

RAVLT and Nonverbal Analog: French Forms and Clinical Findings

Viviane Sziklas, Marilyn Jones-Gotman

ABSTRACT: Background: Objective clinical evaluation of memory frequently requires serial testing but the issue of whether multi-formed tests are equivalent and can be used interchangeably is seldom examined. An added problem in bilingual Canadian settings is the extent to which it is appropriate to measure French speakers' performance on translations of English tests. The present work used the Rey Auditory Verbal Learning Test (RAVLT) and a nonverbal analog, the Aggie Figures Learning Test (AFLT), to examine whether a) different forms of the same test are equivalent, b) performance on the two tests is comparable, c) two language groups perform similarly, and d) the RAVLT can detect dysfunction in patients with temporal lobe epilepsy (TLE). **Methods:** We compared three French versions of the RAVLT and three forms of the AFLT in 114 healthy francophone adults. We subsequently compared the performance of the same francophone subjects to a previously obtained sample of anglophones on both tests, and then administered the RAVLT to anglophone or francophone patients with TLE. **Results:** For both tasks the three forms were equivalent and performance on the RAVLT was comparable to that on the AFLT. Francophone subjects performed slightly worse on the RAVLT compared to anglophones but performance of the two language groups did not differ on the AFLT. Finally, left TLE patients were impaired compared to right on the RAVLT, but no performance differences were observed across the two language groups in the patient sample. **Conclusions:** The RAVLT and AFLT are useful tools for examination of learning and memory in French and English speaking populations. On the RAVLT, the lesion effect in patients is not affected by differences in performance between language groups.

RÉSUMÉ: RAVLT et analogue non verbal : évaluations en français et observations cliniques. Contexte : L'évaluation clinique objective de la mémoire requiert souvent des tests sériés. Cependant, on ne sait pas si les tests multiformes sont équivalents et peuvent être utilisés de façon interchangeable. De plus, on ne sait pas s'il est approprié de mesurer la performance d'un francophone au moyen de traductions de textes anglais, ce qui constitue un problème additionnel. Dans cette étude, nous avons utilisé le Rey Auditory Verbal Learning Test (RAVLT) et un analogue non verbal, le Aggie Figures Learning Test (AFLT), pour évaluer si : a) des formes différentes du même test sont équivalentes; b) les résultats sont comparables lorsque deux tests sont administrés; c) les résultats sont les mêmes pour des groupes de langue différente; d) le RAVLT peut détecter une dysfonction chez les patients atteints d'épilepsie du lobe temporal (ÉLT). **Méthodes :** Nous avons comparé trois versions françaises du RAVLT et trois formes de AFLT chez 114 adultes francophones en bonne santé. Nous avons ensuite comparé les résultats des mêmes sujets francophones à ceux d'un échantillon de sujets anglophones pour lequel ces données avaient déjà été obtenues. Nous avons ensuite administré le RAVLT à des patients anglophones ou francophones atteints d'ÉLT. **Résultats :** Pour ces deux tâches, les trois formes étaient équivalentes et les résultats du RAVLT étaient comparables à ceux de l'AFLT. Les sujets francophones avaient des résultats légèrement moins bons au RAVLT par rapport aux anglophones, mais les résultats dans les deux groupes linguistiques au AFLT n'étaient pas différents. Finalement, le RAVLT objectivait une atteinte chez les patients présentant une ÉLT gauche par rapport à ceux qui avaient une ÉLT droite. Cependant, aucune différence n'a été observée selon la langue chez les patients. **Conclusions :** Le RAVLT et l'AFLT sont utiles pour l'examen de l'apprentissage et de la mémoire dans des populations de langue française et de langue anglaise. L'effet de la lésion chez les patients n'est pas influencé par une différence de résultats entre les groupes linguistiques.

Can. J. Neurol. Sci. 2008; 35: 323-330

Clinical evaluation of learning and memory often requires repeated testing over time to document changes in function. Serial administrations of the same test, however, can confound findings, in part, due to practice effects, which can affect the validity of the results obtained. For this reason, it is critical that tests exist in several equivalent forms for retesting. When different versions of tests exist, their equivalence (here, used interchangeably with "matched" and defined as the extent to which several forms of a test yield comparable performance

measures across groups) has not always been examined empirically.¹ In multilingual clinical settings, an added difficulty

From the Montreal Neurological Institute, Department of Psychology, McGill University, Montreal, Quebec, Canada.

RECEIVED FEBRUARY 16, 2007. FINAL REVISIONS SUBMITTED DECEMBER 21, 2007.
Reprint requests to: Viviane Sziklas, Montreal Neurological Institute, 3801 University Street, Montreal, Quebec, H3A 2B4, Canada.

is that the tests used are frequently translations from English originals and have usually not been validated for the population on which they are used. (but see^{2,3}) In Quebec and other parts of Canada with a large French speaking population, the issue of whether French speakers' performance on translations of English tests can be compared with English-based normative data is also a critical one. Direct translations, for example, can result in poorer scores because stimulus items do not occur with the same frequency in different languages. The importance of the latter issue has been raised for both French-⁴ and Spanish-speaking⁵ populations.

Several tests are currently used to assess memory in the French Canadian population. Many centers, including our own, use French translations of well known English-language tests and often apply English normative data. Few studies have compared the performance of these two language populations directly or attempted to create valid French translations. Bishop et al⁶ compared the performance of French and English-speaking children on the Rey Auditory Verbal Learning Test (RAVLT) and showed that French speakers scored more poorly than the English speaking sample. While the authors suggested that differences in IQ likely contributed to their findings, the need for appropriate translations of tests is clear. A recent study⁷ demonstrated validity of a French translation of the California Verbal Learning Test based on a large sample of subjects and Van der Linden and Coyette (unpublished) reported preliminary norms for a French adaptation of the Selective Reminding Test.⁸ To our knowledge, however, no existing memory tests meet the criteria of having several interchangeable French forms, being completely comparable to English versions, and being validated in both verbal and nonverbal domains for a francophone population.

A further issue in memory testing is the existence of very few analogous verbal and nonverbal tests (i.e., tests that evaluate similar processes but in different modalities). Several visual tests, such as the Rey Complex Figure,^{9,10} figures from the Wechsler Memory Scale (The Psychological Corporation), and the Batterie d'efficience mnésique¹¹ are used clinically but have no specific verbal analogs. In addition, the Rey Visual Design Learning Test¹² is sometimes used as a nonverbal analog to the

RAVLT, but even this test is not directly analogous to the RAVLT because it does not allow for evaluation of as many constructs (e.g., effects of interference, recall) as the RAVLT. Further, that test exists in only one form. The Abstract Word List and Abstract Design List tests¹³ exist in three versions and have been used when analogous tests are required for lateralization of dysfunction in temporal lobe epilepsy (TLE). While these tests have been demonstrated to be sensitive to the effects of damage in the medial temporal region, they are based on a recall of 24 hours, a limiting factor when testing cannot be accomplished over consecutive days.

The present study had several aims. First, using a cross-sectional design, we sought to determine whether three French versions of the RAVLT were similar in difficulty. We also examined whether performance on these verbal measures was comparable to performance on a visuo-perceptual test, the Aggie Figures Learning Test (AFLT), which we constructed and reported previously as an analog to the RAVLT.¹⁴ Our second aim was to compare the performance of native French and English speakers on both tasks to determine whether language plays a role in test performance. This would use data from, and extend, our earlier work, in which we reported that three forms for both RAVLT and AFLT were well-matched for an English speaking population, with generally comparable scores on the RAVLT and the AFLT. Finally, as part of an ongoing study, we examined the performance on the RAVLT of English and French speaking patients with temporal lobe epilepsy to evaluate its use as a clinical test for a French speaking population. The AFLT was not included in that battery and its potential usefulness in the francophone TLE patients is currently being examined.

Experiment 1: Performance of francophone subjects on three forms of RAVLT and AFLT

The aim of this experiment was to examine whether three French versions (see Materials) of the RAVLT were similar in difficulty and to examine whether performance on these verbal measures was comparable to performance on a visuo-perceptual learning test, the AFLT, which was constructed as an analog to the RAVLT.

Table 1: . Demographic composition of subjects in Experiment 1

Group	N	Age (years)		Gender		Education (years)	
		Mean	Range	Men	Women	Mean	Range
Form 1	38	22.5	19-32	19	19	15.9	13-23
Form 2	38	22.0	19-29	19	19	15.6	13-23
Form 3	38	22.1	19-30	19	19	15.2	12-22

METHODS

Subjects

One hundred fourteen healthy volunteers participated in the study. Most were undergraduate university students and all were right-handed native French speakers. They ranged in age from 19 to 32 years, with a mean of 15.5 years of education. All were healthy and free of any known neuropathology (by self-report) or history of head trauma. The subjects were randomly assigned to three groups, each of which received one form of the learning tests. The groups were comparable in gender, age, and level of education (Table 1).

MATERIALS

French forms

Three forms of the RAVLT were used. Two were based on the original four lists (15 words each) of Rey.^{9,12} Of these, one had been translated into English and adapted by Lezak¹⁵ and represents the version commonly used in English. We created a second form using the remaining two of Rey's lists. The last one was a French translation of our third English form, which we had devised by translating the Crawford et al¹⁶ list. Each form consisted of a principal and an interference list. In addition, a recognition test, consisting of all words from the principal and interference lists (15 each) and 20 distracters, was created for each form. Of the distracters 15 were similar to the words of the principal list and five resembled words from the interference list; distracters were chosen to be either semantically or phonetically related to target words in the two lists; approximately one half of the words in each list fell into each of the two categories. We equated all corresponding lists (principal, interference and recognition) for word frequency using Brulex¹⁷ and any repeated words were replaced.

Aggie Figures Learning Test (AFLT)

The three forms of the AFLT were described elsewhere¹⁴ and were analogous to the RAVLT. Each form consisted of a principal and an interference list, and a recognition test using abstract figures as stimuli. The AFLT recognition was comparable to the RAVLT: 15 distracter items resembled the figures in the principal list and five resembled figures from the interference list. Sample figures are illustrated in Figure 1.

Two additional tasks were included in the protocol to help fill the interval between learning and delayed recall of the memory tasks. The first was the Boston Naming Test,¹⁸ which requires the subject to name 60 line drawings of objects of increasing difficulty. The second, the Hooper Visual Organization Test (HVOT;¹⁹), requires the mental rearrangement and identification of objects from disarrayed segments. These tests also served as control tasks should any differences exist amongst the groups on the learning tests (see Experiment 2).

Procedure

The tests were administered to all subjects as follows. The RAVLT was given first (20 minutes), followed by the AFLT (25 minutes), after which delayed recall and recognition of the words were obtained. The two distracting tasks and a brief subject history were then administered, followed by delayed recall and

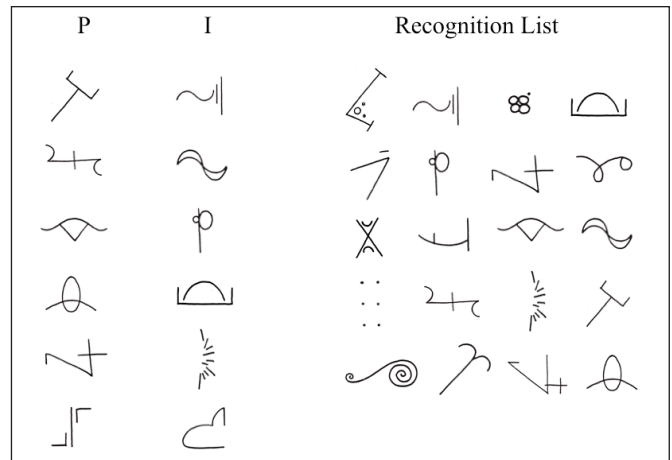


Figure 1: Sample items from the AFLT (P, principal list; I, interference list).

recognition of the figures. The sequence of test administration was the same as that given to Anglophone subjects in our earlier study¹⁴ and thus it was necessary to maintain the protocol for comparative purposes.

The RAVLT was administered according to the protocol described by Lezak.¹⁵ The words were read aloud by the examiner at a rate of one per second and subjects were asked to recall them in any order. Five trials of the principal list were administered first, after which the interference list was read. Subjects were allowed up to two minutes for recall on each trial, with no feedback about accuracy of recall. Immediately following the interference trial, and again after a 30 minute delay interval, subjects were asked to recall the principal list. Finally, subjects were read the recognition list, again without feedback, and asked to state whether the word belonged to the principal or interference lists, or if the word was new (distracter).

Administration of the AFLT differed from that of the RAVLT in that the figures were presented visually and they were exposed at a rate of one every three seconds. Following each presentation of the list, the subjects were asked to draw the figures from memory, in any order and again with a two-minute limit.

Scoring

Each correctly recalled word or figure was allotted one point. Words were accepted if they were identical to the target; differences in pronunciation were accepted as long as the word could be identified as the target. For AFLT, figures that could be unambiguously identified were credited; thus figures that were slightly incomplete (e.g., lines shorter than original drawing, circles not closed, etc.), rotated but correct, or distorted (e.g., curves more concave than target, angles more or less acute than target, etc.) were accepted as long as there was no doubt which figure was represented. These scoring criteria ensured that despite differences in modality, the RAVLT and AFLT were scored for correct responses as comparably as possible. For the recognition test, separate scores were computed for items that

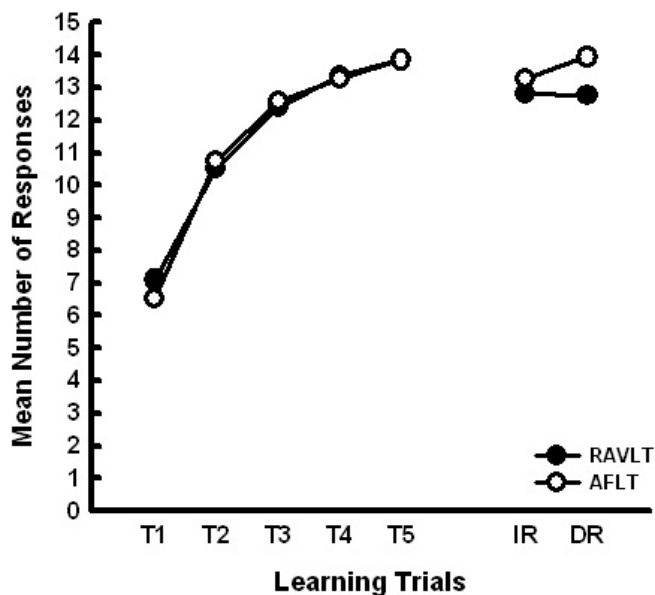


Figure 2: Performance of francophone subjects on the RAVLT and AFLT across five learning trials, immediate (IR) and delayed recall (DR) trials.

had appeared in the principal list and for those that had appeared in the interference list.

RESULTS

A two-way mixed analysis of variance with repeated measures over test and trial revealed significant effects of trial ($F[1,222]=2137.62$, $p=0.000$) and of the test \times trial interaction ($F[1,222]=24.01$, $p=0.000$). Further analysis of this interaction effect (Newman-Keuls) showed that performance on the RAVLT was better than that on the AFLT on trial 1 but worse on trial 6 (immediate recall of principal list, after interference trial) and at delayed recall ($p<0.05$, all tests). No other significant effects were observed, including main effects of test ($F[1,222]=1.269$, $p=0.261$) and form ($F[2,222]=0.186$, $p=0.83$) or of the other interactions, including form \times test ($F[2,222]=0.767$, $p=0.466$) and trial \times form \times test ($F[2,222]=0.724$, $p=0.486$). Figure 2 shows the performance, collapsed across forms, of francophone subjects on RAVLT and AFLT.

The number of items correctly recognized from the principal list during the recognition test for RAVLT and AFLT was compared separately using one-way ANOVAs. There was no significant difference amongst the three test forms for either RAVLT ($F[2,113]=1.22$, $p=0.3$) or AFLT ($F[2,113]=2.21$, $p=0.115$).

Experiment 2: Comparison of francophone and anglophone subjects on RAVLT and AFLT

The aim of Experiment 2 was to determine whether performance across language groups was comparable for the verbal and nonverbal learning tests.

MATERIALS AND METHODS

Subjects

The subjects were the same 114 French speakers tested in Experiment 1. These subjects were compared to our sample of 114 anglophones previously tested on the RAVLT and AFLT. As reported earlier,¹⁴ English speakers ranged in age from 17 to 35 years, with a mean of 14.3 years of education.

The French versions of the RAVLT and AFLT were described in Experiment 1. The three English forms were described previously.¹⁴ Briefly, one was the form based on Rey's original version⁹ as described by Lezak.¹⁵ We created a second form using the two lists from Rey's original four^{9,12} that were not included in what is now commonly used as the RAVLT, and we added a recognition list. A third form was the alternate form published by Crawford et al.¹⁶ The English speaking subjects were also administered the two distracter tasks, the Boston Naming Test and the Hooper Visual Organization Test. The administration of all tests was as described in Experiment 1.

RESULTS

Our earlier results had demonstrated that there was no significant difference between the three English forms of the RAVLT or the AFLT.¹⁴ The learning, immediate and delayed recall trials for the French and English samples were collapsed over forms and submitted to a mixed analysis of variance with repeated measures over trials and tests. This analysis revealed significant effects of language ($F[1,226]=9.6$, $p=0.002$), trials ($F[6,1356]=2125.03$, $p=0.00$), and the test \times trial ($F[6,1356]=19.00$, $p=0.00$) as well as the test \times language ($F[1,226]=3.91$, $p=0.049$) interactions. No other effects were significant. Post-hoc analyses of the test \times language interaction revealed significant differences in performance between francophones and anglophones on the RAVLT ($p<0.05$, all tests) but no difference between the two groups on the AFLT ($p>0.05$). These results are shown separately for each test in Figures 3a and 3b.

Comparisons between language groups for RAVLT ($t(225) = -0.198$, $p=0.844$) and AFLT ($t(226) = -0.143$, $p=0.886$), collapsed over forms, showed no significant differences on recognition performance.

Previously, we showed that amongst anglophone subjects, there were no main effects of gender nor any significant interactions with this variable in repeated measures analyses including tests and learning or memory trials.¹⁴ An analysis of variance with repeated measures over test and trials was carried out to probe gender differences in the francophone sample. The findings revealed significant main effects of gender ($F[1, 226]=8.78$, $p=0.004$) and trials ($F[6,672]=1080.91$, $p=0.00$). Significant interaction effects included gender \times trial ($F[6,672]=3.33$, $p=0.003$), suggesting that the performance between men and women was dependent on particular trials rather than consistently different across trials. The gender \times test \times trials interaction, however, was not significant ($F[6,672]=1.13$, $p=0.34$).

The mean number of correct responses on the control tasks for the two language groups was compared using t tests. The findings showed a significant difference between the groups on the Boston Naming Test ($t(211)=5.4$, $p=0.00$) but not on the Hooper Visual Organization Test ($t(217)=1.8$, $p=0.073$).

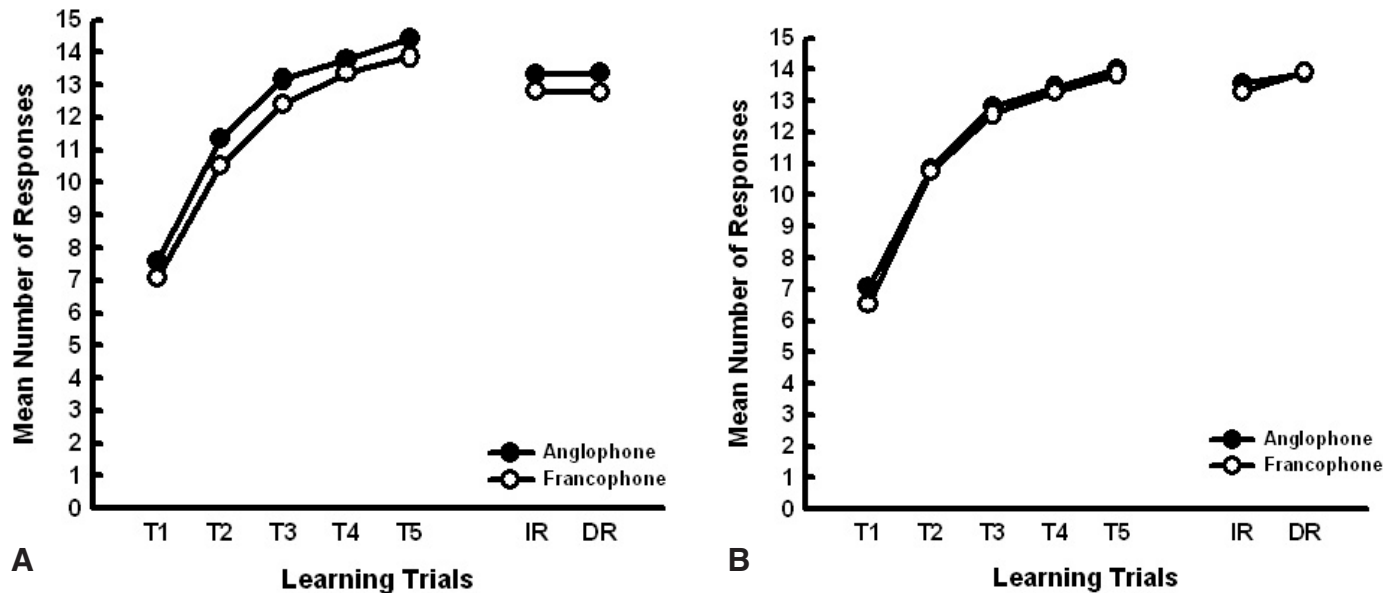


Figure 3: A) Performance of francophone and anglophone subjects on the RAVLT; B) Performance of francophone and anglophone subjects on the AFLT (IR, immediate recall; DR, delayed recall).

Experiment 3: Comparison of francophone and anglophone TLE patients on RAVLT

The RAVLT is commonly used to evaluate learning and memory in patients being considered for surgical treatment of epilepsy. Based on the known verbal learning deficits in patients with resection from the left temporal lobe,^{13,20,21} patients with a left temporal lobe epileptic focus are expected to perform poorly on this test. In Montreal, over half the patients are francophone. The aim of this experiment was to compare the performance of English and French speaking patients with temporal lobe epilepsy to evaluate the RAVLT as a clinical test for a French-speaking population.

SUBJECTS, METHODS AND PROCEDURE

This was a retrospective analysis. Seventy-seven consecutive patients from the Montreal Neurological Hospital who had sustained resection from the anterior temporal region and who had been administered the RAVLT postoperatively were considered in the analysis. Table 2 shows the distribution of gender, age, IQ, and side of surgery for this sample. In most cases, patients were administered version 2 of the RAVLT following operation. The RAVLT, administered as part of the routine clinical test battery, was administered according to the protocol described by Lezak,¹⁵ as summarized in Experiment 1.

RESULTS

A mixed analysis of variance with repeated measures over the five learning and the delayed recall trials was performed for language and side of resection. The results showed significant

main effects of side of resection ($F[1,73]=17.57, p=0.000$) and trials ($F[5,365]=133.49, p=0.00$) but not of language ($F[1,73]=2.58, p=0.112$). A significant interaction was observed for trials \times side ($F[5,365]=13.93, p=0.000$). Further analysis of the interaction effect (Newman-Keuls Test) showed that the performance of the group with left temporal damage was significantly different from that of the right temporal group on trials 2, 3, 4, and 5 as well as on the delayed recall trial ($p<0.05$, all tests). Figure 4 shows the performance of English and French patient groups, collapsed over language, over the learning and delayed recall trials of the RAVLT.

DISCUSSION

This investigation had several aims. First, we sought to determine whether three French forms of the RAVLT could be used interchangeably and whether performance on these tests was comparable to an analogous nonverbal learning measure, the AFLT. We demonstrated equivalence across forms for both tests and showed that performance across modalities was comparable.

Several reports in the literature have constructed or compared alternate versions of the RAVLT in other languages. In English, there are at least seven versions of this test¹ although not all forms have been compared together in any one study. A Dutch adaptation of the RAVLT was developed by Brand and Jolles,³ who created six versions using lists of monosyllabic words. Rey originally created four lists of 15 words^{9,12} but to our knowledge, no study has examined the equivalence of the tests that have been created based on his lists. The study reported in this communication represents the first attempt to compare directly three versions of the RAVLT in a francophone population.

Table 2: . Demographic composition of patient groups in Experiment 3

Side of Surgery	Age (years) (mean \pm SEM)	FSIQ (mean \pm SEM)	Gender		Language		Total
			Men	Women	English	French	
Left	35.9 \pm 1.4	93.2 \pm 1.7	23	21	18	26	44
Right	37.3 \pm 1.9	92.8 \pm 2.0	13	20	12	21	33
Total			36	41	30	47	77

FSIQ=full scale IQ; SEM=standard error of the mean

It is critical to emphasize the need for several forms of a learning test. Repeated testing is often required to document changes over time for a variety of clinical conditions. (e.g.²²) Practice effects (i.e., the gains resulting from re-testing with the same form), however, may confound the data obtained after repeated evaluations. Recently, Rieu et al² constructed six French adaptations of the Hopkins Verbal Learning Test, intended primarily for older populations with memory deficits because of its high tolerability by such patients. The test is brief, consisting of three learning trials although it does have delayed free recall and recognition trials. In view of its reduced demands, however, that test may not be appropriate for younger populations. Thus, the finding in Experiment 1 that the three French versions of the RAVLT can be used interchangeably has important implications for many neurological and other clinical groups in the French Canadian population and fills the need for use in a more intact or younger niche.

The results of Experiment 1 are furthered by the demonstration that performance on the RAVLT is comparable to that observed on the AFLT, a nonverbal companion test.¹⁴ An important drawback of many neuropsychological tests of memory is that the constructs measured in one modality are not comparable to those evaluated in another. The most closely matched pair of verbal and nonverbal tests currently available are the Abstract Word List and Abstract Design List learning tests.^{13,23} These tests evaluate learning and delayed recall by using low-imagery common words and designs that are difficult to verbalize. An added advantage is that stimuli are all presented in the same (visual) modality. In addition to the AFLT, the Rey Visual Design Learning Test (RVDLT¹²) has been used as a nonverbal analog for the RAVLT. The RVDLT, however, is not directly analogous to the RAVLT because it does not allow for evaluation of as many constructs as the RAVLT. (but see²⁴) In addition, one can argue that the stimuli in the RVDLT are highly verbalizable (e.g., combinations of squares, circles, lines, triangles) and as such, do not meet the criteria of a strictly nonverbal learning test. Indeed, there is evidence that performance on the RVDLT can be improved by verbalization strategies.²⁵ Finally, the test exists in only one form. Developed by Majdan et al,¹⁴ the AFLT was designed to meet three criteria required to match the RAVLT: several forms, difficult to verbalize, and measuring the same constructs. Importantly, Experiment 1 has shown that the AFLT yields scores comparable

to the French RAVLT and that it is suitable for use with a francophone population.

Experiment 2 compared the performance of subjects on the RAVLT and the AFLT across language groups. The findings showed that francophone subjects acquired fewer words over the learning trials of the RAVLT but that nonverbal learning (AFLT) was not significantly different between the French- and English-speaking groups. These findings suggest strongly that the poorer performance of the francophone subjects was likely due to difficulty of the items in the French RAVLT rather than any differences between the two language groups. A key issue in the conversion of tests into other languages is that literal translations can often result in changes in word frequency, length, and other linguistic alterations. Many groups have acknowledged this problem in translation of a variety of neuropsychological batteries or tests from English to other languages.^{4,26,27} Differences in word length or duration of the spoken word have

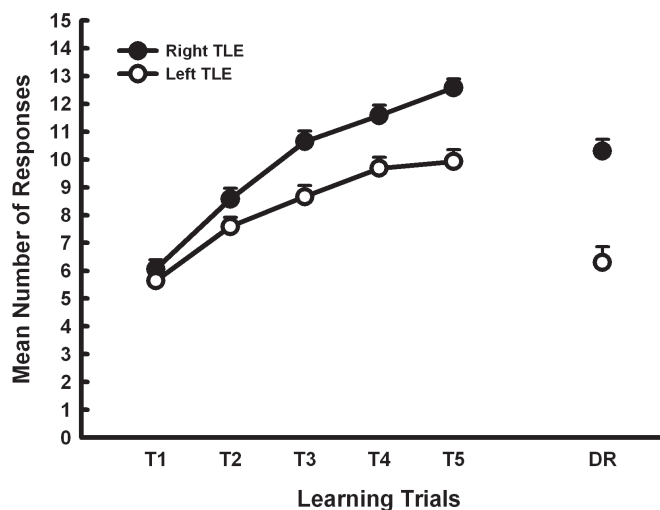


Figure 4: Performance of left and right temporal patients on the RAVLT.

been correlated with reduced performance on a variety of tests (e.g.^{28,29}) and such differences likely account, in part, for the lower scores observed in the French speakers in the present study.

Support for the notion that differences in language may account for the discrepancy between anglophone and francophone performance rather than other factors inherent to the two groups in the present study is provided by the finding that the performance of the two language groups was significantly different on the Boston Naming Test but not on the second control task used in this study, the Hooper Visual Organization Test. To some extent, it may be argued that the HVOT is a “naming” test because subjects are required to identify items verbally and, in fact, it has been suggested that naming ability is modestly associated with HVOT performance.³⁰ Other findings, however, argue that HVOT performance does not depend heavily on language related skills or naming ability.³¹ In clinical practice, we attempt to ensure that we are testing perception and not naming by giving credit if the subject (patient) can provide a clear and unambiguous definition of the object even if he (she) is unable to name it. In the present study, we found a difference between French and English on the Boston Naming Test and RAVLT but no significant difference in performance on HVOT or AFLT. We interpret this to mean that the subjects themselves did not differ in intelligence or other cognitive factors because they did not differ on those two nonlanguage tests. In contrast, the difference between English and French on the two verbal tests suggests that the language factors in these tasks were not perfectly equated in French and English. Further, because there was no main effect of form, as shown in Experiment 1 of the present communication, it is reasonable to suggest that the issue of item difficulty applied equally across the French forms. The finding of poorer performance of francophone subjects on the naming test is consistent with other reports in the literature.³² Further, in a recent study comparing 63 francophone and 63 anglophone subjects’ performance on another language based task, story learning, Djordjevic et al (personal communication) demonstrated that francophones scored poorly in comparison with anglophone counterparts on French versions of this task but did not differ on a test of nonverbal reasoning. Thus, if the HVOT used in the present study was a measure of language, there should have been a difference on it as well but this was not the case. Taken together, such findings argue strongly that the differences in performance between francophone and anglophone populations are based in differences in language rather than other psychological variables. Such findings have important implications for the interpretation of the performance of other language groups against normative data based on standardized English tests.

Several studies have reported the implementation of forward and back translation to minimize the potential biases of direct translation.^{4,27,33,34} This process allows for a better detection of cross-cultural differences in frequency and familiarity of test stimuli. In the case of the present study, only one form required a complete translation from English to French (Form 3). While we did not attempt to carry out forward and back translation when converting that form (i.e., the one based on Crawford, 1989);¹⁶ we can argue that, at least as far as psychometric equivalence across French forms is concerned, this did not

impact performance as all three were demonstrated to be equal.

Despite the possibility that tests translated from English and used in other languages may bias performance unfavourably in the latter population, direct comparisons of performance across language groups are infrequent in the literature. Miranda and Valencia⁵ compared English and Spanish versions of the RAVLT and found no differences in performance in a sample of college students. A potential drawback of their study, however, is that the subjects were bilingual, a factor which likely contributed to the lack of differences in acquisition. To some extent, this issue is unavoidable in regions like Montreal but we made every effort to exclude those subjects who were fluently bilingual from our study. The mean number of items recalled on the RAVLT was worse in the francophone sample but the difference was less than one item for each of the learning and recall trials. Thus, we argue that for clinical purposes, using the French versions of the RAVLT remains a valuable tool.

There is now a large body of evidence demonstrating that patients with damage to the left (speech dominant) medial temporal region are impaired in the learning or retention of verbal material. Deficits have been observed using a variety of verbal stimuli including prose passages³⁵ and word lists.^{13,21,36} The findings of Experiment 3 are consistent with the existing literature, showing a dissociation between the effects of left versus right temporal regions, and continue to suggest that word list learning tests are sensitive to damage in the former. To date, however, there have been no direct attempts to compare the performance of English- and French-speaking patients with temporal lobe epilepsy. This is a particularly important issue for neurological institutions such as those in the Montreal region which cater to a multi-lingual patient population. In view of the slight differences observed between two healthy language groups (Experiment 2, present communication), it was necessary to examine the extent to which performance of patients with damage in the temporal region reflected a true deficit as measured by the RAVLT. As part of an ongoing study, Experiment 3 compared word learning in francophone and anglophone patients who had sustained resection from the right or left temporal region for the relief of epilepsy. The results showed that the performance of francophone and anglophone patient groups was comparable and suggest that the language of testing does not have a significant impact on patient performance. These data were unexpected in view of the finding of Experiment 2, in which we demonstrated a small but significant difference between the two language groups. In contrast, this pattern of findings (i.e., a difference between healthy francophone and anglophone subjects but no differences between patient language groups on the same test) is consistent with those using other cognitive measures. For example, Djordjevic et al³² demonstrated that while the scores on two separate naming tests were significantly lower for healthy francophone subjects than in anglophones, no language effects were observed when English- and French-speaking patients with resection from the left temporal region were compared. Taken together with the findings of Experiment 3, such data suggest that, at least on certain naming and memory measures, the effect of lesion outweighs that of language.

The need for valid measures of cognitive function in several languages is a critical one. Given the prevalence of Spanish in the United States, large efforts have been made to ensure that

neuropsychological tests measure appropriately the constructs they are meant to evaluate. Similar efforts are being made in Canada, in view of the importance of the French-speaking population. The present study has contributed to this endeavour by demonstrating the validity of two learning tests in a young, healthy francophone population and their potential as useful clinical tools in the neuropsychological investigation of French-speaking patients with temporal lobe epilepsy. While our findings provide a suitable gauge for Francophone epilepsy patients close in age and education to the healthy subjects in our study, current efforts are being directed towards extending the data base to match a broader spectrum of French-speaking patients with epilepsy.

ACKNOWLEDGMENTS

Supported by grant MOP 53274 from the Canadian Institutes of Health Research to MJG and VS. We thank Giulia de Prophetis for technical assistance.

REFERENCES

- Hawkins KA, Dean D, Pearlson GD. Alternative forms of the Rey Auditory Verbal Learning Test: a review. *Behav Neurol.* 2004;15(3-4):99-107.
- Rieu D, Bachoud-Levi AC, Laurent A, Jurion E, Dalla Barba G. French adaptation of the Hopkins Verbal Learning Test. *Rev Neurol (Paris).* 2006;162(6-7):721-8.
- Brand N, Jolles J. Learning and retrieval rate of words presented auditorily and visually. *J Gen Psychol.* 1985;112(2):201-10.
- Demers P, Robillard A, Lafleche G, Nash F, Heyman A, Fillenbaum G. Translation of clinical and neuropsychological instruments into French: The CERAD experience. *Age Ageing.* 1994; 23(6):449-51.
- Miranda JP, Valencia RR. English and Spanish versions of a memory test: word-length effects versus spoken-duration effects. *Hisp J Behav Sci.* 1997;19(2):171-81.
- Bishop J, Knights RM, Stoddart C. Rey Auditory-Verbal Learning Test: performance of English and French children aged 5 to 16. *Clin Neuropsychol.* 1990;4:133-40.
- Nolin P. Analyses psychométriques de l'adaptation française du California Verbal Learning Test (CVLT). *Rev Qué Psychol.* 1999;20(1):39-55.
- Grober E, Buschke H. Genuine memory deficits in dementia. *Dev Neuropsychol.* 1987;3:13-36.
- Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique (Les problèmes). *Arch Psychol.* 1941;28:286-340.
- Osterrieth PA. Le test de copie d'un figure complexe: contribution à l'étude de la perception et de la mémoire. *Arch Psychol.* 1944;30:286-356.
- Signoret JL, Whitley A. A memory battery scale. *Int Neuropsychol Soc Bull.* 1979;2-26.
- Rey A. L'examen clinique en psychologie. Paris: Presses Universitaires de France; 1964.
- Jones-Gotman M, Zatorre RJ, Olivier A, Andermann F, Cendes F, Staunton H, et al. Learning and retention of words and designs following excision from medial or lateral temporal-lobe structures. *Neuropsychologia.* 1997; 35(7):963-73.
- Majdan A, Sziklas V, Jones-Gotman M. Performance of healthy subjects and patients with resection from the anterior temporal lobe on matched tests of verbal and visuo-perceptual learning. *J Clin Exp Neuropsychol.* 1996;18(3):416-30.
- Lezak M. Neuropsychological assessment. 2nd ed. New York: Oxford University Press; 1983.
- Crawford JR, Stewart LE, Moore JW. Demonstration of savings on the AVLT and development of a parallel form. *J Clin Exp Neuropsychol.* 1989;16:190-4.
- Content A, Mousty P, Radeau M. Brulex: Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique.* 1990;90:551-66.
- Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
- Hooper HE. Hooper Visual Organization Test. Los Angeles: Western Psychological Service; 1983.
- Corsi P. Human memory and the medial temporal region of the brain. Montreal: Montreal Neurological Institute, McGill University; 1972.
- Helmstaedter C, Elger CE, Hufnagel A, Zentner J, Schramm J. Different effects of a left anterior temporal lobectomy, selective amygdalohippocampectomy, and temporal cortical lesionectomy on verbal learning, memory, and recognition. *J Epilepsy.* 1996; 9:39-45.
- Cunje A, Molloy DW, Standish TI, Lewis DL. Alternate forms of logical memory and verbal fluency tasks for repeated testing in early cognitive changes. *Int Psychogeriatr.* 2007;19(1):65-75.
- Kennepohl S, Jones-Gotman M, McMackin D, Cendes F. Equivalence of three alternate forms of list-learning tasks and indications of task sensitivity to preoperative mesial temporal-lobe atrophy. *Cerebral Bases of Behaviour: A Symposium in Honour of Brenda Milner;* 1993; Montreal, Canada; 1993.
- Senior G, Tehan G. Comparability of performances on the Rey Auditory Verbal and Visual Design Learning Tests. Annual Conference of the Australian Psychological Society; 1993 September 29 - October 2, 1993; Gold Coast, Australia; 1993.
- Wilhelm PW. Reliability and validity of the Rey Visual Design Learning Test in primary school children. *J Clin Exp Neuropsychol.* 2004;26(7):981-94.
- Colombo-Thuillard F, Assal G. Boston Naming Test: French-language adaptation and short forms. *Revue européenne de Psychologie Appliquée.* 1992;42(1):72-3.
- Wilkie FL. HUMANS: an English and Spanish neuropsychological test battery for assessing HIV-1-infected individuals -- initial report. *App Neuropsychol.* 2004;11(3):121-33.
- Ellis NC, Hennesly RA. A bilingual word-length effect: implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *Br J Psychol.* 1980; 71(pt 1):43-51.
- Valencia RR, Rankin RJ. Evidence of content bias on the McCarthy Scales with Mexican American children: implications for test translation and nonbiased assessment. *J Ed Psychol.* 1985; 77(2):197-207.
- Jefferson A, Wong S, Bolen E, Ozonoff A, Green RC, Stern RA. Cognitive correlates of HVOT performance differ between individuals with mild cognitive impairment and normal controls. *Arch Clin Neuropsychol.* 2006; 21, 405-12.
- Merten T. Factor structure of the Hooper Visual Organization Test: a cross-cultural replication and extension. *Arch Clin Neuropsychol.* 2005;20(1):123-8.
- Djordjevic J, Sziklas V, Piper D, Pénicaud S, Jones-Gotman M. Auditory Naming Test: a 60-item version in English and French. Annual Meeting of the American Epilepsy Society; Washington, D.C.; 2005.
- Gonzalez T, Balsa A, Sainz de Murieta J, Zamorano E, Gonzalez I, Martin-Mola E. Spanish version of the Northwick Park Neck Pain Questionnaire: reliability and validity. *Clin Exp Rheumatol.* 2001;19(1):41-6.
- Rouget-Weber B. Screening for bipolar disorders using a French version of the Mood Disorder Questionnaire (MDQ). *J Affect Disord.* 2005;88(1):103-8.
- Frisk V, Milner B. The role of the left hippocampal region in the acquisition and retention of story content. *Neuropsychologia.* 1990;28(4):349-59.
- Helmstaedter C, Grunwald T, Lehnertz K, Gleissner U, Elger CE. Differential involvement of left temporolateral and temporomesial structures in verbal declarative learning and memory: evidence from temporal lobe epilepsy. *Brain Cogn.* 1997;35:110-31.