Journal of the International Neuropsychological Society (2008), **14**, 394–400.

doi: 10.1017/S1355617708080582

Differential neuropsychological test sensitivity to left temporal lobe epilepsy

DAVID W. LORING,[1,2] ESTHER STRAUSS,[3] BRUCE P. HERMANN,[4] WILLIAM B. BARR,[5] KENNETH PERRINE,[6] MAX R. TRENERRY,[7] GORDON CHELUNE,[8] MICHAEL WESTERVELD,[9] GREGORY P. LEE,[10] KIMFORD J. MEADOR,[1] AND STEPHEN C. BOWDEN[11]

[1]Department of Neurology, University of Florida, Gainesville, Florida
[2]Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida
[3]Department of Psychology, University of Victoria, Victoria, British Columbia
[4]Department of Neurology, University of Wisconsin, Madison, Wisconsin
[5]Department of Neurology, New York University, New York, New York
[6]Department of Psychology, Long Island Jewish Hospital, Manhasset, New York
[7]Department of Psychology, Mayo Clinic, Rochester, Minnesota
[8]Department of Neurology, University of Utah, Salt Lake City, Utah
[9]Department of Neurosurgery, Yale University, New Haven, Connecticut
[10]Department of Neurology, Medical College of Georgia, Augusta, Georgia
[11]Department of Psychology, University of Melbourne, Melbourne, Victoria, Australia

(RECEIVED June 5, 2007; FINAL REVISION January 17, 2008; ACCEPTED January 17, 2008)

Abstract

We examined the sensitivity of the Rey Auditory Verbal Learning Test (AVLT), California Verbal Learning Test (CVLT), Boston Naming Test (BNT), and Multilingual Aphasia Examination Visual Naming subtest (MAE VN) to lateralized temporal lobe epilepsy (TLE) in patients who subsequently underwent anterior temporal lobectomy. For the AVLT (n = 189), left TLE patients performed more poorly than their right TLE counterparts [left TLE = 42.9 (10.6), right TLE = 47.7 (9.9); $p < .002$ (Cohen's d = .47)]. Although statistically significant, the CVLT group difference (n = 212) was of a smaller magnitude [left LTE = 40.7 (11.1), right TLE = 43.8 (9.9); ($p < .03$, Cohen's d = .29)] than the AVLT. Group differences were also present for both measures of confrontation naming ability [BNT: left LTE = 43.1 (8.9), right TLE = 48.1 (8.9); $p < .001$ (Cohen's d = .56); MAE VN: left TLE = 42.2, right TLE = 45.6, $p = .02$ (Cohen's d = .36)]. When these data were modeled in independent logistic regression analyses, the AVLT and BNT both significantly predicted side of seizure focus, although the positive likelihood ratios were modest. In the subset of 108 patients receiving both BNT and AVLT, the AVLT was the only significant predictor of seizure laterality, suggesting individual patient variability regarding whether naming or memory testing may be more sensitive to lateralized TLE. (*JINS*, 2008, *14*, 394–400.)

Keywords: Memory, Naming, Neuropsychology, Epilepsy surgery, Anterior temporal lobectomy, Logistic models

INTRODUCTION

Neuropsychological assessment serves different roles that are emphasized to varying degrees across epilepsy surgery institutions. The two primary purposes of neuropsychological testing are (1) to identify focal functional deficits associated with lateralized temporal lobe seizure onset, and (2) to assess the likelihood of postoperative memory and language change following surgery. Despite these slightly different objectives, common neuropsychological instruments are used because both seizure onset laterality determination and cognitive outcome prediction involve verbal learning/memory and naming assessment.

The greatest postoperative cognitive risk following anterior temporal lobectomy (ATL) is memory decline, and in particular, verbal memory (Milner, 1972). Memory risk is greatest following resection of the language dominant temporal lobe, and when the diseased temporal lobe to be resected still actively contributes to memory formation (i.e., high functional adequacy; Chelune, 1995). The risk of naming decline

Correspondence and reprint requests to: David W. Loring, Ph.D., McKnight Brain Institute, University of Florida, PO Box 100236, Gainesville, FL 32610. E-mail: david.loring@neurology.ufl.edu

following left ATL is also well-established (Bell et al., 2000; Langfitt & Rausch, 1996; Saykin et al., 1995; Schwarz et al., 2005), and the role of the hippocampus in naming performance both pre- and postoperatively is increasingly appreciated (Hamberger et al., 2007; Seidenberg et al., 2005).

This study reports the sensitivity of four commonly used neuropsychological tests of left hemisphere function to lateralized temporal lobe seizure onset in temporal lobe epilepsy (TLE). Unlike other diseases in which there have been specific recommendations to include in neuropsychological assessment protocols (e.g., Benedict et al., 2006), there are no similar proposals in the neuropsychology of epilepsy. Differential test sensitivity may inform future practice, either during the development of eventual practice guidelines or affecting test selection for research protocols. Our primary analyses were directed to compare the relative sensitivities of these measures to lateralized dysfunction. As a secondary goal, we examined classification accuracy of the measures, both individually and in combination with a second measure to more fully characterize the contributions of each test to seizure onset classification (Hosmer & Lemeshow, 2000; Strauss et al., 2005).

METHOD

Neuropsychological Tests

Verbal learning and memory were tested with either the Rey Auditory Verbal Learning Test (AVLT) or California Verbal Learning Test (CVLT). The AVLT is a serial word learning task in which 15 words are presented over five learning trials (Schmidt, 1996), followed by a second learning list, and then free-recall of the original list of 15 words.

The CVLT is a serial word learning task with a structure patterned after the AVLT (Delis et al., 1987), with five learning trials, a single presentation of a second list with recall, followed by free recall of the original list. After a delay of approximately 20 min, free and cued recall, and recognition is tested. In contrast to the AVLT, the CVLT contains 16 words from four semantic categories (i.e., spices and herbs, fruits, tools, and clothing).

Visual naming was assessed using either the Boston Naming Test (BNT) or the Multilingual Aphasia Examination Visual Naming (MAE VN) subtest. The BNT (Kaplan et al., 1983) consists of 60 line drawing of objects that vary in their frequency of use (e.g., "bed" to "abacus"). The MAE VN (Benton et al., 1994) also uses line drawings as stimuli, although unlike the BNT, parts of the main object are also used as stimulus items (e.g., "thumb" in addition to "hand"). The test comprises 30 items.

Tests were administered according to standard directions from the test manual for MAE VN and for CVLT, and according to standardized instructions for BNT and AVLT (Spreen & Strauss, 1991). A single dependent variable was analyzed fore each test (AVLT: total recall across trials; CVLT: total recall across trials; BNT: total correct without phonemic cuing; MAE VN; total correct). Delayed recall or recognition measures were not included in the database for both verbal memory tests, precluding analysis of other potential measures of interest. However, total recall across trials is the most reliable measure for either memory test (Strauss et al., 2006). Raw rather than standardized scores were used in all analyses.

Subjects

Subjects were retrospectively identified from the Bozeman Neuropsychology Epilepsy Database. This is a de-identified archival database developed from the informal collaborations from neuropsychology programs at eight epilepsy centers that were willing to share clinical data and neuropsychological findings. The database is named after Bozeman, Montana, the site of the first meeting of participating centers, and has contributed to multiple multicenter epilepsy studies (e.g., Barr et al., 1997; Chelune et al., 1998; Loring et al., 1999; Strauss et al., 2000; Westerveld et al., 2000). Participating centers included Baptist Memorial Hospital (Memphis), Cleveland Clinic Foundation, Long Island Jewish Hospital, New York University, Mayo Clinic, Medical College of Georgia, University of Victoria, and Yale University. These data were collected in compliance with research regulations in place at the time of data entry at each participating institution. Because this was an informal collaboration without independent financial support from an extramural source, the criteria for evaluation at each participating institution were used for patient characterization. Unlike multicenter clinical trials/observational studies, there were no formal mechanisms to standardize and evaluate clinical procedures across participating centers such as case report forms or study monitor visits.

TLE patients were included if they had undergone Wada testing to establish cerebral language representation; only patients determined to be left cerebral language dominant were included. Patients with known lesions other than hippocampal sclerosis (e.g., ganglioglioma, dysembryoplastic neuroepithelial tumor (DNET), arteriovenous malformations) were excluded. However, because data entry began in the late 1980s, there are an unknown number of patients with lesions such as migrational disorders who were not identified using magnetic resonance imaging techniques the time of evaluation. Seizure onset laterality was determined according to clinical criteria in place at each participating institution, but generally consisted of multiple ictal and interictal electroencephalographic (EEG) abnormalities recorded with various combinations of surface and intracranial electrodes. All TLE patients subsequently underwent ATL.

There were 204 patients who subsequently underwent left ATL and 197 patients who underwent right ATL. From this group, 189 patients (left = 91; right = 98) were identified who were administered the AVLT and 212 patients were administered the CVLT (left $n = 113$; right $n = 99$). There were 135 patients who were administered the Boston Nam-

https://doi.org/10.1017/S1355617708080582 Published online by Cambridge University Press

Table 1. Means, standard deviations, and levels of statistical significance for group demographics including WAIS-R scores

	Left TLE			Right TLE				
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*	*t* value	*P* level
Age (years)	204	31.4	8.8	197	31.8	9.9	0.43	.67
Education (years)	195	12.6	2.5	184	12.5	2.4	0.47	.64
Seizure Onset (years)	204	9.9	9.0	197	12.3	10.0	2.56	.01
Seizure Duration (years)	204	21.5	10.6	196	19.6	10.8	1.73	.08
WAIS-R FSIQ	197	87.7	11.8	190	88.4	11.5	0.60	.55
WAIS-R VIQ	130	88.2	12.1	127	90.3	12.9	1.30	.20
WAIS-R PIQ	130	90.3	13.4	127	89.1	12.8	0.70	.48

Note. WAIS-R = Wechsler Adult Intelligence Scale–Revised; FSIQ = Full-Scale IQ; VIQ = Verbal IQ; PIQ = Performance IQ; TLE = temporal lobe epilepsy.

ing Test (left = 69; right = 66), and 173 were administered the MAE Visual Naming Test (left = 79; right = 94).

RESULTS

Sample characteristics and mean Full-Scale IQ (FSIQ) levels are presented in Tables 1 and 2. No significant differences were present with respect to age, education, sex, or handedness. Patients with left TLE had a slightly earlier age of habitual seizure onset compared with right TLE patients with the duration of epilepsy approaching statistical significance. No significant group differences in Wechsler Adult Intelligence Scale-Revised FSIQ, Verbal IQ, or Performance IQ were observed.

Results of verbal memory and confrontation naming tasks are presented in Table 3. Patients with left seizure onset performed significantly more poorly on AVLT score based upon the sum across trials ($p < .002$). Although a significant group difference was present for the CVLT sum across trials ($p < .03$), this was of a smaller magnitude. On confrontation naming, significant differences were seen for both the BNT and MAE VN, although the magnitude of effect was greater for the BNT ($p < .001$) than the MAE VN ($p < .02$). Effect sizes (Cohen's d) are also presented in Table 3.

We next modeled classification using logistic regression (SPSS 14.0). In the first analysis, the AVLT total was used to predict side of seizure onset. Default program values and the enter method of independent variable selection were used. Because left and right TLE patients differed on habitual seizure onset age, we entered this first in the regression analysis. Because it did not contribute significantly to the prediction ($p = .215$), habitual seizure onset age was not included in subsequent analysis. When AVLT was included as a predictor, a nonsignificant Hosmer and Lemeshow statistic was obtained [Hosmer & Lemeshow $\chi^2(N = 189; df = 7) = 8.46; p = .29$]. The Hosmer and Lemeshow statistic is a test of model fit, with significant values indicating lack of fit in the model when tested against the observed data (Hosmer & Lemeshow, 2000). The logistic regression coefficient for AVLT was significant. In logistic regression, the predicted classification value is $1/(1 + e^{-z})$. In this analysis, $z = -2.036 + .046 * (AVLT)$. Standard errors for the intercept and regression coefficients were .711 and .015, respectively. This indicates that, when AVLT total is less than 44, patients are classified as belonging to the left seizure focus group, with a predicted classification value less than .5 (predicted classification values range from 0 to 1).

Details of the correct classifications produced by the regression model are shown in Table 4. Sensitivity, defined as the number of patients with left seizure focus correctly classified by the regression model (AVLT total less than 44) is equal to 44/(44 + 47), or 48.4%. Specificity, defined as the number of patients with right seizure focus correctly classified by the regression model (AVLT total greater than or equal to 44) is equal to (66/(32 + 66), or 67.3%. As shown in Table 3, the positive likelihood ratio (LR+) was 1.48 [95% confidence interval [CI] = 1.04 to 2.11; for calculations see Strauss et al., 2005; see also www.cebm.utoronto.ca/practise/ca/statscal/]. Although this LR+ is significantly different from one, in terms of the 95% confidence interval, the observed value indicates only modest diagnostic value. The negative likelihood ration (LR−) is .77 (95% CI = .60 to .98), which, although also significantly different from one, is of modest diagnostic value (Strauss et al., 2005). Logistic regression analysis was re-run to test for a nonlinear (quadratic) relationship between AVLT

Table 2. Frequencies by seizure onset laterality and levels of statistical significance for sex and handedness

Variable	Left TLE	Right TLE	χ^2(df)	*P* level
Sex				
Male	91	97	0.86 (1)	.35
Female	113	100		
Handedness				
Dextral	182	168	1.75 (1)	.19
Nondextral	21	29		

Note. TLE = temporal lobe epilepsy.

https://doi.org/10.1017/S1355617708080582 Published online by Cambridge University Press

Table 3. Means and standard deviations for verbal memory and confrontation naming performances for both left and right TLE groups

Test	Focus	*n*	*M*	*SD*	*t* value	*P* value	Cohen's d	LR+ (95% CI)	LR− (95% CI)
AVLT	Left	91	42.9	10.6	3.21	.002	.47	1.48 (1.04–2.11)	.77 (.60-.98)
	Right	98	47.7	9.9					
CVLT	Left	113	40.7	11.1	2.20	.03	.30	1.14 (.93–1.39)	.79 (.54–1.14)
	Right	99	43.8	9.9					
BNT	Left	69	43.1	8.9	3.27	.001	.57	1.91 (1.26–2.90)	.60 (.44-.83)
	Right	66	48.1	8.9					
MAE VN	Left	94	42.3	8.8	2.42	.02	.36	1.23 (.99–1.52)	.65 (.42–1.01)
	Right	79	45.6	9.3					

Note. Tests scores are uncorrected raw scores. Also shown are positive and negative likelihood ratios (LR+ and LR−) for each test. Note that the LRs are based on the classification analysis reported in the text and do not relate directly to analysis of group means. AVLT = Auditory Verbal Learning Test sum of recall; CVLT = California Verbal Learning Test sum of recall; BNT = Boston Naming Test spontaneous and semantic cue; MAE VN = Multilingual Aphasia Examination Visual Naming spontaneous.

total score and seizure focus, but forced entry produced a nonsignificant quadratic coefficient ($p = .49$). Because demographic variables (e.g., age, education, gender, and handedness) differed slightly across groups (in absolute terms only), logistic regression analyses were run to test for the inclusion of significant predictions of side of seizure focus. None of these variables were significant in any analyses and, therefore, were not considered further.

We then tested the hypothesis that BNT predicts side of seizure focus with logistic regression using the same default parameters described above. With BNT entered as a single independent variable, the Hosmer and Lemeshow test was significant [Hosmer & Lemeshow $\chi^2(N = 135, \text{df} = 7) = 14.82, p = .04$], although the intercept and regression coefficients were also significant. Because the significant Hosmer and Lemeshow test suggested some lack of fit, a quadratic term for BNT was examined using forward inclusion. However, when the quadratic term for BNT entered the equation as a significant predictor, the linear term for BNT was no longer significant. Therefore, the quadratic model for BNT was not considered viable. Although BNT was moderately negatively skewed, examination of various data transformations failed to identify any way to improve model fit over the simple linear model.

The classification table resulting from the simple logistic regression model containing only the linear term for BNT yielded a sensitivity of 58% and specificity of 70%. The associated LR+ was 1.91 (95% CI = 1.26 to 2.90). The LR− was .60 (95% CI = .44 to .83). Although these LRs are significantly different from one, the 95% CIs for both the LR+ and the LR− include the values for the respective LRs for AVLT total reported above. LRs and confidence intervals for tests examined are also summarized in Table 2. Therefore, the LRs for the BNT do not provide a better prediction of side of seizure focus than that provided by the AVLT.

To directly compare AVLT and BNT, we performed additional logistic regression using the same default parameters except that both AVLT and BNT scores were included as predictors and analyzed using forward inclusion in the sample of 108 patients who had scores on both tests. The analysis produced a nonsignificant Hosmer and Lemeshow test [Hosmer & Lemeshow $\chi^2(N = 108; \text{df} = 8) = 6.43; p = .60$]. The only predictor entered on forward inclusion was AVLT total (regression coefficient: $p = .001$). BNT just failed to meet the entry significance criterion ($p = .055$). Method of forward entry did not alter the results. While it is tempting to assume that a larger sample might facilitate identification of a significant regression coefficient associated with BNT, forced entry into the equation of BNT only improved model fit by a small amount (difference in Cox and Snell $R^2 = .03$, and difference in Nagelkerke $R^2 = .04$) compared with the model with AVLT score only (Cox and Snell $R^2 = .11$, and Nagelkerke $R^2 = .14$). In addition, computation of the LRs after forced entry of the BNT score did not reveal any useful increments in correct classification.

Table 4. Classification table for prediction of seizure focus from the logistic regression model including Auditory Verbal Learning Test sum of recall

		Observed seizure focus		
		Left	Right	Likelihood ratios (95% CI)
Predicted seizure focus	Left	44 (48.4%)[1]	32 (32.7%)	1.48 (1.04–2.11)
	Right	47 (51.6%)	66 (67.3%)	0.77 (0.60–0.98)

Note. CI = 95% confidence intervals associated with predicted scores.
[1]Frequency (%).

https://doi.org/10.1017/S1355617708080582 Published online by Cambridge University Press

Prediction of seizure focus laterality was then examined in the subset of patients administered the CVLT. Using the same default analysis values as above, age of seizure onset was entered first followed by CVLT [Hosmer & Lemeshow $\chi^2(N = 212; \text{df} = 8) = 11.40, p = .18$]. In this analysis, age of seizure onset was significant ($p = .034$) and CVLT just failed to significantly predict side of seizure onset ($p = .055$). Although not significant, a separate logistic regression was conducted on the CVLT score alone to derive classification performance for comparison with the other tests reported in this study. Prediction of left seizure focus from CVLT alone resulted in sensitivity = 69% and specificity = 39%. The associated LR+ was 1.14 (95% CI = .93 to 1.39). The LR− was .79 (95% CI = .54 to 1.14). Both of these confidence intervals included one, and, therefore, illustrate in terms of explicit diagnostic efficiency that the CVLT cannot be regarded as providing useful diagnostic information.

In a final logistic regression analysis, MAE VN also entered into a prediction of side of seizure onset. Again, age at seizure onset was initially tested for inclusion but was not significant in this sample ($p = .144$) and was dropped from this analysis. Prediction of seizure onset laterality by MAE VN alone [Hosmer & Lemeshow $\chi^2(N = 173; \text{df} = 8) = 9.04; p = .34$] was associated with a significant regression coefficient ($p = .02$). The logistic classification resulted in sensitivity = 75% and specificity = 39%. The associated LR+ was 1.23 (95% CI = .99 to 1.52). The LR− was .65 (95% CI = .42 to 1.01). Because both confidence intervals included 1, the LRs do not indicate useful diagnostic information.

DISCUSSION

These findings suggest differential neuropsychological test sensitivity to lateralized temporal lobe epilepsy for common tests used to assess verbal learning and confrontation naming. There were also differences in the magnitude of statistical results. Specifically, AVLT appears superior to CVLT in discriminating left from right temporal lobe seizure onset (Cohen's d of .47 *vs*. .30). When contrasting confrontation language measures, BNT was superior to MAE VN (Cohen's d of .57 *vs*. .36). Although three tests were associated with significant logistic regression equations and significant regression coefficients and were able to statistically differentiate left from right TLE, the results were of modest significance. Although there are significant theoretical implications of whether verbal learning or confrontation naming may be more sensitive to left TLE, this cannot be adequately addressed in this data set and must await future research.

These group differences can also be interpreted in terms of diagnostic utility and LRs. On the individual patient level, all of the tests except CVLT provided statistically significant logistic regression formula. However, when logistic regression classifications for all tests were examined, sensitivity and specificity were disappointing, reflecting the overlapping distributions of scores on all of these tests in patients with left and right TLE. When the sensitivity and specificity information was converted to LRs, only AVLT and BNT produced LRs that were significantly different from 1.

Although we anticipated that AVLT and BNT would provide a more powerful diagnostic model when combined into a single regression equation, this was not the case, with the only significant predictor being AVLT. Although BNT just failed to meet the forward entry criteria, this failure did not appear to be a function of low statistical power. Forced entry of the BNT score added only a small increment in logistic regression approximations to R^2 (3–4% change) and, although not reported, there was no useful change in the LRs after inclusion of BNT scores. Therefore, on the basis of the present results, the likelihood of having left TLE focus appears to be identified by the AVLT better than by any other test examined. Nevertheless, all of the correct classification data produced LRs of only modest magnitude and the best predictor, the AVLT, could only be described as a slightly useful diagnostic test (Strauss et al., 2005). Although both CVLT and MAE VN scores produced statistically significant predictions of side of seizure onset, the resulting sensitivity and specificity statistics were suboptimal, and the LRs did not indicate significant improvement over base rate classifications.

Consistent with the absolute, albeit modest, differences in the Cohen's d's associated with BNT and MAE VN (Table 1), the LR+ for the MAE VN fell below the lower bound of the 95% confidence interval for the LR+ derived from the BNT. This finding indicates that BNT is a significantly better diagnostic test of naming than MAE VN, perhaps because the BNT has twice as many items, although the BNT is still of only modest usefulness in absolute terms. A similar interpretation applies to differences in respective LRs associated with the AVLT and CVLT. These examples illustrate that LRs can be use as a direct test of incremental validity. If one test produces a LR that falls above the confidence interval for a second test, then the first test can be considered significantly more useful in diagnostic terms. That is, interval estimation using LRs may provide the most direct method for evaluating the hitherto nebulous concept of incremental validity.

It was impossible to examine all possible combinations of tests to determine which might produce the best discrimination between seizure onset laterality, because the same tests were not administered to all patients. Furthermore, because subjects were not randomly assigned to tests, potential differences in the criteria for surgery across centers may be contributing to our results. In addition, although the CVLT uses a longer list length and semantic relationships among its words, fewer words were recalled across trials compared with the AVLT. There are no systematic comparisons between the CVLT and AVLT in clinical populations, so whether this reflects a systematic difference in test difficulty or is an artifact of sample specific characteristics cannot be determined.

It is unfortunate that, when this database was being constructed, delayed recall or recognition were not included for both the CVLT and AVLT. Thus, no conclusion about

https://doi.org/10.1017/S1355617708080582 Published online by Cambridge University Press

whether differences exist for other variables such as delayed recall or recognition, or the sensitivity of these measures relative to confrontation naming. A further limitation is the potential for criterion contamination because these tests were used to varying degree to determine surgical candidacy. We note, however, that the numbers of patients excluded from candidacy at any center due to neuropsychological test findings was extremely small (<1%), with inclusion criteria relying primarily on EEG and clinical semiology.

Our data were collected using the original 1987 CVLT. The CVLT was revised in 2000, and in addition to increasing the normative sample, a new list of words was included in the revision as well as new approaches to analyze aspects of learning, memory, and motivational status. Although the original and revised versions of the test are informally treated as equivalent with respect to task difficulty and sensitivity to clinical disease, there may be systematic performance differences between the two test versions. Thus, given the difference in stimulus materials between CVLT editions, it is possible that different sensitivities between test versions may exist.

The overall test structure of the AVLT and CVLT is the same, with the primary difference between the two word learning tasks being the semantic relationships between words present in the CVLT and no obvious semantic link between words in the AVLT. The increased sensitivity of the AVLT to left temporal abnormality may reflect its increased sensitivity to deficits in relational learning. In contrast to the CVLT, AVLT words do not show a clear semantic relationship and subjects may have to rely on more effortful strategies (e.g., temporal tagging) to form relationships among items.

The relative contribution of the BNT and AVLT was not consistent in our analyses. When examined on a test by test basis regarding their ability to classify individual patients, only the BNT significantly predicted patients with left TLE. However, BNT classification fell within the 95% confidence interval for the AVLT, indicating that these two tests do not necessarily differ in classification ability. Further clouding an interpretation of the relative contribution of these tests is our analysis in which both tests were examined in the subset of patients in whom both tests were administered. In this analysis, the AVLT was actually superior to the BNT. It is now well established that the disruptive effects of focal temporal lobe seizure onset, demonstrated by both functional and structural changes, extend far beyond a single area of focal abnormality associated with the seizure focus (Burneo et al., 2004; Sawrie et al., 2000; Seidenberg et al., 2005; Theodore & Gaillard, 2002; Vinton et al., 2007) and that nontemporal lobe seizure effects will also contribute the sensitivity of individual neuropsychological tests to temporal lobe epilepsy. Thus, prospective studies in which both neuropsychological tests are administered to the entire sample using contemporary imaging measures (i.e., magnetic resonance imaging, positron emission tomography, magnetic resonance spectroscopy) will be necessary to address this important issue.

ACKNOWLEDGMENTS

An earlier version of this study was presented at the Annual Meeting of the International Neuropsychological Society, Portland, Oregon, February 10, 2007. These data have not been previously published, either electronically or in print media. There are no financial relationships associated with these results that may reflect a conflict of interest or be perceived to reflect a conflict of interest. This project was not supported by extramural funding.

REFERENCES

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

Bell, B.D., Davies, K.G., Hermann, B.P., & Walters, G. (2000). Confrontation naming after anterior temporal lobectomy is related to age of acquisition of the object names. *Neuropsychologia*, *38*, 83–92.

Benedict, R.H., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., & Weinstock-Guttman, B. (2006). Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society*, *15*, 549–558.

Benton, A.L., Hamsher, K.D., & Sivan, A.B. (1994). *Multilingual Aphasia Examination-Third Edition*. Odessa, FL: Psychological Assessment Resources.

Burneo, J.G., Knowlton, R.C., Faught, E., Martin, R., Sawrie, S., & Kuzniecky, R.I. (2004). Chronic temporal lobe epilepsy: Spatial extent and degree of metabolic dysfunction studied with magnetic resonance spectroscopy (MRS). *Epilepsy Research*, *62*, 119–124.

Chelune, G.J. (1995). Hippocampal adequacy versus functional reserve: Predicting memory functions following temporal lobectomy. *Archives of Clinical Neuropsychology*, *10*, 413–432.

Chelune, G.J., Naugle, R.I., Hermann, B.P., Barr, W.B., Trenerry, M.R., Loring, D.W., Perrine, K., Strauss, E., & Westerveld, M. (1998). Does presurgical IQ predict seizure outcome after temporal lobectomy? Evidence from the Bozeman Epilepsy Consortium. *Epilepsia*, *39*, 314–318.

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

Hamberger, M.J., Seidel, W.T., Goodman, R.R., Williams, A., Perrine, K., Devinsky, O., & McKhann, G.M., Jr. (2007). Evidence for cortical reorganization of language in patients with hippocampal sclerosis. *Brain*, *130*, 2942–2950.

Hosmer, D.W. & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: John Wiley & Sons, Inc.

Kaplan, E.F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia: Lea & Febiger.

Langfitt, J.T. & Rausch, R. (1996). Word-finding deficits persist after left anterotemporal lobectomy. *Archives of Neurology*, *53*, 72–76.

Loring, D.W., Strauss, E., Hermann, B.P., Perrine, K., Trenerry, M.R., Barr, W.B., Westerveld, M., Chelune, G.J., Lee, G.P., & Meador, K.J. (1999). Effects of anomalous language representation on neuropsychological performance in temporal lobe epilepsy. *Neurology*, *53*, 260–264.

Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery*, *19*, 421–446.

Sawrie, S.M., Martin, R.C., Gilliam, F.G., Faught, R.E., Maton, B., Hugg, J.W., Bush, N., Sinclair, K., & Kuzniecky, R.I. (2000). Visual confrontation naming and hippocampal function: A neural network study using quantitative (1)H magnetic resonance spectroscopy. *Brain*, *123*(Pt 4), 770–780.

Saykin, A.J., Stafiniak, P., Robinson, L.J., Flannery, K.A., Gur, R.C., O'Connor, M.J., & Sperling, M.R. (1995). Language before and after temporal lobectomy: Specificity of acute changes and relation to early risk factors. *Epilepsia*, *36*, 1071–1077.

Schmidt, M. (1996). *Rey Auditory and Verbal Learning Test: A Handbook*. Los Angeles: Western Psychological Services.

Schwarz, M., Pauli, E., & Stefan, H. (2005). Model based prognosis of postoperative object naming in left temporal lobe epilepsy. *Seizure*, *14*, 562–568.

Seidenberg, M., Geary, E., & Hermann, B. (2005). Investigating temporal lobe contribution to confrontation naming using MRI quantitative volumetrics. *Journal of the International Neuropsychological Society*, *11*, 358–366.

Spreen, O. & Strauss, E. (1991). *A Compendium of Neuropsychological Tests*. New York: Oxford University Press.

Strauss, E., Semenza, C., Hunter, M., Hermann, B., Barr, W., Chelune, G., Lavdovsky, S., Loring, D., Perrine, K., Trenerry, M., & Westerveld, M. (2000). Left anterior lobectomy and category-specific naming. *Brain & Cognition*, *43*, 403–406.

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd ed.). New York: Oxford University Press.

Strauss, S.E., Richardson, W.S., Glasziou, P., & Haynes, R.B. (2005). *Evidence-based Medicine: How to Practice and Teach EBAM* (3rd ed.). Edinburgh, UK: Elsevier Churchill-Livingstone.

Theodore, W.H. & Gaillard, W.D. (2002). Neuroimaging and the progression of epilepsy. *Progress in Brain Research*, *135*, 305–313.

Vinton, A.B., Carne, R., Hicks, R.J., Desmond, P.M., Kilpatrick, C., Kaye, A.H., & O'Brien, T.J. (2007). The extent of resection of FDG-PET hypometabolism relates to outcome of temporal lobectomy. *Brain*, *130*, 548–560.

Westerveld, M., Sass, K.J., Chelune, G.J., Hermann, B.P., Barr, W.B., Loring, D.W., Strauss, E., Trenerry, M.R., Perrine, K., & Spencer, D.D. (2000). Temporal lobectomy in children: Cognitive outcome. *Journal of Neurosurgery*, *92*, 24–30.