

# An Evaluation of the CVLT-II Substitution on the WMS-IV Amongst a Mixed Clinical and Nonclinical Population

Jessica Holster, Ph.D., Christine Corsun-Ascher, Psy.D., Traci Olivier, B.S., Charles Golden, Ph.D.

## Abstract

*The Wechsler Memory Scale (WMS) is one of the most commonly used psychological measures in clinical practice and research. The newly revised fourth edition (WMS-IV) has recently been published, and introduced a novel substitution rule in which scores from the California Verbal Learning Test, Second Edition (CVLT-II) may replace WMS-IV subtest scores for Verbal Paired Associates (VPA) I and II subtests for the ease of administration. Yet, the validity of this substitution remains unclear. The purpose of the paper was to determine if the replacement of CVLT-II scores for WMS-IV scores affects WMS-IV index scores, thus affecting subsequent analyses and diagnostic decision making, within a mixed clinical and nonclinical population. Correlations between the measures and paired t-tests were conducted between the WMS-IV Index Scores with and without the CVLT-II substitution. Results indicated that performance on the Delayed Memory Index (DMI) was indicated to be significantly better when scored using WMS-IV VPA II relative to CVLT-II Long Delay Free Recall. This is theorized to be attributed to the cued-recall paradigm underlying VPA administration, and mandated visual interference tasks on the WMS-IV. The authors suggest that caution should be taken when implementing this score substitution.*

## Introduction

Memory functioning is often a core cognitive ability evaluated by clinicians during clinical and neuropsychological evaluations. The variability in memory impairment across neurological and psychiatric disorders makes the assessment of specific components of memory functioning critical to diagnostic decision-making (Delis, Cullum, Butters, Cairns, & Prifitera, 1988). One of the most popular memory test batteries is the Wechsler Memory Scale (WMS). First developed in 1945, the WMS has grown to become one of the most well-known and utilized measures of memory (Loring & Bauer, 2010). The fourth edition of the WMS (WMS-IV; Wechsler, 2009), was introduced in 2009, with such modifications as updated normative data, newly developed subtests and items, and flexible scoring guidelines.

One of novel scoring rules of this updated measure involves the optional substitution of WMS-IV subtests with scores from the California Verbal Learning Test, Second Edition (CVLT-II; Delis, Kramer, Kaplan & Ober, 2000). Specifically, the WMS-IV manual (Wechsler, 2009) denotes that the CVLT-II Total Trial 1-5 can be substituted for Verbal Paired Associates (VPA) I

in both the Auditory and Immediate Memory Indices. Further, the CVLT-II Long Delay Free Recall is suggested to be a reliable substitute for VPA II, contributing to both the Auditory and Delayed Memory Indices. This replacement is suggested to shorten test administration and be optimal for use amongst low-functioning clients. Despite the suggested benefits, the ability for the CVLT-II to accurately substitute for WMS-IV VPA I and II subtests remains unclear.

In order for one task to accurately substitute for another task, it should be assumed that the two are measuring similar underlying cognitive abilities. Yet, there are multiple construction differences between WMS-IV VPA subtests and the CVLT-II that implicate differences between what each is evaluating. Possibly the most noticeable difference regards how participants may initially appraise the task. Wechsler (2009) specifies that some participants may have difficulty understanding VPA, and with such participants, the CVLT-II may be preferred. Thus, it is implied that the CVLT-II may be easier to approach and consequently complete than VPA. This may not only be relevant for populations who demonstrate difficulty with complex directions, but those who are bilingual or whose primary language is not English (Restrepo & Gutiérrez-Clellen, 2004). Thus, if one test is easier to understand than the other, it may be assumed that the two tests are placing different demands on the participant from the onset of administration.

In addition, the two tests appear to differ regarding the underlying approaches to memorization. For instance, the way in which the words are initially encoded may differ based on the degree of association amongst the words chosen for each test (Wible et al., 2006), order of stimulus presentation (Albuquerque, Loureiro, & Martins, 2008), and feedback given by the examiner during testing (i.e., corrective feedback and cuing; Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Butler, Karpicke, & Roediger, 2008; Pashler, Cepeda, Wixted, & Rohrer, 2005; Smith & Kimball, 2010). Delayed recall may also differ due to such initial encoding differences as well as the differing level of interference before the delayed task between the two tests (Fernandes & Grady, 2008). Generally, the CVLT-II has been suggested to be more of a learning task than VPA, providing less organization and meaning cues than the latter (Golden, Espe-Pfeifer, and Wachslar-Felder, 2000).

Normative differences are also noted. In order for tests to be adequately compared, they should possess similar normative samples (Golden et al., 2000). The CVLT-II norms are based on both gender and age (Delis et al., 2004), while the WMS-IV norms are based only on age (Wechsler, 2009). Research has documented that on the CVLT-II females possess greater use of semantic clustering strategies, are less susceptible to interference, and are less skilled at encoding contextual cues (Lengenfelder, Kemenoff, Kramer, & Delis, 2000). In addition, males show better immediate recall from the beginning and end of the list, while females demonstrate better recall from the middle portion (Lengenfelder et al., 2000). Since gender was not incorporated in the construction of the WMS-IV norms, this negatively impacts the ability to make accurate comparisons between the measures. Further, the WMS-IV age bands are smaller than the CVLT-II, thus implicating that the WMS-IV norms may be more sensitive to the changes in semantic memory often associated with the aging process (Byrd, 1984; Lovden et al., 2004).

Multiple studies have evaluated the relationship between the first edition of the CVLT (Schear & Bruce, 1989) and revised edition of the WMS (WMS-R; Delis et al., 1988), generally documenting low to moderate correlations. Early studies often did not illustrate promising results regarding the ability for one test to substitute for the other. For instance, Randolph et al. (1994) found that that the same individual taking both the CVLT and WMS-R would look more impaired on the CVLT than the WMS-R. Subsequent studies continued to document discrepant results between the two tests, such as enhanced performance on the WMS-R compared to the

---

CVLT amongst clinical Norwegian patients (Bosnes & Bjorn, 2003), higher sensitivity on the first edition of the WMS to memory dysfunction than the CVLT amongst schizophrenic relatives (Trandafir, Meary, Schurhoff, Leboyer, & Szoke, 2006), and different demands on semantic processing and memory organization between the WMS-R and CVLT, as documented within a sample of epilepsy patients (Helmstaedter, Wietzke, & Lutz, 2009). Such differences were generally attributed to poor normative samples with the earlier CVLT and WMS editions, but were suggestive of several underlying differences between the two tests.

Regarding the most recent editions of the two tests, low to moderate correlations have been documented by Wechsler (2009). Most notably, a correlation of 0.54 was reported between the CVLT-II Total Trial 1-5 and WMS-IV VPA I subtest, the highest correlation found between measures. Further, a correlation of 0.49 was found between Long-Delay Free Recall on the CVLT-II and the WMS-IV VPA II subtest. Despite only moderate correlations, the authors determined that the VPA subtest from the WMS-IV and CVLT-II share similar verbal content, response processes, semantic association, and task demands.

Overall, it appears from the literature that the replacement of the WMS-IV subtests with CVLT-II scores is questionable. In this study, we attempted to assess if the substitution of scores from the CVLT-II for WMS-IV VPA I and II would alter index scores on the WMS-IV, consequently affecting interpretation of memory performance. First, it was hypothesized that only moderate correlations would be found between the CVLT-II and WMS-IV, particularly between CVLT-II Trials 1-5 and Long Delay and WMS-IV VPA I and II, respectively. Second, differences regarding the Auditory Memory Index (AMI), Immediate Memory Index (IMI), and Delayed Memory Index (DMI) were hypothesized with CVLT-II score substitution. Third, it was hypothesized that the largest difference would be found regarding the AMI, where scores from both the immediate and delayed recall tasks are substituted for VPA I and II.

## **Methods**

**Participants.** Participants included 70 adults from a mixed clinical ( $N = 24$ ) and normal ( $N = 46$ ) population. Participants were predominantly female (54.3%) and right-handed (85.7%), with a mean age of 30.24 years ( $SD = 11.01$ ) and education level of 15.26 ( $SD = 2.39$ ) years. Participants were predominantly Caucasian (64.3%), with 21.4% endorsing Hispanic origin, 5.7% endorsing Black origin, and 8.6% endorsing some other primary ethnic background. Diagnoses regarding the clinical sample were made by a licensed clinical psychologist based on behavioral observations, records, and testing results according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000). Of the 24 clinical participants, diagnoses included Learning Disorder (37.5%), Generalized Anxiety Disorder (16.7%), Personality Disorder (16.7%), ADHD (12.5%), Major Depressive Disorder (8.3%), Adjustment Disorder (4.2%), Obsessive Compulsive Disorder (4.2%), Borderline Intellectual Functioning (4.2%), Cognitive Disorder-Not Otherwise Specified (4.2%), Alcohol Dependence (4.2%), Mild Mental Retardation (4.2%), and Dysthymic Disorder (4.2%). Several (33%) met criteria for more than one DSM-IV-TR diagnosis. Diagnosis was deferred on Axis I and II regarding three (12.5%) of the clinical participants.

**Materials.** The CVLT-II (Delis et al., 2004) is an individually administered word list task designed to assess the ability to learn and remember verbally presented information amongst individuals between the ages of 16 and 89. The CVLT-II involves the oral presentation of two word lists, each containing 16 items, which can be broken down into four words from each of

---

four semantic categories. Initially, there are five trials of presentation of the first list (i.e., List A), in which the individual is asked to recall as many words as possible. An interference list (i.e., List B) is then presented, followed by a free recall and short-delay cued recall trial of List A. After a 20-minute delay, in which only nonverbal testing may be administered, the individual is then asked to again remember as many words as he or she can from List A via free and cued recall trials. This is followed by a yes/no recognition trial and, 10 minutes thereafter, a forced choice recognition trial.

The WMS-IV Adult Battery (Wechsler, 2009) is a test of an individual's visual and verbal memory functioning administered to individuals between the ages of 16 and 69. The measure contains subtests that may be grouped together, amounting to five Index Scores: the AMI (i.e., Logical Memory I and II, VPA I and II), Visual Memory Index (VMI; i.e., Designs I and II, Visual Reproduction I and II), Visual Working Memory Index (VWMI; i.e., Spatial Addition, Symbol Span), IMI (i.e., Logical Memory I, VPA I, Designs I, Visual Reproduction I), and DMI (i.e., Logical Memory II, VPA II, Designs II, Visual Reproduction II).

The current study focused on the WMS-IV VPA I and II subtests and related indices (i.e., AMI, IMI, and DMI). VPA is composed of a list of 14 word pairs that are both novel ( $N = 7$ ) and semantically related ( $N = 7$ ). On VPA I, the list of word pairs is read to the individual. Then the person is orally presented with one of the words from the pair and asked to recall the associated word from the list. Five seconds is allotted for a response before corrective feedback. This trial is followed by three identical trials of presentation and cued recall. VPA II is administered about 20-30 minutes after VPA I. During VPA II, the first word of each word pair is read out loud by the examiner, and the individual is prompted for the associated word. He or she is allotted 10 seconds to respond. No corrective feedback is given. There is a subsequent recognition trial and word recall task, which were not administered in the present study.

**Design and Procedure.** The current study included both clinical and normal participants. Clinical participants were recruited through a university-affiliated mental health clinic by student clinicians. Alternatively, normal participants were recruited by graduate level students from South Florida by means of snowball sampling and fliers distributed throughout the South Florida area. Both clinical and normal participants volunteered to permit the use of their data for research purposes, and reviewed and signed an IRB approved consent form detailing the procedures, risks, and benefits of participating in the study.

Participants were then screened regarding sample-specific inclusionary/exclusionary criteria prior to testing. Clinical participants were required to be seeking neuropsychological testing for cognitive, attentional, memory, or emotional difficulties, be between the ages of 18 and 69, and speak English. Those under legal guardianship were excluded from participation. Regarding normal participants, inclusion criteria required individuals to be between the ages of 18 and 69 and endorse a first language and predominant language of English. Exclusion criteria included being under legal guardianship, and having a history of neurological disease (e.g., Parkinson's Disorder, Alzheimer's Disease, Pick's Disease), treated Axis I and Axis II disorders (e.g., Major Depression, Schizophrenia, Substance Dependence, Borderline Personality Disorder, Mathematics Disorder, Mental Retardation), terminal illness that may lead to cognitive deterioration (e.g., cancer, human immunodeficiency virus (HIV)), and/or brain injury (e.g., anoxia, hemorrhage, aneurism, traumatic brain injury). In addition, if overall testing results were determined by the licensed psychologist to demonstrate an abnormal profile (e.g., presence of a Learning Disorder), the participant's data was consequently excluded from data analysis.

---

If no exclusionary criteria were met participants were subsequently administered a standard battery of neuropsychological tests. Standard administration was followed for the CVLT-II and WMS-IV. The order of administration was random, with some individuals being administered the CVLT-II prior to the WMS-IV and vice versa. In total, testing for clinical participants totaled approximately 18 hours, which was broken into multiple three-hour testing session blocks. Normal participants received an abbreviated battery, totaling approximately six hours over the course of two three-hour testing sessions.

The WMS-IV AMI, IMI, and DMI indices were scored both with and without the substitution of CVLT-II scores for VPA I (i.e., CVLT-II Total Trial 1-5) and VPA II (i.e., CVLT-II Long Delay Free Recall), using the score conversion depicted in the WMS-IV manual (Wechsler, 2009).

**Results**

All analyses were conducted at the .05 level of significance. Correlations were calculated involving WMS-IV and CVLT-II scores using Fischer’s *z* transformation. Weak to strong linear associations were found between WMS-IV and CVLT-II variables. In particular, moderate positive linear associations were found between CVLT-II Trial 1-5 and WMS-IV VPA I,  $z' = .58$ , and CVLT-II Long-Delay Free Recall and VPA II,  $z' = .51$ . The means and standard deviations for the CVLT-II and WMS-IV scores can be found in Tables 1.0. Correlations at the subtest and index level are shown in Table 2.0.

Table 1  
*Descriptive of CVLT-II and WMS-IV Scores*

	Score	<i>M</i>	<i>SD</i>
CVLT-II	Trials 1-5 Correct	56.31	11.03
	Trial 1 Correct	7.41	2.59
	Trial 5 Correct	13.67	2.11
	Trial B Correct	7.19	2.05
	Short-Delay Free Recall Correct	11.87	3.14
	Short-Delay Cued Recall Correct	12.69	2.82
	Long-Delay Free Recall Correct	12.63	3.28
	Long-Delay Cued Recall Correct	13.14	2.96
	Free-Recall Intrusions	3.01	3.24

	Cued-Recall Intrusions	1.40	1.56
WMS-IV	VPA I	42.43	10.10
	VPA II	12.76	2.25
	AMI	47.06	9.01
	IMI	44.43	7.90
	DMI	45.43	7.75

Note. *M* = Mean. *SD* = Standard Deviation.

Table 2

*Correlations Between the WMS-IV and CVLT-II Subtest Scores and Indices\**

CVLT-II Scores	WMS-IV Score				
	VPA I	VPA II	AMI	IMI	DMI
Trials 1-5 Correct	.58	.61	.68	.70	.86
Trial 1 Correct	.35	.36	.42	.34	.44
Trial 5 Correct	.55	.62	.77	.78	.92
Trial B Correct	.45	.43	.53	.51	.56
Short-Delay Free Recall Correct	.40	.49	.51	.57	.72
Short-Delay Cued Recall Correct	.47	.51	.59	.68	.79
Long-Delay Free Recall Correct	.49	.51	.60	.61	.72
Long-Delay Cued Recall Correct	.47	.48	.58	.61	.74
Free-Recall Intrusions	-.14	-.07	-.08	-.07	-.18
Cued-Recall Intrusions	-.25	-.16	-.19	-.13	-.23

\*Correlations were computed using Fischer's *z* transformation.

Paired samples t-tests were executed regarding the AMI, IMI, and DMI index scores to compare index scores with and without the CVLT-II substitution (i.e., CVLT-II Total Trial 1- 5 score for VPA I and CVLT-II Long Delay Free Recall score for VPA II). No significant results were found regarding the AMI,  $t(69) = 1.59, p = .349$ , or IMI Index Scores,  $t(69) = .94, p = .076$ . Regarding the AMI Index Score, participants generally performed the same with ( $M = 108.21, SD = 13.99$ ) and without ( $M = 110.34, SD = 15.53$ ) CVLT-II scores substitution. Results were similar for the IMI Index with ( $M=105.94, SD=13.54$ ) and without ( $M=106.87, SD = 16.62$ ) CVLT-II score substitution.

A significant difference was found regarding the DMI Index Score,  $t(69) = -2.02, p = .047$ , when comparing indices with and without CVLT-II score substitution. Participants generally appeared to perform better without the CVLT substitution ( $M = 109.46, SD = 13.60$ ) than with the score substitution ( $M = 108.43, SD = 14.65$ ).

## **Discussion**

The purpose of the present study was to address if the CVLT-II Trial 1- 5 and Long-Delay Free Recall scores could be substituted for VPA I and II scores on the WMS-IV, respectively. If the CVLT-II scores are an accurate substitution for VPA, measuring similar underlying cognitive abilities, then it would be assumed that performance by participants would be similar between the two tasks. Consequently, strong associations should be demonstrated between the related subtests on the CVLT-II and WMS-IV and non-significant differences between WMS-IV AMI, IMI, and DMI indices should be found when index scores generated with and without score substitution are compared.

Correlations between the related subtests indicated moderate positive linear associations between variables. Such results are similar in comparison to correlations obtained by Wechsler (2009), but generally indicated that the CVLT-II and the WMS-IV VPA task share some similarity regarding the underlying constructs. Specifically, regarding both the short-term and long-term memory tasks, only about 34% and 26% of the variability in performance was shared between tasks, respectively. Thus, approximately 66% and 74% of the variability in short-term and long-term tasks, respectively, were explained by some other means (e.g., different task demands, the use of cuing trials versus learning trials, and/or the number of words involved in each trial). Stronger associations between related CVLT-II and WMS-IV variables would be expected given that one test would accurately substitute for another. Such results offer support for the lack of interchangeability between the two measures.

Index score comparisons further indicated that this score substitution is not effective with regards to delayed memory. When CVLT-II Long Delay Free Recall was substituted for WMS-IV VPA II according to the score substitution guidelines, detailed in Wechsler (2009), performance regarding the WMS-IV DMI significantly declined; participants generally performed better on the WMS-IV VPA delayed memory task than on the CVLT-II. Such results are in contradiction to what would be expected given scoring guidelines detailed in Wechsler (2009). When clinicians implement the CVLT-II substitution it may be assumed that they risk misinterpretation regarding the subject's ability to encode, store, and relay information upon delay. Further, these results suggest that WMS-IV VPA and the CVLT-II are measuring two different underlying cognitive abilities.

Differences in performance between WMS-IV VPA and the CVLT-II may be attributed to the cued recall paradigm utilized in the WMS-IV; the first word from each of the 14 pairs is

---

provided by the examiner and the individual is subsequently asked to recall the associated word from memory. In VPA I and II, words that compose each pair are both novel and semantically related. The semantically related word pairs should be encoded readily; the first word of the semantically related word pairs should naturally cue retrieval and recall of the matching word pair. Research has suggested that each word stimulus should initiate activation of a series of networks of neighboring semantic associates (Collins & Loftus, 1975; Meyer & Schvaneveldt, 1971; Schacter & Buckner, 1999). Thus, when comparing VPA to the CVLT-II, it can be assumed that word associations on VPA should be easier to encode based on the priming given throughout the task and short inter-stimulus intervals for at least half of the word pairs. The remaining half require the person to develop a learning strategy (e.g., the “ball” can bounce over the “moon”), which may again follow similar activation patterns.

Additionally, corrective feedback is given on VPA, but not on the CVLT-II. Corrective feedback assists in initial storage of information and enhances long-term retention (Bangert-Drowns et al., 1991; Smith & Kimball, 2010). Research has furthermore suggested that feedback may also strengthen initially correct responses (Butler et al., 2008), and correct initial errors in learning (Pashler et al., 2005). Learning theory suggests that feedback functions as a reinforcer of correct responses, particularly when given shortly after the response (Skinner, 1953). Thus, immediate feedback is likely to reinforce correct responses, consequently increasing retention (Smith & Kimball, 2010). Kang, McDermott, and Roediger (2007) determined that corrective feedback may be more effective in enhancing learning and memory more than repeating a series of facts, which is similar in nature to the repetition of words on the CVLT-II.

Penultimately, the CVLT-II possesses a unique interference task that may decrease performance upon delayed recall. After immediate recall trials of List A on the CVLT-II, a second list of items (i.e., List B) is presented, with two of the four categories overlapping with List A (i.e., animals and vegetables). The presentation of List B is followed by an immediate recall trial. With inclusion of the second list, the examiner is afforded the ability to assess interference on recall, both proactive (i.e., from List A into List B recall) and retroactive (i.e., from List B into delayed recall and recognition trials). Yet, the absence of a similar verbal interference list from the WMS-IV VPA subtests may make the retrieval of words during delayed recall easier. Research has suggested that interference tasks have a significant effect on memory (Fernades & Grady, 2008). This is particularly relevant for tasks in which individuals can benefit from strategic processes such as chunking or organization (i.e., categorized word lists), which is dependent on frontal lobe-mediated processes (Fernades & Grady, 2008).

Lastly, the delay between the immediate and delayed tasks on the CVLT-II follows no set pattern of administration; the CVLT-II instructs the examiner to administer any type of visual task before delayed recall. Common sense suggests that perhaps the order of administration or the types of interference tasks affect delayed recall. Cassel, Johnson, & Burns (1962) noted, however, that the order in which tests in a battery were presented to patients did not have a significant effect on resulting performance. Furthermore, Williams and Donovan (2008) challenged the CVLT-II manual’s instruction that only non-verbal tasks be administered before delayed recall. Results indicated that the variability in tasks given during the delay may negatively impact recall, as compared to a set standard of administration, as with WMS-IV VPA.

Limitations to the current study include the small sample size and mixed clinical and non-clinical sample. Consequently, future research may examine this score substitution within a larger sample size and different populations which have been demonstrated to show variability in cognitive functioning on standardized measures, such as individuals diagnosed with a Mood

---

Disorder (Hermens, Naismith, Redoblado Hodge, Scott, & Hichie, 2010) and non-native English-speakers (Mungas, Reed, Haan, & Gonzalez, 2005). In addition, researchers may wish to examine the validity of the CVLT-II score substitution in different age ranges, particularly in an older adult population. Wechsler (2009) published both an Adult Form (utilized in the present study) and Older Adult Form to accommodate for the variances in memory attributed to age. Research has suggested that older adults demonstrate greater difficulty with binding target items and contextual features into a coherent memory representation (Chalfonte & Johnson, 1996), and this deficit in the creation and retrieval of associative links appears to extend to all memory tasks (i.e., Associative Deficit Hypothesis; Naveh-Benjamin, 2000).

Regarding the CVLT-II score substitution, it was determined that, while both immediate and delayed tasks were moderately associated between the CVLT-II and WMS-IV, the CVLT-II score substitution is not recommended regarding delayed tasks. When looking at the magnitude of the obtained correlations and index score comparisons with and without CVLT-II substitution, it was concluded that innate differences are present between the measures. Differences regarding priming and administration regarding interference tasks are implicated as causes.

## About the Authors

**Jessica Holster, Ph.D.** Jessica Holster, Ph.D., completed her degree in Clinical Psychology at Nova Southeastern University and is affiliated with San Antonio Military Medical Center, San Antonio, TX.

Email: [holster@nova.edu](mailto:holster@nova.edu)

**Christine Corsun-Ascher, Psy.D.** completed her degree in Clinical Psychology at Nova Southeastern University and is affiliated with a private practice in Naperville, IL.

Email: [ccorsun@nova.edu](mailto:ccorsun@nova.edu)

**Traci Olivier, B.S.** is currently pursuing a doctoral degree in Clinical Psychology at Nova Southeastern University, Ft. Lauderdale, FL.

Email: [to141@nova.edu](mailto:to141@nova.edu)

**Charles Golden, Ph.D.** is with Nova Southeastern University, Ft. Lauderdale, FL.

Email: [goldench@nova.edu](mailto:goldench@nova.edu)

## References

- Albuquerque, L., Loureiro, C., Martins, I. P. (2008). Effect of lesion site on serial position during list learning: A study with the CVLT. *International Journal of Neuroscience*, *118*, 917-933. doi: 10.1080/00207450701591081
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4<sup>th</sup> ed., text revision). Washington, D.C.: Author.
- Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428. doi: 10.1037/0033-295X.82.6.407
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effects of feedback in test-like events. *Review of Educational Research*, *61*, 213-238. doi: 10.2307/1170535
- Bosnes, O. & Bjorn, E. (2003). California Verbal Learning Test and Wechsler Memory Scale Revised: Performance in a Norwegian clinical sample. *Magazine for Norwegian Psychology*, *40*, 590-597.
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 918-928. doi: 10.1037/0278-7393.34.4.918
- Byrd, M. (1984). Age differences in the retrieval of information from semantic memory. *Experimental Aging Research*, *10*, 29-33. doi: 10.1006/brcg.1997.0872
-

- Cassell, R.H., Johnson, A.P., & Burns, W.H. (1962). The order of tests in the battery. *Journal of Clinical Psychology, 18*(4), 464-465. doi: 10.1002/1097-4679(196210)18:43.0.CO;2-W
- Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403-416.
- Delis, D. C., Cullum, M., Butters, N., Cairns, P., Prifitera, A. (1988). Wechsler Memory Scale-Revised and California Verbal Learning Test: Convergence and divergence. *The Clinical Neuropsychologist, 2*, 188-196. doi: 10.1080/13854048808520100
- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test-Second Edition*. San Antonio, TX: Psychological Corporation.
- Delis, D. C., Wetter, S. R., Jacobson, M. W., Peavy, G., Hamilton, J., Gongvatana, A., Kramer, J. H., Bondi, M. W., et al. (2004). Recall discriminability: Utility of a new CVLT-II measure in the differential diagnosis of dementia. *Journal of the International Neuropsychological Society, 11*, 708-715. doi: 10.1017/S1355617705050812
- Fernandes, M. A., & Grady, C. (2008). Age differences in susceptibility to memory interference during recall of categorizable but not unrelated word lists. *Experimental Aging Research, 34*, 297-322. doi: 10.1080/03610730802273860
- Golden, C. J., Espe-Pfeifer, P., & Wachsler-Felder, J. (2000). *Neuropsychological interpretations of objective psychological tests*. New York, NY: Kluwer Academic/Plenum Publishers.
- Helmstaedter, C., Wietzke, J., & Lutz, M. T. (2009). Unique and shared validity of the “Wechsler Logical Memory Test”, the “California Verbal Learning Test”, and the “Verbal Learning and Memory Test” in patients with epilepsy. *Epilepsy Research 87*, 203-212. doi: 10.1016/j.eplepsyres.2009.09.002
- Hermens, D. F., Naismith, S. L., Redoblado Hodge, M. L., Scott, E. M., & Hickie, I. B. (2010). Impaired verbal memory in young adults with unipolar and bipolar depression. *Early Intervention in Psychiatry, 4*, 227-233. doi: 10.1111/j.1751-7893.2010.00194.x
- Kang, S. H., McDermott, K. B., and Roediger, L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*, 528-558. doi: 10.1080/09541440601056620
- Lengenfelder, J., Kemenoff, L. A., Kramer, J. H., & Delis, D. C. (2000). Gender differences in verbal learning: California Verbal Learning Test-II. *Archives of Clinical Neuropsychology, 15*. 682. doi: 10.1093/arclin/15.8.682
- Loring, D. W., & Bauer, R. M. (2010). Testing the limits: Cautions and concerns regarding the new Wechsler IQ and memory scales. *Neurology, 74*, 685-690. Doi: 10.1212/WNL.0b013e3181d0cd12
-

- Lovden, M., Ronnlund, M., Wahlin, A., Backman, L., Nyberg, L., & Nilsson, L. (2004). The extent of stability and change in episodic and semantic memory in old age: Demographic Predictors of level and change. *The Journal of Gerontology: Series B: Psychological Sciences & Social Sciences*, *59*, 130-134.
- Mungas, D., Reed, B. R., Haan, M. N., & Gonzalez, H. (2005). Spanish and English neuropsychological assessment scales: Relationship to demographics, language, cognition, and independent function. *Neuropsychology*, *19*, 466-475. doi: 10.1037/0894-4105.19.4.466
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227-234. doi: 10.1037/h0031564
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology, Learning, Memory, & Cognition*, *26*, 1170-1187. doi: 10.1037/0278-7393.26.5.1170
- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 3-8. doi: 10.1037/0278-7393.31.1.3
- Randolph, C., Gold, J. M., Kozora, E., Munro, C., Hermann, B. P., & Wyler, A. R. (1994). Estimating memory function: Disparity of Wechsler Memory Scale-Revised and California Verbal Learning Test indices in clinical and normal samples. *The Clinical Neuropsychologist*, *8*, 99-108. doi: 10.1080/13854049408401547
- Restrepo, M. A., & Gutiérrez-Clellen, V. F. (2004). Grammatical impairments in Spanish-English bilingual children. In B. A. Goldstein, *Bilingual language development and disorders in Spanish-English speakers* (pp. 213-234). Baltimore, MD: Paul H Brookes Publishing.
- Schacter, D. L., & Buckner, R. L. (1999). On the relations among priming, conscious recollection, and intentional retrieval: Evidence from neuroimaging research. *Neurobiology of Learning & Memory*, *70*, 284-303. doi: 10.1006/nlme.1998.3854
- Schear, J. M., & Bruce, C. R. (1989). Examination of the concurrent validity of the California Verbal Learning Test. *Clinical Neuropsychologist*, *3*, 162-168. doi: 10.1080/13854048908403289
- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Macmillan.
- Smith, T. A., & Kimball, D. R. (2010). Dynamics of thematic activation in recognition testing. *Psychonomic Bulletin & Review*, *17*, 355-361. doi: 10.3758/PBR.17.3.355
-

- Trandafir, A., Meary, A., Schurhoff, F., Leboyer, M., & Szoke, A. (2006). Memory tests in first-degree adult relatives of schizophrenic patients. *Schizophrenia Research*, *81*, 217-226. doi: 10.1016/j.schres.2005.09.005
- Wechsler, D. (2009). *Wechsler Memory Scale- Fourth Edition*. San Antonio, TX: Pearson.
- Wible, C. G., Han, S. D., Spencer, M. H., Kubicki, M., Niznikiewicz, M. H., Jolesz, F. A., McCarley, R. W., & Nestor, P. (2006). Connectivity among semantic associates: An fMRI study of semantic priming. *Brain & Language*, *97*, 294-305. doi: 10.1016/j.bandl.2005.11.006
- Williams, B.R. & Donovanick, P.J. (2008). Questioning the rule of thumb: Can verbal tasks be administered during the CVLT-II delay interval? *The Clinical Neuropsychologist*, *22*, 807-812. doi: 10.1080/13854040701562833
-