Revising the Rey–Osterrieth: Rating Right Hemisphere Recall

David W. Loring, Gregory P. Lee, and Kimford J. Meador

Medical College of Georgia

Recall performance of the Rey-Osterrieth Complex Figure was examined in patients with partial complex seizures originating from either the right or left temporal lobe and who underwent subsequent unilateral temporal lobectomy. A scoring system was developed to assess the types of errors frequently observed in the recall of patients with right temporal lobe epilepsy (TLE), but absent in left TLE patients. The scoring system was initially developed on a single group of patients, and then "cross-validated" on an independent sample. Performance analysis of the cross-validation group revealed a significant difference in the frequency of right hemisphere errors. In contrast, no significant difference using standard quantitative scoring was present. By applying the new scoring critiera alone, a rater blind to seizure onset correctly predicted seizure laterality in 15/18 of the cross-validation patients. These results suggest that evaluation of qualitative errors may be a valuable adjunct to standard scoring criteria, thereby extending the range of applications for this test.

The assessment of material specific recent memory functions mediated by mesial temporal lobes of the dominant and non-dominant hemispheres is an important aspect of contemporary neuropsychology. In general, the ability to learn verbal information is dependent on the dominant hemisphere temporal lobe and the learning of visual-spatial material requires intact non-dominant mesial temporal lobe structures.

The use of verbal and figural tests to assess specific memory deficits in patients with partial complex seizures originating from a single temporal lobe has produced inconsistent results, even when identical tests and procedures are employed. For example, Glowinski (1973) failed to detect right/left

differences in temporal lobe epilepsy (TLE) patients for both immediate and delayed Logical Memory and Visual Reproduction performance on the Wechsler Memory Scale. In contrast, Delaney, Rosen, Mattson, and Novelly (1980) report right/left group differences on both subtests.

This lack of agreement may be due to several factors. When clinical groups are studied using an *ex post facto* design, group differences are related to subject inclusion characteristics (Campbell & Stanley, 1963; Meehl, 1975). As the ability to diagnose seizure disorders improves, along with improved classification of seizure types, the probability of group differences increases. However, we do not consider subject variables to be the sole explanation for the reported inconsistencies since they exist across diverse patient populations using a variety of tasks (e.g., Mayeux, Brandt, Rosen, & Benson, 1980; Milner, 1968, 1975; Ladavus, Umiltà, & Provinciali, 1979).

We believe the contradictory reports of material specific memory function in unilateral TLE result from a combination of non-comparable scoring criteria across studies in conjunction with tests that may not be sufficiently sensitive to figural memory deficits. For example with verbal memory, the use of Logical Memory can produce different results depending on whether verbatim, gist, or half-credit scoring criteria are employed (see Loring & Papanicolaou, 1987). It has been further suggested that the visual-spatial stimuli may be of insufficient complexity to produce right/left TLE differences (Milner, 1975). However, even using the Rey–Osterrieth Complex Figure (CF), which is more complex than either WMS Visual Reproduction or Benton Visual Retention tests, no right/left group differences have been detected (Mayeux et al., 1980).

It has been our experience that although standard CF quantitative scoring does not consistently predict site of seizure focus in patients undergoing evaluation for temporal lobectomy, subjective analysis of patient's qualitative performance frequently suggests laterality of seizure onset. Although other authors have devised scoring systems to assess qualitative performances (e.g., Binder, 1982; Waber & Holmes, 1986), item distortion and misplacement are not directly assessed. Consequently, we undertook the task of developing a scoring system sensitive to the types of errors frequently observed in patients with right TLE, but not adequately captured using standard scoring criteria. We report here our initial attempt at such a scale.

METHODS

Twenty-nine patients undergoing pre-operative evaluation for temporal lobectomy served as subjects. Table 1 shows demographic data including IQ measures for right and left TLE patients. Group assignment was determined from the side of subsequent temporal lobectomy, which in turn was based upon multiple ictal EEG recordings from scalp, sphenoidal and/or intracere-

TABLE 1
Demographic Information

	Left Temporal	Right Temporal
Age	28.1 (7.1)	32.6 (13.1)
FSIQ	87.6 (10.44)	93.5 (7.8)
VIQ	87.8 (10.6)	97.5 (10.4)
PIQ	89.1 (10.8)	91.2 (6.7)

bral electrodes. Exclusion criteria consisted of evidence of right hemisphere language function during the intracarotid sodium amytal evaluation or evidence of frontal lobe seizure discharges. In addition, subjects failing to obtain a score of at least 34/36 on the CF copy trial were excluded so that constructional deficits would not confound examination of memory performance. Details of our diagnostic workup for temporal lobectomy can be found elsewhere (King et al., 1986).

Patients were randomly divided into two groups of right and left TLE patients for the development of the rating scale, one group for the scale development (R=6, L=5) and a larger group for "cross-validation" (R=9, L=9). The standard administration of the CF was performed (Lezak, 1983). Each patient was instructed to copy the CF in its entirety and to make the drawing approximately the same size and shape as the standard. An immediate and 30-minute delayed recall condition were administered following completion of the CF copy.

Delayed recall performances by the initial group were studied for right TLE patients, and those errors suggestive of right temporal lobe dysfunction were noted. We chose to examine the delayed performances since it had been our impression that delayed recall contains more qualitative errors. The left TLE reproductions were then analyzed to determine how frequently the presumed "right temporal lobe errors" were present. Those errors that did not appear to discriminate between groups were discarded. The list of remaining errors was then applied to a second "cross-validation" sample, and is presented in Table 2. Because our qualitative ratings refer to standard CF units, this information is provided in Table 3 (from Lezak, 1983). To avoid confusion with the standard scoring system, qualitative errors are indicated by Roman Numerals.

RESULTS

Using standard quantitative scoring, the right TLE patients obtained a mean score of 15.4 (10.3) and the left TLE group averaged 16.9 (7.7). This difference is not statistically significant. In contrast, comparison of the two groups using the qualitative error rating revealed a statistically significant

TABLE 2
Scoring System of Qualitative Errors

I. *Diamond (#14) on stem.* The stem is defined as either a vertical or horizontal line which connects the diamond to the main figure. This feature is scored regardless of whether the stem is correctly attached to main figure at #13. This is *not* scored if line #16 is extended beyond triangle #13 without the diamond attached.

II. *Misplacement of diamond (#14).* This feature is scored if the diamond (or if #1 is scored, the stem) is not attached to #13 and is incorrectly placed. This error is scored even with minor misplacements near #13 if it appears that the patient did not intend for the two points to meet. However, this is *not* scored if the diamond, although correctly attached, does not hang down in a predominantly vertical direction.

III. *Rotation of horizontal lines (#8).* Rotation is defined as greater than 30° deviation from the horizontal.

IV. *Distortion of overall configuration.* This is scored if any basic shape other than the rectangle (#2) is present as the major shape. This error is also scored if additional shapes roughly the size of one of the rectangle quadrants is added to the rectangle. If individual elements are drawn with no central figure, the error is *not* scored (e.g., circle or diamond by themselves in isolation).

V. *Inversion, misplacement, or distortion of upper right triangle (#9).* This error is scored if a "mirror image" triangle is present, if the triangle is misplaced (e.g., above #2 on the upper left), or if it is distorted (e.g., no intersection of lower points at appropriate vertices).

VI. *Additional horizontal lines (#8).* This error is scored if six or more lines are drawn for item #8.

VII. *Additional parallel lines.* This is scored if parallel lines similar to #8 are repeated elsewhere. This is scored even if the position of #8 is incorrect. There must be three or more lines to qualify as additional parallel lines in the repeated figure. This may consist of two sets of similar lines, reflecting some confusion on the part of the subject by attempting to combine #8 and #12.

VIII. *Misplacement of upper left cross (#1) or lower cross (#17).* This error is scored if the crosses are attached to wrong elements within the design. The most common errors are at the intersection of triangle #13 with the main rectangle (#2). This error is *not* scored if the cross "points" in the wrong direction, or if it forms an extension of the rectangle.

IX. *Major mislocation.* Some examples of this error are the appearance of #12 in any other quadrant other than the lower right, or if the line on which marks #12 are placed (the RR track) is not attached to the center of rectangle #2 (intersection of lines #4, #5, and #3). Do not double score for II and VIII since mislocation is already a component of those errors.

X. *Additional lines for cross (#1 and #17).* This is scored if either cross has more than one hash mark.

XI. *Incorporation of pieces into a larger element.* This is scored when closing off the figure with extra lines (e.g., connecting cross #1 to triangle #9).

effect (Mann–Whitney $U = 12.5$; $p < .01$). Individual patient ratings may be found in Table 4. In order to examine predictive utility, we classified patients as either right or left TLE based upon rating scores. Two or more errors was our criterion for right TLE classification. Using only ratings derived from the initial group, 15/18 (83%) of the patients were correctly classified in the cross-validation sample ($p < .01$, sign test).

Following the blinded categorization of the second temporal lobe group, those patients who were misclassified were examined to determine what characteristics lead to their misassignment. No consistent errors were ob-

TABLE 3
Traditional Scoring Units for the Complex Figure

1. Cross upper left corner, outside of rectangle
2. Large rectangle
3. Diagonal cross
4. Horizontal midline of 2
5. Vertical midline
6. Small rectangle, within 2 to the left
7. Small segment above 6
8. Four parallel lines within 2, upper left
9. Triangle above 2 upper right
10. Small vertical line within 2, below 9
11. Circle with three dots within 2
12. Five parallel lines within 2 crossing 3, lower right
13. Sides of triangle attached to 2 on right
14. Diamond attached to 13
15. Vertical line within triangle 13 parallel to right vertical of 2
16. Horizontal line with 13, continuing 4 to right
17. Cross attached to 5 below 2
18. Square attached to 2, lower left

served. Examination of Table 4 reveals errors VIII and IX to be relatively common in the left temporal group. However, for both these errors, their occurrence is over twice as frequent in the right temporal group suggesting it would be premature to exclude these items from the scale at this time. With a larger series of patients, a weighted scoring system could be developed giving secondary importance to those items that, although still discriminative, are occasionally observed in left TLE patients.

Examples of right and left TLE reproductions with similar quantitative results, but different qualitative ratings, are presented in Figures 1 and 2. Although both reproductions in Figure 1 are in the non-impaired range according to standard quantitative scoring, the presence of qualitative scoring errors including horizontal line rotation, misplacement of upper left cross, and additional hash marks on the lower cross in 1A suggests right

TABLE 4
Error Analysis of Cross-Validation Sample

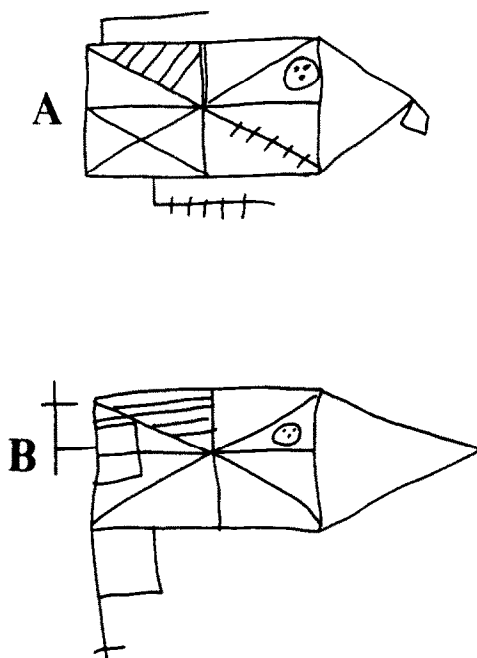| Left Temporal	Right Temporal
1. IX	1. III, VIII, X
2. –	2. I, VI
3. IX	3. II, III, IV, VIII, IX
4. IX	4. IX
5. VIII	5. IV, VIII, IX, XI
6. VIII	6. IV, VIII, IX, XI
7. –	7. II, IV, VI, VII, IX
8. VIII	8. –
9. III, V	9. IV, VIII, IX

FIGURE 1 (A, B). Example of unimpaired recall on quantitative scoring of a right (A) and left (B) temporal lobe patient. Note the presence of scoring errors III, VIII, and X for the right TLE drawing. Both right and left drawing are in the unimpaired range using standard quantitative scoring (18/36 and 21.5/36 respectively).

temporal dysfunction, which corresponds to the patient's seizure onset. A reproduction with similar quantitative scoring, but without qualitative errors and obtained from a left TLE patient, is presented in Figure 1B for comparison purposes. Figure 2 shows impaired design recall using quantitative scoring for a right (2A) and left (2B) TLE patient. Distinct qualitative differences are present in the right TLE reproduction including significant distortion of the overall configuration, misplacement of the upper left cross which is not simply an extension of the figure, major mislocation, and incorporation of pieces into a larger element.

DISCUSSION

The present report illustrates that standard CF memory scoring criteria are inadequate to characterize the types of errors observed in patients with right TLE. As can be seen from Table 2, many of the types of responses that we observe with a right hemisphere seizure focus involve distortion or misplacement. Obviously, a scoring system that scores principally for presence

or absence of elements, with little or secondary weight to misplacement, cannot adequately capture the quality of errors observed in right TLE patients. For example, with perseveration of the horizontal lines to other areas (scoring error VII), the standard scoring system is unable to score this error if the original horizontal lines (#8) are correctly placed. This is in addition to the problem that many "scoring units" appear arbitrarily defined and are not equally susceptible to the effects of brain dysfunction.

The mesial temporal lobe structures, from which partial-complex seizures typically originate, are critical for the acquisition of new material into recent memory. Consequently, TLE patients have the advantage of relatively isolated damage to recent memory structures. In our patients, seizure onset was determined to arise primarily from a single temporal lobe since otherwise, unilateral temporal lobectomy would not have been performed.

The TLE population has the additional advantage that significant language and visual-constructional deficits are generally absent. Seizure onset must primarily originate from a single temporal lobe prior to temporal
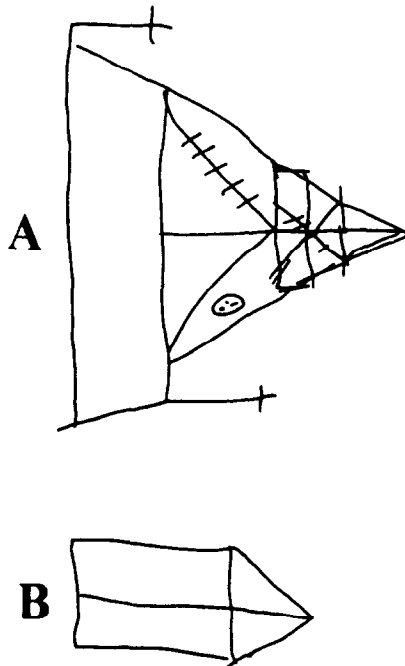

FIGURE 2 (A, B). Example of impaired quantitative scoring of a right (A) and left (B) temporal lobe patient. Note that although both drawings are in the impaired range, drawing 2A contains qualitative errors IV, VIII, IX, and XI. No qualitative errors are seen in the left TLE drawing. Standard quantitative scores for the right and left patients were 9/36 and 8/36 respectively.

lobectomy; however, spontaneous interictal spikes may be observed originating from the contralateral temporal lobe suggesting some degree of bilateral dysfunction. In addition, seizures are occasionally recorded that appear to begin in the contralateral temporal lobe. Many TLE patients also develop seizures relatively early in life and a certain amount of functional reorganization may be expected. Given the above considerations, decreasing the likelihood of finding group differences, and the relatively subtle nature of an epileptic lesion (e.g., mesial temporal sclerosis), we believe our demonstration of right/left TLE group differences suggests a strong probability of similar sensitivity in other neurologic populations.

A caveat is indicated by our decision to restrict the sample to only those patients with scores of at least 34 on the CF copy. We have observed patterns in other populations with predominant left hemisphere injury who perform below 34 on the CF copy and who display similar "right-sided" qualitative errors on delayed recall. Consequently, without adequate CF copy performance, the signs should not be considered as evidence of lateralized temporal dysfunction. However, when the initial copy is performed satisfactorily, the qualitative erros appear as a function of memory processing and transformation. Thus, in the latter case, qualitative errors may reflect lateralized temporal dysfunction.

REFERENCES

Binder, L. M. (1982). Constructional strategies on complex figure drawings after unilateral brain damage. *Journal of Clinical Neuropsychology, 4*, 51–58.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental design for research*. Chicago: Rand McNally.

Delaney, R. C., Rosen, A. J., Mattson, R. H., & Novelly, R. A. (1980). Memory function in focal epilepsy: A comparison of non-surgical, unilateral temporal lobe and frontal lobe samples. *Cortex, 16*, 103–117.

Glowinski, H. (1973). Cognitive deficits in temporal lobe epilepsy. *Journal of Nervous and Mental Disease, 157*, 129–137.

King, D. W., Flanigin, H. F., Gallagher, B. B., So, E. L., Murvin, A. J., Smith, D. B., Oommen, K. J., Feldman, D. S., & Power, J. (1986). Temporal lobectomy for partial complex seizures: Evaluation, results, and 1-year follow-up. *Neurology, 36*, 334–339.

Ladavus, E., Umiltà, C., & Provinciali, L. (1979). Hemisphere-dependent cognitive performances in epileptic patients. *Epilepsia, 20*, 493–502.

Lezak, M. D. (1983). *Neuropsychological assessment 2nd Edition*. New York: Oxford.

Loring, D. W., & Papanicolaou, A. C. (1987). Memory assessment in neuropsychology: Theoretical considerations and practical utility. *Journal of Clinical and Experimental Neuropsychology, 9*, 340–358.

Mayeux, R., Brandt, J., Rosen, J., & Benson, D. F. (1980). Interictal memory and language impairment in temporal lobe epilepsy. *Neurology, 30*, 120–125.

Meehl, P. E. (1975). Nuisance variables and the ex post facto design. In M. Radner & S. Winokar (Eds.), *Analyses of theories and methods of physics and psychology*. Minneapolis, MN: University of Minnesota Press.

Milner, B. (1968). Visual recognition and recall after right temporal lobe excision in man. *Neuropsychologia*, **6**, 191–209.

Milner, B. (1975). Psychological aspects of focal epilepsy and its neurosurgical management. In D. P. Purpura, J. K. Penry, & R. D. Walter (Eds.), *Advances in Neurology* (Vol. 8). New York: Raven Press.

Waber, D. P., & Holmes, J. M. (1986). Assessing children's memory productions of the Rey-Osterrieth complex figure. *Journal of Clinical and Experimental Neuropsychology*, **8**, 563–580.