Journal of the International Neuropsychological Society (2007), **13**, 354–358.

DOI: 10.1017/S1355617707070439

BRIEF COMMUNICATIONS

Utility of California Verbal Learning Test, Second Edition, recall discriminability indices in the evaluation of traumatic brain injury

JACOBUS DONDERS[1] AND JACOB B. NIENHUIS[2]

[1]Psychology Service, Mary Free Bed Rehabilitation Hospital, Grand Rapids, Michigan
[2]Department of Psychology, Calvin College, Grand Rapids, Michigan

(Received June 15, 2006; Final Revision October 2, 2006; Accepted October 4, 2006)

Abstract

The performance of 23 patients with moderate–severe traumatic brain injury on the California Verbal Learning Test, Second Edition (CVLT-II; Delis et al., 2000) was compared with that of 23 matched healthy controls to determine whether recall discriminability indices, which take into account both correct target recall and intrusive errors, would provide better diagnostic classification than traditional variables that are based exclusively on correct recall. Patients with traumatic brain injury recalled fewer correct words, and also made more intrusive errors, on CVLT-II short and long delay, free and cued recall trials ($p < .02$ for all variables after Stepdown Bonferroni correction). However, recall discriminability indices yielded a classification of clinical *versus* control participants (72%) that was not significantly different from one based on traditional variables (74%). We conclude that CVLT-II recall discriminability indices do not routinely provide an advantage over traditional variables in patients with traumatic brain injury. (*JINS*, 2007, *13*, 354–358.)

Keywords: Assessment, Learning, Memory, Traumatic brain injury, Sensitivity, Specificity

INTRODUCTION

The California Verbal Learning Test (CVLT; Delis et al., 1987) was used widely for many years in neuropsychological assessments to evaluate quantitative and process aspects of learning and memory. It had demonstrated sensitivity to a wide range of clinical conditions, ranging from various forms of amnesia and dementia (Delis et al., 1991; Hamilton et al., 2004) to traumatic brain injury (Curtiss et al., 2001; Wiegner & Donders, 1999). In addition, the CVLT had been shown to be a significant predictor of job performance after acquired brain dysfunction (Kibby et al., 1998). The revision of this instrument, the California Verbal Learning Test, Second Edition (CVLT-II), included standardization on a much more representative normative base (Delis et al., 2000). Some of the component indices intended to reflect specific underlying cognitive processes were modified (Stricker et al., 2002). Another important change was the incorporation of indices of discriminability for short and long delay, free and cued recall trials. These indices incorporate in a single score the number of correct words recalled as well as the number of intrusive errors (i.e., words that were not on the original list), with the goal of providing a potentially more accurate assessment of memory characteristics than traditional variables that are based exclusively on the number of target items recalled. The availability of recall discriminability variables presented a potential advantage because some patients may provide an average number of correct items during recall trials, not because they truly remember them selectively but because they are making plausible guesses, including a high number of intrusions along with correct responses.

Delis and colleagues (2005) demonstrated that the new CVLT-II recall discriminability indices were superior to traditional variables in distinguishing the performances of patients with Alzheimer's disease *versus* Huntington's dis-

Correspondence and reprint requests to: Jacobus Donders, Ph.D., Psychology Service, Mary Free Bed Rehabilitation Hospital, 235 Wealthy S.E., Grand Rapids, MI 49503. E-mail: jacobus.donders@maryfreebed.com

ease. However, this potential advantage has not yet been demonstrated with other diagnostic groups. The purpose of the current investigation was to determine the degree to which these recall discriminability indices would improve the distinction between patients with traumatic brain injury (TBI) and matched healthy controls. At this time, there is no published research concerning the criterion validity of the CVLT-II in patients with TBI. Several studies involving the original CVLT suggested that not all patients with TBI make a lot of recall errors, but that significantly elevated levels of intrusions may occur in those with more severe injuries (Wiegner & Donders, 1999) and who demonstrate more overall memory impairment (Curtiss et al., 2001). For these reasons, it is not yet clear if the new CVLT-II recall discriminability indices are routinely more informative than traditional measures that are based exclusively on target items correctly recalled after delays, in the evaluation of patients with TBI. The current exploratory investigation was pursued in an attempt to clarify this possibility.

Table 1. Demographic characteristics of patients with moderate–severe traumatic brain injury (TBI; $n = 23$) and controls ($n = 23$)

Variable	TBI		Control	
Years of age (*M*, *SD*)	34.44 (20.06)		35.26 (20.01)	
Gender (*n*, %)				
Female	8	(34.78)	8	(34.78)
Male	15	(65.22)	15	(65.22)
Level of education (*n*, %)				
9–11 years	6	(26.09)	5	(21.74)
12 years	8	(34.78)	8	(34.78)
13–15 years	4	(17.39)	6	(26.09)
16+ years	5	(21.74)	4	(17.39)
Ethnicity (*n*, %)				
African	1	(4.35)	2	(8.70)
Caucasian	20	(86.96)	18	(78.26)
Latino	2	(8.70)	3	(13.04)

Note. Data for control participants are from the standardization sample of the California Verbal Learning Test, Second Edition (CVLT-II).

METHODS

Research Participants

Following institutional review board approval, the 23 clinical participants were selected from a 2-year series of consecutive referrals to a regional Midwestern rehabilitation hospital, according to the following criteria: (a) diagnosis of moderate–severe TBI, defined as an external force to the head with witnessed loss of consciousness and associated neuroimaging evidence for an acute intracranial lesion; (b) age between 16 and 80 years at the time of psychometric assessment (to allow applicability of available test norms); (c) evaluation with the CVLT-II within 1 year after injury; (d) absence of any premorbid neurological, psychiatric, special education, or substance abuse history; and (e) absence of current financial compensation-seeking. During the time period that these data were collected, the CVLT-II was routinely administered to all referred patients with TBI, unless there were circumstances that would have argued against it (e.g., non-English language background, uncorrected hearing impairment). This investigation was limited to patients with moderate–severe TBI, because uncomplicated mild TBI is typically not associated with significant memory impairment (Iverson, 2005; Schretlen & Shapiro, 2003).

Following selection of the 23 clinical patients, an equal number of healthy control participants were selected from the CVLT-II standardization sample. These controls were matched to the clinical participants on the basis of age, gender, and (when possible) educational level and ethnicity. None of them had known neurological or psychiatric histories. Demographic characteristics of both groups are presented in Table 1.

The clinical patients were seen at an average of 118.74 days postinjury ($SD = 71.07$; median = 99; range, 42–270 days); always on an outpatient basis, and only when they were medically stable. The majority ($n = 17$; 74%) had sustained TBI as the result of a motor vehicle accident, with other injuries including falls, recreational activities, and assaults. Initial Glasgow Coma Scale scores spanned the full range of 3–15, but all participants had positive neuroimaging findings, including diffuse (e.g., edema, shear injury; $n = 10$) and/or focal (e.g., contusion, hemorrhage; $n = 17$) lesions. Median duration of coma, defined as the days until verbal commands were followed, was 1 day (range, 0–10 days).

MATERIALS AND PROCEDURE

The CVLT-II is an individually administered test of the ability to learn and remember verbally presented information. It involves the oral presentation of two word lists that each contain 16 items, including four words from each of four semantic categories. There are five trials of full presentation and immediate recall of the first list (A), followed by one-time presentation and immediate recall of a second, interference list (B). Measures of free recall and semantically cued recall are obtained immediately after the trial with list B, and again after a 20-min delay during which nonverbal tasks are administered. Then, a recognition trial is presented in which the examinee is asked to identify the 16 items from list A from a larger list that contains various distractor items. After another 10-min delay with nonverbal tasks, a final forced-choice trial is administered that is intended to assess level of effort and motivation on the test.

Overall performance on the CVLT-II is characterized in terms of a summary *T* score ($M = 50$; $SD = 10$), reflecting level of correct immediate recall over the five successive

https://doi.org/10.1017/S1355617707070439 Published online by Cambridge University Press

trials of the first list (Total A1–5), with higher scores reflecting better performance. In addition, the instrument allows for the computation of numerous *z* scores ($M = 0$; $SD = 1$) for a variety of other variables. The main variables of interest in the present study were those from the CVLT-II recall trials, including Short Delay Free Recall, Short Delay Cued Recall, Long Delay Free Recall, and Long Delay Cued Recall. For each of these recall trials, two types of *z* scores were obtained: (1) the traditional scores based on the number of correct target words reported, and (2) the new recall discriminability indices that reflect level of target recall relative to intrusion rate. Higher *z* scores reflect better performance on all of these variables. Finally, CVLT-II Intrusions were also considered. Higher *z* scores reflect worse performance on this particular variable.

RESULTS

The average composite CVLT-II *T* score of the clinical patients ($M = 44.57$; $SD = 11.68$) was statistically significantly worse than that of the control participants [$M = 52.08$; $SD = 7.03$; $F(1,44) = 7.01$; $p < .02$, $\eta^2 = .14$]. This finding reflects the general sensitivity of the instrument to the sequelae of moderate–severe TBI. However, of greater interest for the purposes of this investigation was the performance of the two groups on the recall trials because the clinical patients also had a higher *z* score for Intrusions ($M = .87$; $SD = 1.47$) than the control participants [$M = .02$; $SD = .68$; $F(1,44) = 6.29$; $p < .02$, $\eta^2 = .13$]. This finding reflects that, as a group, patients with moderate–severe TBI made more recall errors.

Figure 1a presents the performance of both groups on the traditional CVLT-II variables that are based exclusively on target words correctly recalled, whereas Figure 1b presents the performance of these groups on the CVLT-II recall discriminability indices for these same trials. Inspection of these figures suggests that, with both sets of variables, the patients with TBI typically did worse than the control participants. Independent samples *t* tests with the Stepdown Bonferroni method to balance the risk of Type I and Type II errors revealed that these group differences were statistically significant ($p < .02$) for all of the eight variables.

Because group differences in average performance do not convey information about classification accuracy that is important for clinical diagnosis, two separate discriminant function analyses were performed, one using the traditional CVLT-II recall indices as independent variables, and one using the recall discriminability indices. Group status (TBI *vs*. control) was the dependent variable in both analyses. The group of traditional indices classified 74% (34/46) of the participants correctly, with a sensitivity of 70% and a specificity of 78%. The group of recall discriminability indices classified 72% of the participants correctly, with a sensitivity of 74% and a specificity of 70%. The difference in overall classification accuracy fell far short of statistical significance ($z = .22$; $p > .40$).

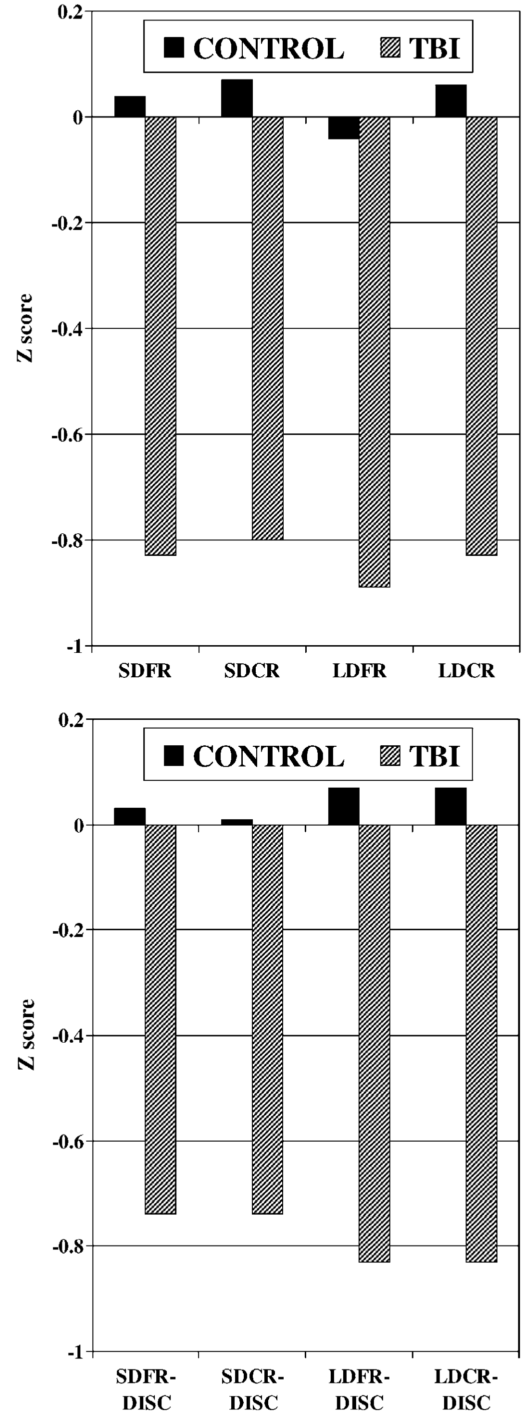

Fig. 1. a: Comparison of patients with moderate–severe traumatic brain injury (TBI; $n = 23$) and controls ($n = 23$) on the California Verbal Learning Test, Second Edition (CVLT-II), using traditional recall variables. *Note*. Data for control participants are from the standardization sample of the CVLT-II; copyright © 2000 by Harcourt Assessment, Inc. Used by permission. All rights reserved. **b:** Comparison of patients with moderate–severe TBI ($n = 23$) and controls ($n = 23$) on the CVLT-II, using new recall discriminability variables. *Note*. Data for control participants are from the standardization sample of the CVLT-II. Copyright © 2000 by Harcourt Assessment, Inc. Used by permission. All rights reserved. SDFR, short delay free recall; SDCR, short delay cued recall; LDFR, long delay free recall; LDCR, long delay cued recall.

DISCUSSION

The purpose of this investigation was to determine whether new CVLT-II indices that take into account both correct and incorrect responses on the short and long delay, free and cued recall trials would provide a distinct diagnostic advantage in the evaluation of patients with moderate–severe TBI over traditional variables that are based exclusively on correct target words recalled. The findings indicate that clinical *versus* control mean group differences were statistically significant for both the traditional scores and the recall discriminability indices. The indices had comparable sensitivity and specificity.

The current findings contrast with those of Delis and colleagues (2005), who demonstrated much more pronounced differences between patients with Alzheimer's disease *versus* Huntington's disease on the new CVLT-II recall discriminability scores than on the traditional recall variables. This difference in findings may be due to the fact that patients with Alzheimer's disease can often be expected to have fairly high levels of Intrusions (Delis et al., 1991), but this finding is not necessarily ubiquitous after TBI. Although the patients with TBI in the current investigation did have a statistically significantly higher level of Intrusions than the controls, their average z score on this variable was still less than 1 *SD* from the normative mean. *Post hoc* frequency counts suggested that only a minority ($n = 9$; 39%) of the participants with TBI had a z score ≥ 1.5 on this variable; a level of impairment that was considered unusual because it was not found in any of the control participants. These nine participants also had worse overall performance in terms of accurate recall, as reflected in the composite T score ($M = 39.01$; $SD = 10.74$), than the rest of the clinical patients ($M = 48.14$; $SD = 11.16$). Due to small subgroup sizes, it was not considered appropriate to subject these data to formal statistical analyses but these findings seem consistent with those from Curtiss et al. (2001) who found that high levels of Intrusions on the original CVLT typically occurred in patients who had worse overall memory impairment. Thus, elevated Intrusion rates in the absence of deficits in accurate recall may simply not be common enough after moderate–severe TBI for recall discriminability variables to yield a consistent or routine diagnostic advantage over traditional scores.

Another possibility is that the location of the lesion incurred in TBI may have relevance, because at least one prior study has reported that increased intrusions on the original CVLT tended to occur only in patients with dominant temporal lobe pathology (Crosson et al., 1993). Other studies with the CVLT-II have suggested a role of frontal, and especially ventromedial prefrontal, dysfunction in this regard (Baldo et al., 2002; Cato et al., 2004). We did not have the opportunity to perform detailed localization or volumetric lesion analyses with the current patient sample, partly because we often did not have access to the original neuroimaging scans and had to rely on radiology reports as to the presence or absence of acute intracranial pathology. Inclusion of more advanced morphometric analyses in future research might shed more light on the neuroanatomical bases of accurate as well as inaccurate recall on the CVLT-II after TBI.

A potential limitation of this investigation is that we used a referred convenience sample, consisting only of patients with moderate–severe injuries who had neuroimaging evidence for an acute intracranial lesion. Thus, our findings cannot be generalized to patients with mild TBI, and the potential of the CVLT-II to discriminate between various levels of injury severity still needs to be addressed in future research. At the same time, a relative strength of this investigation was that the participants were screened carefully for complicating premorbid factors, and were matched to controls with similar demographic backgrounds.

With these reservations in mind, we conclude that the new CVLT-II recall discriminability variables do not provide a diagnostic advantage in most patients with moderate–severe TBI. As a guideline for clinical practitioners, we suggest that, unless patients have highly elevated levels of Intrusions (e.g., $z \geq 1.5$), reliance on the traditional variables will likely suffice. However, because of somewhat modest sensitivity and specificity, regardless of which CVLT-II variables are considered, they should never be relied upon in isolation to determine the presence or absence of acquired memory impairment. A specific goal for future research is to evaluate in greater detail the construct validity of the CVLT-II in a much larger clinical sample of patients with TBI through confirmatory factor analysis, and to explore the possibility of subtypes of memory impairment on this test (e.g., limited encoding, retrieval problems) through cluster analysis. Such research is currently under way in our laboratory. Exploration of the utility of CVLT-II recall discriminability indices with other diagnostic groups that are at increased risk for memory impairment (e.g., multiple sclerosis, Korsakoff syndrome) is also desirable.

ACKNOWLEDGMENTS

This research was supported in part by a grant from the Campbell Foundation. We thank Harcourt Assessment for the provision of the data concerning the demographically matched controls. This research was based in part on standardization data of the California Verbal Learning Test, Second Edition (CVLT–II), copyright © 2000 by Harcourt Assessment Inc. Used with permission. All rights reserved. "California Verbal Learning Test" and "CVLT" are trademarks of Harcourt Assessment, Inc.; registered in the United States of America and other jurisdictions.

REFERENCES

Baldo, J.V., Delis, D.C., Kramer, J.H., & Shimamura, A.P. (2002). Memory performance on the California Verbal Learning Test–II: Findings from patients with focal frontal lesions. *Journal of the International Neuropsychological Society*, *8*, 539–546.

Cato, M.A., Delis, D.C., Abildskov, T.J., & Bigler, E. (2004). Assessing the elusive cognitive deficits associated with ventromedial prefrontal damage: A case of modern-day Phineas Gage. *Jour-*

nal of the International Neuropsychological Society, *10*, 453–465.

Crosson, B., Sartor, K.J., Jenny, A.B., Nabors, N.A., & Moberg, P.J. (1993). Increased intrusions during verbal recall in traumatic and nontraumatic lesions of the temporal lobe. *Neuropsychology*, *7*, 193–208.

Curtiss, G., Vanderploeg, R.D., Spencer, J., & Salazar, A.M. (2001). Patterns of verbal learning and memory in traumatic brain injury. *Journal of the International Neuropsychological Society*, *7*, 574–585.

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

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., Massman, P.J., Butters, N., Salmon, D.P., Cermak, L.S., & Kramer, J.H. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications of the assessment of memory disorders. *Psychological Assessment*, *3*, 19–26.

Delis, D.C., Wetter, S.R., Jacobson, M.W., Peavy, G., Hamilton, J., Gongvatana, A., Kramer, J.H., Bondi, M.W., Corey-Bloom, J., & Salmon, D.P. (2005). Recall discriminability: Utility of a new CVLT-II measure in the differential diagnosis of dementia. *Journal of the International Neuropsychological Society*, *11*, 708–715.

Hamilton, J.M., Salmon, D.P., Galasko, D., Delis, D.C., Hansen, L.A., Masliah, E., Thomas, R.G., & Thal, L.J. (2004). A comparison of episodic memory deficits in neuropathologically-confirmed dementia with Lewy bodies and Alzheimer's disease. *Journal of the International Neuropsychological Society*, *10*, 689–697.

Iverson, G.L. (2005). Outcome from mild traumatic brain injury. *Current Opinions in Psychiatry*, *18*, 301–317.

Kibby, M.Y., Schmitter-Edgecombe, M., & Long, C.J. (1998). Ecological validity of neuropsychological tests: Focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology*, *13*, 523–534.

Schretlen, D.J. & Shapiro, A.M. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry*, *15*, 341–349.

Stricker, J.L., Brown, G.G., Wixted, J., Baldo, J.V., & Delis, D.C. (2002). New semantic and serial clustering indices for the California Verbal Learning Test—second edition: Background, rationale, and formulae. *Journal of the International Neuropsychological Society*, *8*, 425–435.

Wiegner, S. & Donders, J. (1999). Performance on the California Verbal Learning Test after traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *21*, 159–170.