One

OVERVIEW

INTRODUCTION

During development in childhood, verbal learning and memory are key to success in school and language acquisition and development. A large portion of time is spent in activities requiring retention of verbal information, such as conversations with parents and peers, reading, and interacting with digital devices. The ability to encode and retain verbal information remains elemental throughout the lifespan as much learning and social interaction requires adequate verbal learning and memory ability. Therefore, the assessment of verbal memory is an essential component of clinical and neuropsychological evaluations. This is particularly relevant for evaluations of individuals with known or suspected cognitive or language impairments. Many developmental, neurological, and psychiatric disorders involve disruption or impairment of memory processes (e.g., attention-deficit hyperactivity disorder, dementia, traumatic brain injury). Although some areas of memory decline normally with age (Salthouse, 1996; Tromp, Dufour, Lithfous, Pebayle, & Després, 2015), the prevalence of memory disorders increases with age and memory complaints are a frequent concern cited by older adults. Evaluation of memory processes can assist in differentiating normative development and decline across the lifespan from memory deficits related to insidious origins.

The California Verbal Learning Test, Third Edition (CVLT3; Delis, Kramer, Kaplan, & Ober, 2017) provides a detailed look into verbal learning and memory processes for individuals ages 16–90. It introduced some major scoring changes and updates from the CVLT, second edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000). However, since the CVLT-II is still widely used in clinical and research settings, both editions are described in this book. The CVLT children's version (CVLT-C; Delis, Kramer, Kaplan, & Ober, 1994) provides a similar assessment of memory for individuals ages 5–16. This book provides an easy-to-use reference for individuals learning the essentials of administration, scoring, and interpretation of the CVLT3, CVLT-II, or the CVLT-C. It maintains the direct, systematic approach to presenting material that is characteristic of the Essentials series. In addition, guidelines for selecting the best CVLT product for a specific client and administrative and interpretive guidelines are provided. The latest research on the CVLT products and on verbal memory processes is also included throughout the chapters to assist in applying results obtained with the most recent CVLT editions.

Essentials of California Verbal Learning Test: CVLT-C, CVLT-II, & CVLT3 covers topics that emphasize the appropriate administration, scoring, interpretation, and application of

each test. Each chapter includes several "Rapid Reference," "Caution," or "Don't Forget" boxes to highlight important points for easy reference and clarification. At the end of each chapter, a short "Test Yourself" quiz is provided to help you solidify what you have read. The information in this book is provided to help you understand the nuances of each of the CVLT instruments and become a proficient user.

HISTORY AND DEVELOPMENT

Although the concept of investigating the processes underlying memory is standard in memory assessment instruments today, at the time the original CVLT (Delis, Kramer, Kaplan, & Ober, 1987) was developed in 1987 it was a novel approach to assessment. The authors combined research and clinical experience with patients with memory deficits to create a novel approach to assessing learning and memory processes. Memory deficits were linked to specific tasks within a 16-word memory recall and recognition test. The 16 words were derived from four semantic categories, allowing the examinee to group words semantically as an aid to recalling the words. This approach allows examination of not only how much an individual can learn and retain but also the strategies used to learn and recall information and the types of memory errors made. These additional measures provide detailed information on the processes involved in learning and memory to aid in diagnosis and intervention planning.

In the original CVLT, the examiner read a Monday shopping list that contained 16 items to the examinee. The 16 items consisted of four words from four different semantic categories (tools, fruit, clothing, spices, and herbs). The Monday shopping list was repeated in five learning trials, followed by an interference list trial. A Tuesday shopping interference list also comprised 16 words from four categories. Two of the categories overlapped with categories from the Monday list and two were novel categories. Following recall of the Tuesday list, the examinee was asked to recall the Monday list in both a free-recall and a cued-recall trial. During the cued-recall trial, the examiner provided each of the four semantic categories and asked the examinee to name the items from each. The examiner than waited 20 min before administering delayed free- and cued-recall trials, as well as a delayed recognition trial (see Figure 1.1). The examiner was encouraged to administer nonverbal tasks during this 20-min delay. Scores derived on the CVLT assessed attention, learning strategies, recall accuracy, interference effects, recall errors, and recognition. The original normative sample consisted of 273 neurologically intact individuals (104 males, 169 females) recruited from other research studies being conducted at the time of data collection.

Since 1987, three revisions and expansions of the CVLT have been published, although the administration has remained remarkably stable. The expansion and evolution of the CVLT reflect the growing influence and use of the process approach to memory assessment. Its wide use in research on memory also ensures several editions of the CVLT will remain in use for a long time. The currently available editions include the CVLT3 (published in 2017), the CVLT-II (published in 2000), and the CVLT-C (published in 1994). Each of the editions reflects the knowledge and theories of memory at the time of its development as well as the developmental needs of the population assessed. This chapter provides an overview of the content of each of the currently available CVLT instruments,

Learning trials	Interference trial	Short-delay	Long-delay (follows 20-min delay)	Yes/No recognition	Forced-choice recognition (follows 10-min delay)[a]
Trial 1 Free Recall	List B Free Recall	Short-Delay Free Recall	Long-Delay Free Recall	Yes/No Recognition	Forced-Choice Recognition
Trial 2 Free Recall		Short-Delay Cued Recall	Long-Delay Cued Recall		
Trial 3 Free Recall					
Trial 4 Free Recall					
Trial 5 Free Recall					

[a]Only on CVLT-II and CVLT3.

Figure 1.1. CVLT-C, CVLT-II standard and alternate, and CVLT3 standard and alternate forms structure.

and later chapters provide a review of the research literature related to memory and learning as measured by the CVLT.

OVERVIEW AND ORGANIZATION OF THE CVLT-C

The CVLT was expanded to younger ages with the publication of the CVLT-C. It not only extended the age downward to 5, it also modified the content to be more relevant for children. The word list was shortened to 15 words derived from three semantic categories. The semantic categories reflect categories common in childhood. Two of these categories overlap with the semantic categories contained in the adult forms. Similar to the administration of the original CVLT, the child is read the list of 15 words across five learning trials, followed by an interference list trial. Just as in the original CVLT, the lists are presented as Monday and Tuesday shopping lists. Following the interference trial, the child recalls the original list in both short-delay free- and cued-recall trials. The examiner then completes nonverbal testing for 20 min before administering long-delay free- and cued-recall trials, as well as a delayed recognition trial (see Figure 1.1). All trials are required to derive the primary scores. Scores derived on the CVLT-C assess auditory attention, learning strategies and characteristics, recall accuracy and consistency, proactive and retroactive interference, recall errors, and recognition.

A *T* score is derived for the learning trials (Trials 1–5 Total) with a mean of 50 and a standard deviation of 10. The process scores provided in the CVLT-C provide detailed information about the learning and memory processes required to recall verbal information. For these scores, age-corrected *z* scores are provided with a mean of 0 and a standard

deviation of 1. The scores derived in CVLT-C are listed in Rapid Reference 1.1 by condition. Detailed information on calculating scores is provided in Chapter 3.

Rapid Reference 1.1

Primary Scores Derived in CVLT-C, by Condition

Learning trials	Recall scores	Recognition scores
Trial 1 Free-Recall Correct	Short-Delay Free Recall	Recognition Hits
Trial 5 Free-Recall Correct	Short-Delay Cued Recall	Discriminability
Total Trials 1–5 (*T* score)	Short-Delay Free Recall vs. List A Trial 5 (difference score)	Discriminability vs. Long-Delay Free Recall
List B Free-Recall Trial	Long-Delay Free Recall	False Positives
List B Recall vs. List A Trial 1 Recall (difference score)	Long-Delay Cued Recall	Response Bias
Semantic Cluster Ratio	Long-Delay Free Recall vs. Short-Delay Free Recall (difference score)	
Serial Cluster Ratio	Total Perseverations	
Expected Serial Clustering	Total Free-Recall Intrusions	
Percent of Total Recall from Primacy Region	Total Cued-Recall Intrusions	
Percent of Total Recall from Middle Region	Total Intrusions	
Percent of Total Recall from Recency Region		
Learning Slope		
Recall Consistency		

Difference scores compare performance on one task to performance on another task. On CVLT-C, difference scores are derived using two methods (see Chapter 3 for detailed information on calculating the change scores): raw percentage change and scaled score difference. The raw percentage change scores are not normed due to the heavy influence of the raw scores on the calculation of percentage retained, such that low overall recall can result in higher retention percentages. Difference (or savings) scores utilize the age-corrected scaled scores. Means and standard deviations for the normative sample are provided to provide context for these scores. Change and difference scores should not replace the primary scores but are used to guide the interpretation of differences observed across conditions. Detailed information on the interpretation of scores is provided in Chapter 4.

OVERVIEW AND ORGANIZATION OF THE CVLT-II

The CVLT-II was the first major revision to the CVLT. The revision was guided by neuropsychological and cognitive research on memory, feedback from users and reviewers of the CVLT (both personal and published), and clinical experience by the authors. One of the major criticisms of the CVLT was the small, nonrepresentative sample used for the norms. Multiple studies demonstrated that the norms resulted in lower than expected scores in individuals with average to low educational attainment (Paolo, Troster, & Ryan, 1997; Wiens, Tindall, & Crossen, 1994). A large representative sample of the population aged 16–89 was collected to update the norms across seven age bands. In terms of content, the 16 items used in the word lists were modified to be easier to understand and were not presented as a shopping list. The word lists were derived using high-frequency words commonly spoken within the United States from four different semantic categories. The interference list also comprised 16 words from four categories, two of which overlapped with the original list. In addition, a forced-choice recognition trial was added to assess insufficient effort. Two new forms were also introduced with the CVLT-II: an Alternate Form and a Short Form. The Alternate Form is equated to the Standard Form, utilizes the same administration format, and provides the same scores as the Standard Form but uses alternate word lists from the Standard Form. The Short Form uses a nine-word list with a shortened format to accommodate use as a screener or with individuals that cannot tolerate lengthy testing. The Standard and Alternate Forms take 30 min to administer (in addition to the 20-min delay and an optional 10-min delay to administer the forced-choice recognition condition), whereas the Short Form takes around 20 min (in addition to the 10-min delay).

The development of the CVLT-II incorporated analysis of the performance of each score to evaluate the influence of demographic variables and general cognitive ability on performance, reliability and stability, score range and distribution for floor and ceiling problems, and clinical utility. Norms are corrected for age and sex because these demographic factors contributed more than 5% of the variance in primary scores. Detailed information on the psychometric properties and the clinical utility of the CVLT-II are described in the *CVLT-II Manual* (Delis, Kramer, Kaplan, & Ober, 2000).

Just as on the original CVLT, for the Standard and Alternate Forms, the examinee is read a list of 16 words across five learning trials, followed by an interference list trial. Following the interference trial, the examinee recalls the original list in both free-recall and cued-recall trials. The examiner then completes nonverbal testing for 20 min before administering delayed free-recall and cued-recall trials, as well as a delayed yes/no recognition trial. The examiner may then administer an optional forced-choice recognition trial following a 10-min delay (see Figure 1.2). All trials except the forced-choice recognition trial are required to derive the primary scores. It is highly recommended that examiners routinely administer the forced-choice recognition condition to assess performance validity.

For the Short Form, the examinee is read a list of nine words (from three semantic categories) across four learning trials, followed by a 30-s distractor task. Following the distraction task, the examinee is asked to recall the list in a short-delay free-recall trial. The examiner than completes nonverbal testing for 10 min before administering long-delay free- and cued-recall trials, as well as a delayed yes/no recognition trial. The examiner may then administer an optional forced-choice recognition trial following a 5-min delay (see Figure 1.3). All trials except the forced-choice recognition trial are required to derive

Learning trials	30-second distractor task	Short-delay	Long-delay (follows 10-min delay)	Yes/No recognition	Forced-choice recognition (follows a 5-min delay)
Trial 1 Free Recall		Short-Delay Free Recall	Long-Delay Free Recall	Yes/No Recognition	Forced-Choice Recognition
Trial 2 Free Recall			Long-Delay Cued Recall		
Trial 3 Free Recall					
Trial 4 Free Recall					

Figure 1.2. CVLT-II and CVLT3 structure, brief form.

the primary scores. See Rapid Reference 1.2 for an overview of differences between the Standard/Alternate Forms and the Short Form.

Rapid Reference 1.2

Differences Between the CVLT-II Standard/Alternate Forms and the Short Form

Standard/alternate form	Short form
Administration	
16-word list from 4 categories	9-word list from 3 categories
5 Learning Trials	4 Learning Trials
Interference List and Recall	30-s distraction task
Short-Delay Free and Cued Recall	Short-Delay Free Recall
20-min delay	10-min delay
Long-Delay Free and Cued Recall	Long-Delay Free and Cued Recall
Long-Delay Yes/No Recognition	Long-Delay Yes/No Recognition
10-min Delay	5-min Delay
Forced-Choice Recognition	Forced-Choice Recognition
Scoring	
18 Hand-Scored Variables	15 Hand-Scored Variables
66 Normed Variables (Software)	51 Normed Variables (Software)
Multiple Raw Scores (Software)	

Scores derived on the CVLT-II assess attention, learning strategies, recall accuracy and consistency, proactive and retroactive interference, recall errors, recognition, and performance validity. Normative data are provided for 27 primary scores and 39 expanded scores in the scoring software. Eighteen of the primary variables can be hand scored. For the Short Form, 23 primary scores are provided along with 28 expanded scores in the scoring software. Fifteen of the primary variables can be hand scored. In addition, the scoring software provides raw data for numerous nonnormed variables in a research report.

A *T* score is derived for the learning trials (Trials 1–5 for Standard/Alternate Forms and Trials 1–4 for the Short Form) with a mean of 50 and a standard deviation of 10. The process scores provided in the CVLT-II provide detailed information about the learning and memory processes required to encode, recall, and recognize verbal information. For these scores, age- and gender-corrected *z* scores are provided with a mean of 0 and a standard deviation of 1. In addition, cumulative percentages are provided for some scores with highly skewed distributions. The normative scores derived in CVLT-II are listed in Rapid Reference 1.3 by condition.

Rapid Reference 1.3

Primary Scores Derived in CVLT-II, by Condition

Learning trials	Recall scores	Recognition scores
Trial 1 Free-Recall Correct[a]	Short-Delay Free-Recall Correct[a]	Long-Delay Yes/No Recognition Total Hits[a]
Trial 2 Free-Recall Correct[a]	Short-Delay Cued Recall Correct[a,b]	Long-Delay Yes/No Recognition Total False Positives[a]
Trial 3 Free-Recall Correct[a]	Short-Delay Free Recall vs Trial 5 (contrast score)	Total Recognition Discriminability (*d'*)
Trial 4 Free-Recall Correct[a]	Long-Delay Free-Recall Correct[a]	Total Recognition Discriminability vs. Long-Delay Free Recall (contrast score)
Trial 5 Free-Recall Correct[a,b]	Long-Delay Cued Recall Correct[a]	Total Response Bias
Trials 1–5 Free-Recall Correct (T score)[a]	Long-Delay Free Recall vs. Short-Delay Free Recall (contrast score)	Long-Delay Forced-Choice Recognition Percent Total Accuracy[a]
List B Free-Recall Correct[a,b]	Free-Recall Intrusions[a]	
List B vs. Trial 1 (contrast score)	Cued-Recall Intrusions[a]	
Semantic Clustering (Chance-Adjusted)		
Semantic Clustering Bidirectional (Chance-Adjusted)		

(*continued*)

Learning trials	Recall scores	Recognition scores
Subjective Clustering (Chance-Adjusted)	Total Intrusions[a]	
Percent Recall from Primacy	Total Repetitions[a]	
Percent Recall from Middle	Total Recall Discriminability	
Percent Recall from Recency		
Total Learning Slope Trials 1–5		
Across-Trial Recall Consistency		

[a]Scores that are easily hand scored.
[b]Scores not on the Short Form.

Difference scores compare performance on one task to performance on another task. On CVLT-II, difference (or savings) scores are derived using the sex- and age-corrected *z*-scores. Change and difference scores should not replace the primary scores but are used to guide interpretation of differences observed across conditions. Detailed information on the interpretation of scores is provided in Chapter 4.

OVERVIEW AND ORGANIZATION OF THE CVLT3

The CVLT3 is the most recent revision to the CVLT. The revision was guided by the need for updated normative data, feedback from users and reviewers of the CVLT-II (both personal and published), and research utilizing the CVLT-II. In terms of content, the administration instructions and word lists were not modified but remain the same as in the CVLT-II. However, a few items on the forced-choice recognition trial were modified to increase sensitivity as a measure of performance validity. The alternate and short forms were also retained with similar modifications made to the forced-choice recognition trial.

The CVLT3 addressed many needs identified in the research literature and customer feedback. Revisions made in the CVLT3 included:

- updated norms for ages 16–90, using a nationally stratified sample matched to the U.S. population;
- application of a scaled score metric (mean = 10, SD = 3) over the *T* score and *z*-score metric to allow easier comparison to other measures;
- introduction of index scores (mean = 100, SD = 15) on the key memory and learning scores (Trials 1–5, total delayed recall, total recall);
- updated intrusion measures to reflect different types of memory errors;

- digital administration, recording, and scoring; and
- provision of demographic adjustments to age-adjusted scores for education and sex presented as *T* scores in the scoring software.

Detailed information on the modifications and improvements to the content, psychometric properties, and clinical utility of the CVLT3 are described in the *CVLT3 Manual*. An overview of the changes is provided in Rapid Reference 1.4.

Rapid Reference 1.4

Changes from CVLT-II to CVLT3

- New normative sample reflective of 2015 U.S. Census data
- Increased age range to 90
- Index scores on standard score metric (mean = 100, SD = 15)
- Primary scores on scaled score metric (mean = 10, SD = 3)
- Trials 1–5 Total score computed by summing scaled scores for Trials 1–5
- Norms provided are age corrected only
- Demographic adjustments to normative scores available for sex and education, presented as *T* scores
- Forced-choice recognition items modified to include only concrete distractors (abstract distractors were removed)
- New measures of across- and within-trial intrusions
- New measures of intrusion error types
- New yes/no recognition scores that describe the types of recognition errors

The CVLT3 introduced significant changes to the scoring of the CVLT-II. Although scores still measure the processes underlying attention, learning, and memory, the traditional *T* score and *z*-score metrics were changed to standard and scaled scores. This allows direct comparison to other measures commonly used in evaluations. Three index scores are derived by summing the scaled scores for the learning trials (Trials 1–5), for the delayed recall trials (free and cued recall), and for all recall trials (learning, interference, short delay, and long delay). For the primary process scores, age-corrected scaled scores were derived with a mean of 10 and a standard deviation of 3. In addition, cumulative percentages are provided for some scores that had highly skewed distributions. The normative scores derived in CVLT3 are listed in Rapid Reference 1.5 by condition.

Contrast scores are utilized for difference scores in CVLT3. Contrast scaled scores provide information about performance on one task adjusted for performance on another relevant task. Similar to the manner in which demographic adjustments are derived for normative scores, one score is adjusted to account for performance on a related but separate score. For example, the Long-Delay Free-Recall Correct vs. Short-Delay Free-Recall

Correct Contrast Scaled Score adjusts the long-delay score based on performance on short-delay recall. This accounts for differences in performance on delayed memory due to differences in immediate recall. The new score represents the examinee's performance on delayed memory in comparison to individuals of similar immediate memory ability. Because contrast scores utilize age-adjusted scaled scores, they are not further adjusted by age. Contrast scaled scores are provided at the scaled score level and are presented as scaled scores with a mean of 10 and a standard deviation of 3. Contrast scores are used to interpret scores in relation to similar ability peers; they do not replace subtest scaled scores and should not be substituted for primary scores in reports or to compute index scores. Detailed information on the interpretation of contrast scaled scores is provided in Chapter 4.

The CVLT3 introduces the use of demographic adjustments to norms. In addition to the age-adjusted normative scores, education, and sex adjustments are provided in the scoring software. The demographic adjustments are applied to the normative scores to produce *T* scores that account for education and sex differences.

Rapid Reference 1.5

Primary Scores Derived in CVLT3, by Condition

Learning trials	Recall scores	Recognition scores
Trial 1 Correct[a]	Short-Delay Free-Recall Correct[a]	Long-Delay Yes/No Recognition Total Hits[a]
Trial 2 Correct[a]	Short-Delay Cued-Recall Correct[a,b]	Long-Delay Yes/No Recognition Total False Positives[a]
Trial 3 Correct[a]	Short-Delay Free-Recall Correct vs. Trial 5 Correct (contrast score)	Recognition Discriminability (*d'*)
Trial 4 Correct[a]	Long-Delay Free-Recall Correct[a]	Recognition Discriminability Nonparametric
Trial 5 Correct[a,b]	Long-Delay Cued-Recall Correct[a]	Long-Delay Free-Recall vs. Recognition Discriminability (*d'*) (contrast score)
Trials 1–5 Correct (standard score)[a]	Delayed Recall Correct (standard score)	Long-Delay Forced-Choice Recognition Hits[a]
List B Correct[a,b]	Total Recall Correct (standard score)	Long-Delay Free-Recall Discriminability vs. Recognition Discriminability (*d'*) (contrast score)
List B Correct vs. Trial 1 Correct (contrast score)	Total Recall Responses (standard score)	

Learning trials	Recall scores	Recognition scores
	Long-Delay Free-Recall Correct vs. Trial 5 Correct (contrast score)	
	Long-Delay Free-Recall Correct vs. Short-Delay Free-Recall Correct (contrast score)	
	Total Intrusions[a]	

[a]Scores that are easily hand scored.
[b]Scores not on the Short Form.

THEORETICAL FOUNDATION

Learning and memory have long been of interest to psychologists, with William James first proposing the concepts of short-term and long-term memory in 1890. Short-term memory was described as finite with low durability unless it was encoded into long-term memory that was more lasting and of infinite capacity. Ebbinghaus' (1885) famous experiments on memory described the processes of learning and forgetting and introduced the concepts of the learning curve (rate at which information is acquired over repeated trials), serial position effects (how position within a series of words impacts recall), and the forgetting curve (rate at which one forgets information with most decay occurring within the first 20 min after learning). These early descriptions and investigations into memory laid the groundwork for modern memory assessment.

Despite Ebbinghaus' experiments on the processes of learning and memory, the assessment of these concepts has historically focused on the amount of information an individual could encode, consolidate, and recall. *Encoding* is the process of taking external information and transforming it into mental representations or memories. *Consolidation* is the process through which information in immediate memory is moved into long-term memory, and *retrieval* is the process of recalling information from storage. Focus on the amount of recalled information allows a global picture of memory ability, including the determination of the presence of memory disorders.

With the use of immediate and delayed recall, memory assessments often provide scores for short-term memory and long-term memory, again measuring the amount of information retained. *Short-term memory* is the momentary storage of information, lasting from a few seconds to a few minutes. Memories lasting from hours to years are considered stored in *long-term memory*. Long-term memory can be categorized as either *implicit* or *explicit* memory. *Implicit* or procedural memory involves involuntary learning from experiences without conscious awareness, such as learning to ride a bike or drive a car. *Explicit* or declarative memory involves the purposeful storage and retrieval of information. Explicit memory can further be divided into *semantic* (factual) and *episodic* (personal events and

context) memory. When described in these terms, the CVLT versions are measures of explicit episodic verbal memory.

Processes of Memory

Edith Kaplan and colleagues introduced the process approach to cognitive assessment (Kaplan, 1988; Libon, Swenson, Ashendorf, Bauer, & Bowers, 2013) through their work at the Boston Veterans Administration Medical Center. This approach places importance on assessing not just *what* an individual is able to do but also *how* they do it. The approach utilizes overall test scores to assess the severity of impairment but emphasizes the analysis of the process through which test scores are achieved and errors made. Similar overall test scores can be obtained through very different processes. The development of the CVLT applied this process approach to learning and memory. Through studies involving individuals with brain injuries or disorders, strengths and weaknesses of memory processes were identified that differed across and within clinical populations. The instrument was refined to capture these processes and allow further investigation. So, although importance is placed on the primary measures of recall accuracy; critical attention is also given to the processes underlying performance.

Performance on memory specific measures is not only influenced by learning and memory processes but also by other cognitive functions. For example, attentional processes and learning strategies have an impact on the encoding and retrieval of verbal information. The interrelated nature of processes utilized across cognitive abilities requires that the assessment of memory include the assessment of these related processes. The CVLT versions measure the specific processes of learning and memory and processes related to the success and failure of encoding and retrieval of information. This allows examiners to examine differences in performance due to specific cognitive deficits. Figure 1.3 displays the processes measured directly within the CVLT editions.

The ability to encode information is highly dependent upon the ability to perceive and attend to information. Auditory attention is a key precursor for adequate memory retention. Individuals who have difficulty focusing on information long enough to encode it will not be able to retain that information over time. The first trial of the CVLT provides information on auditory attention span. It correlates highly with other measures of attention and is impaired in clinical groups with known attentional deficits, such as those with anxiety or mood disorders. Most individuals improve in overall recall across the learning trials, so comparison of performance across trials can assist in teasing out the influence of attention on memory problems.

Learning is the acquisition of new information. Historically it has been measured through overall recall across learning trials or by assessing the learning slope, the amount of information gained after the initial learning trial. These measures are included in the CVLT editions but the CVLT editions go further to evaluate learning strategies and characteristics and consistency of recall. Learning involves both passive and active strategies. Rote repetition is a passive strategy for recalling information that increases consolidation but does not involve increasing efficiency of encoding information. Repetition requires that information is repeated multiple times for encoding to occur. This type of learning

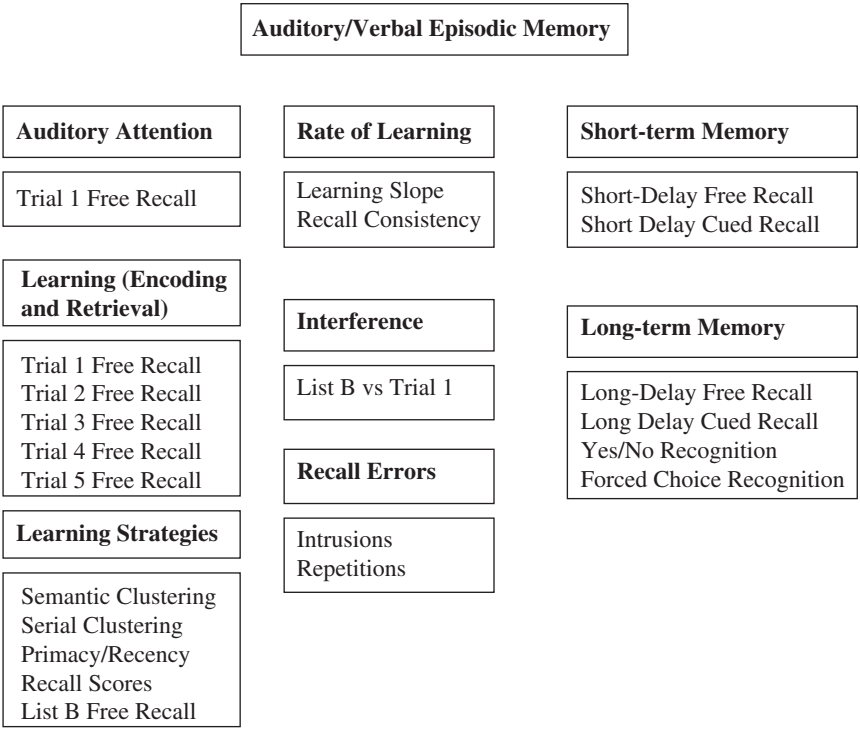

Figure 1.3. Memory processes measured in the CVLT.

strategy is represented in serial clustering or recalling words together that were close serially in the original list. Alternatively, active strategies of organizing information into meaningful groups aid the encoding and retrieval of information. The semantic grouping of words by an examinee represents utilization of this active learning strategy. Information on strategies used to encode information can aid in developing interventions for improving active learning.

In addition to evaluating learning strategies, data on the recall of words from different sections of the word list are provided. It is typical for individuals to recall more information from the beginning of a list (primacy effect) and from the end of a list (recency effect) than from the middle of a list. The primacy effect is generally attributed to greater rehearsal time for information whereas the recency effect is potentially related to the last words being held in recent or working memory. Examinees with encoding deficits often show higher recency effects than observed in the normative sample. Finally, consistency of recall provides information on the consistent application of recall strategies to retrieval of information. Individuals with executive functioning difficulties (e.g., poor planning, poor organization) often produce inconsistent profiles of responding. While this affects overall recall, executive functioning may be a key deficit in these cases.

The introduction of a second word list introduces active interference into the assessment of memory. Other measures utilize unrelated tasks to fill in between immediate and delayed recall but the CVLT actively introduces a second, similar list to explicitly assess interference. Two of the four semantic categories overlap across the two lists further eliciting interference effects. Comparison of performance on Trial 1 and the interference trial provides a measure of proactive interference, the decline in performance on learning material due to prior learning.

Whereas the level of recall on the immediate and delayed recall trials reflects overall retrieval, errors in recall provide insight into specific memory dysfunction and are invaluable in differential diagnoses. Repetitions or repeated responses are common in individuals with normal memory functioning when the repetition is used as self-cueing to promote further recall. However, repetitions can also be a sign of perseverative responding, the failure to inhibit previous responses, or poor self-monitoring. Intrusions, responses not from the current word list, are relatively uncommon in individuals without memory difficulties. Examination of the specific types of intrusions, which trials they occur on, their semantic relationships to the items in the word lists, and whether they are repeated across or within trials is critical to differentiating clinical groups. All clinical groups demonstrate deficits in memory recall; however, intrusions provide specific information about performance that aids in differentiating clinical groups.

Performance Validity

The assessment of performance validity has become standard practice in forensic and neuropsychological evaluations in both adult and pediatric populations (Brooks, Ploetz, & Kirkwood, 2016; Heilbronner et al., 2009; Holcomb, 2018; Martin, Schroeder, & Odland, 2015). The forced-choice task on the CVLT-II and CVLT3 was developed as an embedded measure of performance validity. In addition, Lichtenstein, Holcomb, and Erdodi (2018) presented data on a forced-choice measure developed for use with the CVLT-C. Several other scores across all three instruments have also demonstrated worth as indicators of performance validity, including recognition discriminability, Trials 1–5 Correct, Long-Delay Cued Recall, and Yes/No Recognition Hits (Bauer, Yantz, Ryan, Warden, & McCaffrey, 2005; Brooks & Ploetz, 2015; Shura, Miskey, Rowland, Yoash-Gantz, & Denning, 2016; Whiteside et al., 2015). It is important to note that low scores on these measures alone do not indicate invalid performance but suggest the possibility of symptom exaggeration or other factors that could influence performance.

RESEARCH FOUNDATION

Standardization and Psychometric Properties

Prior to evaluating the reliability data on the CVLT editions, it is important to note that estimates of reliability pose particular difficulties in measures of learning and recall. Measures of internal reliability do not accurately describe the reliability of memory measures due to item score interdependence. Recalling one word on a trial influences the recall of

other words on a trial and also increases the likelihood of recalling the same word on further trials. For this reason, measures of test-retest reliability or alternate form reliability provide greater insight into the reliability of memory measures, although they are influenced by practice effects. Error measures or scores with limited variability also produce lower reliabilities due to skewed distributions. These limitations of traditional measures of reliability should be considered when interpreting the reliabilities described for the CVLT-C, CVLT-II, and CVLT3.

CVLT-C: The standardization sample for the CVLT-C consisted of 920 children selected to form a representative sample of the U.S. population, based on March 1988 U.S. Census data. It was stratified based on age, sex, race/ethnicity, education level, and geographic region. Twelve normative age bands were created, each included 1 year of age. Each age band for ages 5–12 included 80 children and bands for ages 13–16 included 70 children. Sex was roughly equal within each age group; all other demographic variables roughly matched the U.S. Census data.

Due to the interdependent nature of responses on word list recall, the CVLT-C utilized several measures of internal consistency. Internal consistency was evaluated using three approaches: comparing overall performance on odd and even numbered learning trials, across-semantic category consistency, and across-word consistency. The odd-even and across-word approaches yielded average correlations of 0.88 and 0.83, respectively, for Trials 1–5. The across-semantic category approach yielded an average correlation of 0.72. Detailed information on how these consistency estimates were defined and derived is provided in the *CVLT-C Manual*.

The test-retest sample consisted of 106 children tested between 10 and 42 days apart. Results are reported for three age groups: 8-, 12-, and 16-year-olds. Memory and learning measures are particularly susceptible to practice effects that lower test-retest correlations (Strauss, Sherman, & Spreen, 2006) due to repeated exposure of the stimuli to be recalled. Stability coefficients for 13 CVLT-C scores are listed in Rapid Reference 1.6. Test-retest coefficients ranged from 0.61 to 0.73 for the Trials 1–5 *T* score, from 0.26 to 0.77 for the recall *z*-scores, and from 0.17 to 0.90 for the error *z*-scores.

CVLT-II: The standardization sample for the CVLT-II consisted of 1,087 individuals selected to form a representative sample of the U.S. population, based on the March 1999 U.S. Census data. It was stratified based on age, sex, race/ethnicity, education level, and geographic region. Seven normative age bands were created: 16–19, 20–29, 30–44, 45–59, 60–69, 70–79, and 80–89. Each age band included between 107 and 200 individuals. Sex was evenly represented for ages 16–59; in ages 60–89 more females were included than males, reflecting the sex distribution in the population at the older ages.

Internal consistency was evaluated using the three approaches introduced in the CVLT-C: comparing overall performance on odd and even numbered learning trials, across-semantic category consistency, and across-trial word consistency. The odd-even and across-semantic category approaches yielded average correlations of 0.94 and 0.83, respectively, for Trials 1–5. The across-trial word consistency approach yielded an average correlation of 0.79. Estimates obtained in a clinical sample of 124 neuropsychiatric patients produced similar reliability coefficients. Detailed information on how these consistency estimates were defined and derived is provided in the *CVLT-II Manual*.

Rapid Reference 1.6

Stability Coefficients for 13 CVLT-C Scores, by Age

Score	Age 8 average test-retest r_{12}	Age 12 average test-retest r_{12}	Age 16 average test-retest r_{12}
List A Trials 1–5 Total	0.73	0.73	0.61
List B Free Recall Total	0.59	0.26	0.66
Short-Delay Free Recall	0.40	0.77	0.48
Short-Delay Cued Recall	0.75	0.49	0.59
Long-Delay Free Recall	0.59	0.62	0.60
Long-Delay Cued Recall	0.69	0.69	0.59
Semantic Cluster Ratio	0.56	0.58	0.53
Perseverations	0.90	0.32	0.31
Free-Recall Intrusions	0.74	0.56	0.85
Cued-Recall Intrusions	0.59	0.17	0.74
Recognition Hits	0.38	0.24	0.80
Discriminability	0.55	0.37	0.78
False Positives	0.62	0.35	0.78

Standardization data from the *California Verbal Learning Test, Children's Version (CVLT-C)*.

The test-retest sample consisted of 78 individuals tested between 9 and 49 days apart. Stability coefficients for 18 CVLT-II scores are listed in Rapid Reference 1.7. Test-retest coefficients were 0.82 for the Trials 1–5 *T* score, ranged from 0.57 to 0.88 for the recall *z*-scores, and from 0.27 to 0.86 for the error *z*-scores.

Due to the addition of the Alternate Form, alternate form reliability was also provided. The Alternate Form sample consisted of 288 nonclinical adults tested between 0 and 77 days apart. The administrations were counterbalanced, with 155 receiving the Standard Form first followed by the Alternate Form and 133 receiving the Alternate Form first followed by the Standard Form. Correlation coefficients for the 18 scores that can be hand scored are listed in Rapid Reference 1.8. Correlation coefficients were 0.79 for the Trials 1–5 *T* score, ranged from 0.51 to 0.76 for the recall *z* scores, and ranged from 0.27 to 0.76 for the error *z* scores.

Rapid Reference 1.7

Stability Coefficients for CVLT-II Scores, by Age

Score	Average test-retest r_{12}
Trials 1–5 Correct (*T* score)	0.82
Trial 1 Correct	0.57
Trial 2 Correct	0.60
Trial 3 Correct	0.58
Trial 4 Correct	0.82
Trial 5 Correct	0.76
Trial B Correct	0.61
Short-Delay Free-Recall Correct	0.81
Long-Delay Free-Recall Correct	0.88
Semantic Clustering (Chance-Adjusted) Trials 1–5	0.74
Total Learning Slope, Trials 1–5	0.27
Total Intrusions	0.63
Total Repetitions	0.30
Long-Delay Yes/No Recognition Hits	0.79
Long-Delay Yes/No Recognition False Positives	0.72
Total Recognition Discrimination	0.86
Total Response Bias	0.57

Standardization data from the *California Verbal Learning Test, Second Edition Adult Version (CVLT-II)*.

CVLT3: The standardization sample for the CVLT3 consisted of 700 individuals selected to form a representative sample of the U.S. population, based on 2015 U.S. Census data. It was stratified based on age, sex, race/ethnicity, education level, and geographic region. Seven normative age bands were created: 16–19, 20–29, 30–44, 45–59, 60–69, 70–79, and 80–90. Each age band included 100 individuals. Sex was evenly represented for ages 16–59 and in proportion to the U.S. Census in ages 60–90.

Internal consistency was evaluated using alternate form reliability. The Alternate Form sample consisted of 213 nonclinical adults administered the Standard and Alternate Forms

Rapid Reference 1.8

Alternate Form Reliability for the CVLT-II

Score	Average test-retest r_{12}
Trials 1–5 Correct (*T* score)	0.79
Trial 1 Correct	0.52
Trial 2 Correct	0.61
Trial 3 Correct	0.71
Trial 4 Correct	0.70
Trial 5 Correct	0.71
Trial B Correct	0.51
Short-Delay Free-Recall Correct	0.73
Long-Delay Free-Recall Correct	0.76
Semantic Clustering (Chance-Adjusted) Trials 1–5	0.61
Semantic Clustering (Chance-Adjusted) Trials 1–5	0.35
Total Learning Slope, Trials 1–5	0.38
Total Intrusions	0.55
Total Repetitions	0.46
Long-Delay Yes/No Recognition Hits	0.64
Long-Delay Yes/No Recognition False Positives	0.76
Total Recognition Discrimination	0.72
Total Response Bias	0.55
Forced-Choice Recognition Percent Total Accuracy	0.68

of the CVLT3 between 10 and 51 days apart. The administrations were counterbalanced, with 107 receiving the Standard Form first followed by the Alternate Form and 106 receiving the Alternate Form first followed by the Standard Form. Correlation coefficients are reported in the *CVLT3 Manual* for all normed scores for two age groups: 16–44 and 45–90. Alternate form correlations for the 18 core and index scores are listed in Rapid Reference 1.9. Correlation coefficients ranged from 0.79 to 0.83 for the index scores, from 0.50 to 0.71 for the recall scaled scores, from 0.49 to 0.67 for the recognition scaled scores, and from 0.15 to 0.66 for the error and process scaled scores.

Rapid Reference 1.9

Alternate Form Reliability for the CVLT3 Core and Index Scores

Score	Ages 16–44 average test-retest r_{12}	Ages 45–90 average test-retest r_{12}
Trials 1–5 Correct (Standard score)	0.75	0.80
Trial 1 Correct	0.51	0.56
Trial 2 Correct	0.53	0.59
Trial 3 Correct	0.70	0.63
Trial 4 Correct	0.58	0.67
Trial 5 Correct	0.50	0.71
List B Correct	0.57	0.53
Short-Delay Free-Recall Correct	0.71	0.71
Short-Delay Cued-Recall Correct	0.56	0.69
Long-Delay Free-Recall Correct	0.61	0.71
Long-Delay Cued-Recall Correct	0.60	0.65
Delayed Recall Correct (Standard Score)	0.77	0.80
Total Recall Correct (Standard Score)	0.79	0.83
Total Intrusions	0.25	0.57
Yes/No Recognition Hits	0.49	0.62
Yes/No Recognition False Positives	0.57	0.55
Recognition Discrimination	0.67	0.54
Recognition Discrimination Nonparametric	0.65	0.58

Standardization data from the *California Verbal Learning Test, Third Edition (CVLT3)*.

Effect sizes ranged from −0.09 to 0.06 across the index scores, from −0.11 to 0.11 across the recall scaled scores, from −0.15 to 0.06 across the recognition scaled scores, and from −0.18 to 0.29 across the error and process scaled scores, indicating small changes in performance across forms. Therefore, it is unlikely that

CAUTION 1.1

Memory and learning measures are particularly susceptible to practice effects. Practice effects can be observed for more than 6 months.

performance will vary across the two forms greatly due to experience with the test. It is important to note that when the same form is administered, memory tests show practice effects (i.e., performance improvement related to prior experience with taking a test) for a long time after administration, with patients recalling information for months to years after administration (Goldberg, Harvey, Wesnes, Snyder, & Schneider, 2015).

COMPREHENSIVE REFERENCES ON TEST

The most detailed and comprehensive information for each edition of the CVLT can be found in the corresponding test manual: *CVLT-C Manual, CVLT-II Manual,* and *CVLT3 Manual*. Each manual provides an overview of the test, descriptions of each condition and score, and detailed information on administration and scoring, calculating the raw scores, and deriving normative scores. Information on the theoretical underpinnings, development and standardization, reliability, validity, and interpretation are also provided. In addition, the *CVLT3 manual* provides an overview of the use of demographic adjustments to the CVLT3 norms.

Editions of the CVLT have been used in multiple research studies incorporated into thousands of published articles. The CVLT-C has been used to assess memory functioning in typically developing children, children with acquired disorders, such as brain tumors and injuries, and children with neurological disorders, such as epilepsy. The CVLT-II has been used to examine memory functioning in typically developing and aging adults, individuals with degenerative diseases, such as mild cognitive impairment and dementia, individuals with acquired injuries, such as stroke or traumatic brain injury, individuals with psychiatric disorders, such as schizophrenia, and other populations with suspected or known memory disorders. It is expected that more research on the CVLT3 will become available as the revision is more widely used. Rapid Reference 1.10 provides basic reference and publication information for the CVLT-C, CVLT-II, and CVLT3.

Rapid Reference 1.10

Publication Data for CVLT-C

Authors: Dean C. Delis, Joel H. Kramer, Edith Kaplan, and Beth A. Ober
Publication Date: 1994
What test measures: Verbal and Auditory Learning and Memory
Age Range: 5–16
Administration Time: 30 min administration plus a 20-min delay
Qualification of Examiners: Graduate- or professional-level training in psychological assessment
Publisher: NCS Pearson, Inc.
5601 Green Valley Drive
Bloomington, MN 55437
Order Phone Number: 1-800-627-7271
www.PearsonClinical.com
Price: Complete Kit: (as of March 2019) $235.75 (paper kit); 107.00 (scoring software)

Publication Data for CVLT-II

Authors: Dean C. Delis, Joel H. Kramer, Edith Kaplan, and Beth A. Ober
Publication Date: 2000
What test measures: Verbal and Auditory Learning and Memory
Age Range: 16–89
Administration Time:
Standard and Alternate Forms: 30 min administration plus 20-min delay (optional 10-min delay before optional 5-min forced-choice recognition condition)
Short Form: 20 min administration plus 10-min delay (optional 5-min delay before optional 5-min forced-choice recognition condition)
Qualification of Examiners: Graduate- or professional-level training in psychological assessment
Publisher: NCS Pearson, Inc.
5601 Green Valley Drive
Bloomington, MN 55437
Order Phone Number: 1-800-627-7271
www.PearsonClinical.com
Price: Complete Kit: (as of March 2019) $136.50 (manual); 105.00 (25 Standard/Alternate Record Forms); 84.00 (25 Short Record Forms); Digital Scoring priced by usage and subscription

Publication Data for CVLT3

Authors: Dean C. Delis, Joel H. Kramer, Edith Kaplan, and Beth A. Ober
Publication Date: 2017
What test measures: Verbal and Auditory Learning and Memory
Age Range: 5–16
Administration Time:
Standard and Alternate Forms: 30-min administration plus 20-min delay (optional 10-min delay before optional 5-min forced-choice recognition condition)
Short Form: 20-min administration plus 10-min delay (optional 5-min delay before optional 5-min forced-choice recognition condition)
Qualification of Examiners: Graduate- or professional-level training in psychological assessment
Publisher: NCS Pearson, Inc.
5601 Green Valley Drive
Bloomington, MN 55437
Order Phone Number: 1-800-627-7271
www.PearsonClinical.com
Price: Complete Kit: (as of March 2019) $409.00 (paper kit); 430.00 (paper kit plus 1-year online scoring subscription)

A NOTE ON CVLT NOMENCLATURE

There is a great deal of overlap in the structure, administration, scoring, and interpretation of the CVLT versions. For ease of communication, CVLT is used throughout this guide to refer to all three versions presented. This allows for quick communication of information without listing each individual measure. In instances where CVLT refers to the original

publication, this is indicated within the text. Moreover, when information is specific to one version of the CVLT, this is clearly indicated.

TEST YOURSELF

1. **Memory measures like the CVLT help differentiate normal, age-related decline from deficits related to disruption or impairment of memory processes.**
   True or False
2. **The California Verbal Learning Test was developed using in the process approach. This means that although _____ are important in determining severity of a memory impairment or what a person can do, _____ provide key information on how a person achieves those scores.**
   (a) overall recall scores, process scores
   (b) process scores, behavioral characteristics
   (c) error scores, immediate and delayed recall scores
   (d) error scores, process scores
3. **The CVLT-C was designed specifically for children ages 5–16. What is the main difference between the structure of the CVLT-C and the CVLT-II?**
   (a) The word lists are divided into semantic categories on CVLT-II.
   (b) The word lists are shorter on CVLT-C and involve categories for children.
   (c) CVLT-C reflects developmental processes specific to children.
   (d) The structure of the CVLT-C and CVLT-II are the same
4. **The Forced-Choice Recognition condition was added after the original CVLT. What was this condition specifically designed to measure?**
   (a) Recall errors
   (b) Recognition memory
   (c) Learning characteristics
   (d) Performance validity
5. **Semantic clustering, serial clustering, and serial position effects are all:**
   (a) learning strategies and characteristics
   (b) ways to enhance recall of information
   (c) measures of immediate and delayed memory
   (d) measures of recognition memory
6. **Why is reliability in memory assessments difficult to assess?**
   (a) Items are not scored individually
   (b) Trials are not scored individually
   (c) Items are not independent
   (d) Scores on memory measures are skewed

7. **All scores on CVLT-C, CVLT-II, and CVLT3 can be hand scored. True or False**
8. **Which of the following is a major change introduced in the CVLT3?**
   (a) new word lists were created for CVLT3
   (b) an alternate form and a short form were introduced
   (c) the age range was extended down to 5
   (d) normative scores are not corrected for gender only
9. **Memory processes are influenced by cognitive abilities other than specific encoding, storage, and retrieval functions. True or False**
10. **Practice effects on memory tests:**
    (a) are negligible.
    (b) can be observed for months or years.
    (c) are not relevant in clinical populations.
    (d) are short lived, similar to other cognitive measures.

Answers: 1. True, 2. a, 3. b, 4. d, 5. a, 6. c, 7. False, 8. d, 9. True, 10. b