Antonio, TX, 28759
Telephone: 800 627 7271
www.PearsonAssessments.com

WMS–III to WMS–IV: Rationale for Change

Since the publication of the *Wechsler Memory Scale–Fourth Edition* (WMS–IV) in January 2009, we have received many questions regarding the changes introduced in the revision. What follows is an overview of the changes in question and the supporting rationale that led to the construction of the WMS–IV.

When we began developing WMS–IV, we gathered customer inquiries and criticisms from internal and external sources to help guide revision decisions. Although many decisions were relatively straightforward, in some areas, expert opinions were mixed, and additional sources of information were sought before making our decisions. As development proceeded, it became clear that the lack of concensus meant that some of our development decisions would not be accepted equally by some customers. We attempted to address the criticisms and complaints that our customers had with the WMS–III and balance the resulting effects of these changes for the benefit of our customers and their clients.

The chief customer concern with the WMS–III was related to the sensitivity of the assessment, particularly the Faces subtest and its associated norms, to brain injury and dementia. We implemented three changes to address this concern. One of the issues with Faces was that it appeared to be sensitive primarily to disorders associated with social perception impairment (e.g., schizophrenia, autism, Asperger's syndrome), which is consistent with areas of specialized processing for human faces in the brain.Faces, as presented in the WMS–III, had some psychometric issues with relatively low reliability (due to a high guess rate) and clinical sensitivity issues with floor problems (random responding resulted in a low average score). We attempted to fix the psychometric and clinical sensitivity issues by making the format more clinically sensitive. The result was a more clinically sensitive Faces subtest that had better psychometric qualities, in general, but was still more sensitive to clinical groups with social perception deficits. For this reason, it was not included in the core WMS–IV package. The subtest can be found in the social cognition section of the *Advanced Clinical Solutions* product. We understood that some customers

were satisfied with Faces as it existed in the WMS–III and apologize to those customers who were negatively affected by the change.

The second change we made to address WMS–III sensitivity issues was to add easy items to the Verbal Paired Associates subtest. In the WMS–III, there were floor and ceiling issues with this subtest, which made it more difficult to identify impaired association memory. The easy items introduced in the WMS–IV allowed for a better floor across all age ranges, improving clinical sensitivity. Also, having more items to recall allowed for better scaling of delayed recall (fewer jumps between scaled score points). For the older adult battery, the overall differences are much smaller. In the WMS–III, there were 4 learning trials and 8 items. The WMS–IV has 4 learning trials and 10 items so the time difference should be marginal.

The third change we made was to improve the normative data by screening standardization cases for possible cognitive impairment or suboptimal effort. Using the tools we developed for malingering detection and the results from the Brief Cognitive Status Exam, all cases were evaluated for possible effort or cognitive difficulties. This screening was not done on prior editions of the WMS.

Another major criticism we had from WMS–III customers was directed to the assessment of visual memory. Many customers told us that they did not use the Family Pictures subtest because they did not view it as a measure of visual memory. They felt Family Pictures primarily measured verbal memory, was redundant with Logical Memory, was not culturally representative of their patient population, and was confusing to older adults who often identified the *grandfather* or *grandmother* as the *father* or *mother*. In fact, many customers used Visual Reproduction even though it was not a core subtest. However, these customers did not like the scoring for the subtest in WMS–III and requested a scoring system that was shorter and easier. During WMS–IV pilot phase development, we produced a more culturally diverse edition of the Family Pictures subtest but still saw some character confusion among older adults. This could have been resolved with extensive exceptions within the scoring rules but that made scoring more complex. We decided to drop the Family Pictures subtest at the pilot stage, with the understanding that some of our customers who found the test satisfactory, would have preferred to have it remain part of the core battery.

Given that many of our customers did not like Family Pictures and were using Visual Reproduction, we made the decision to make Visual Reproduction a core subtest as long as we had scoring rules that, on average, could be completed in 5–10 minutes. We tried multiple scoring systems, and the version that was included in the WMS–IV was reliably scored in less than 10 minutes by our scorers and had psychometric and clinical sensitivity data comparable to longer scoring systems. We also felt that for younger subjects, Visual Reproduction alone may not be sensitive enough to capture visual memory deficits. We needed to develop an additional visual memory subtest.

We undertook development of a visual memory task understanding that there are no *pure* visual memory tests and there will be confounding factors no matter how we approached the test (there are also no pure verbal memory tests since all verbal memory tests assume some level of language functioning). In reviewing the literature, it was clear that object memory (particularly, nameable objects) operates differently than spatial memory. This is not surprising since visual information is processed in two visual streams: a dorsal stream, which evaluates the spatial location of objects and is related to occipital-parietal lobe functioning, and a ventral stream, which evaluates visual details and is related to occipital-temporal lobe functioning (attaching language to a visual object). We felt that the best chance of a subtest being visual rather than verbal would be to develop items that were hard to verbalize and required processing of both spatial and detail elements. The immediate reaction of the examinee to the items on the Designs subtest is that they can't do it. This is because the language centers of the brain are overwhelmed by the stimuli; however, because there are item difficulty gradients, most subjects can get at least some points. If the subtest had used nameable objects or simply used designs, the language centers of healthy adults would be able to attach labels and turn the test into a visual-verbal association test. Patients with left temporal epilepsy/lobectomy, would be put at a disadvantage compared to normal controls. In addition, the right temporal epilepsy/lobectomy patients would be able to use language processing to facilitate performance on this test. The purpose is not to inform the surgeon where to operate but to track the impact of left or right surgery or injury on memory functions. Our results found that the right temporal lobectomy group had difficulty with spatial memory but not memory for details, and the left temporal group did well on both. In addition to providing better information about left-right differences than the previous edition, this version of the Designs subtest is one of the more sensitive tests for indentifying

memory deficits in patients with moderate to severe TBI. The Designs subtest also correlates highly with ratings of daily functioning after TBI. Overall, the subtest has a very good floor and ceiling, good reliability, and good clinical sensitivity, and meets the need for a visual memory subtest for younger subjects.

Our polling of WMS–III customers also indicated that they rarely used supplemental measures, other than Visual Reproduction, so most of these were dropped or incorporated into the cognitive screener. We also found that there were issues with the WMS–III Working Memory Index. The index was comprised of one subtest and overlapped with the WAIS–III Working Memory Index, making the two indexes statistically dependent. This created problems when comparing the two indexes because the standard significance level calculations assume the scores are independent. Also, since the WMS–III index was made up of a visual and auditory measure, differences between the WAIS–III and WMS–III indexes could be driven by one score and not by working memory as a general construct. The decision was made to separate the WAIS–IV and WMS–IV indexes, where the WAIS–IV Working Memory Index would include auditory working memory and the WMS-IV Working Memory Index would include visual working memory.

Because the WMS–III had only one visual working memory measure, it was clear that at least one more visual working memory measure would be required. In addition, we wanted to develop tasks that required some mental manipulation of visual information. Although Spatial Span is a good measure of visual spatial span, it does not require a lot of mental manipulation. Two tasks were developed as analogs to the WAIS–IV Arithmetic and Digit Span subtests. The two subtests would also reflect the two streams of vision noted previously. The Spatial Addition subtest has psychometric qualities very similar to Arithmetic and correlates with academic functioning similar to Arithmetic. The Symbol Span subtest has psychometric qualities similar to the WAIS–IV Digit Span subtest and correlates with academic functioning in a similar manner. It was not expected that examiners would need to use both indexes on a regular basis because they are highly correlated and function similarly. It was expected that examiners would use the Visual Working Memory index only when they thought that the WAIS–IV Working

PEARSON

Memory Index was not an adequate measure of working memory due to language processing problems or difficulties with arithmetic skills, or if there was a specific hypothesis about Visual Working Memory.

We have also heard complaints about the administration time being long. We have some suggestions to offer in response. First, the WMS–IV contains many optional procedures that are used to answer specific questions. The recognition trials are optional and should be used only when there is a clinical question regarding encoding versus retrieval deficits. The VPA Delayed Free Recall task is also optional and should be used when a more general word recall measure is desired rather than an associative memory measure. The administration of every task of every subtest of the WMS–IV would be needed only for unusual cases. Second, not using the WMS-IV Working Memory subtests can save a substantial amount of time. The need for a second delay is caused by the need to administer the Symbol Span subtest after Visual Reproduction plus the administration time required for Spatial Addition. When not using these subtests, the memory tests can be administered sequentially with only a single delay.

We recognize that time and billing constraints challenge all psychologists to provide the most comprehensive, yet cost-efficient evaluation, and we are working towards providing solutions to further meet those needs.

We hope that this information has provided insight into the development process of the WMS–IV and the decisions and trade-offs that were made to best serve our customers and their clients. If you have additional questions regarding the WMS–IV or any of our products, you can always contact us at 800.627.7271.