Normative data and construct validity of the Rey Auditory Verbal Learning Test in a Brazilian elderly population
Normative data and construct validity of the Rey Auditory Verbal Learning Test in a Brazilian elderly population

Helenice Charchat Fichman1,2,3, Luciana Brooking Teresa Dias1,2, Conceição Santos Fernandes2, Roberto Lourenço2, Paulo Caramelli3,4 and Ricardo Nitrini3

1 Pontifícia Universidade Católica do Rio de Janeiro, Brazil
2 Universidade do Estado do Rio de Janeiro, Brazil
3 Universidade de São Paulo, Brazil
4 Universidade Federal de Minas Gerais, Brazil

Abstract

The Rey Auditory Verbal Learning Test (RA VLT) is an episodic memory test that is influenced by demographic variables, such as age, education, and gender. Classic word list learning has also been well documented to be an experimental cognitive paradigm that is very useful for the detection of Alzheimer’s disease. The present study analyzed the performance of healthy elderly subjects on the RA VLT, describing the effects of age, education, and gender. To verify its construct validity, the RA VLT was compared with delayed recall and recognition scores on the Brief Cognitive Screening Battery (DR-BCSB). The sample comprised 183 cognitively healthy elderly subjects from outpatient care units of two university reference centers in Rio de Janeiro and São Paulo, Brazil. All participants were subjected to physical and neurological evaluation and neuropsychological assessment. Pearson’s correlation and Student’s $t$-test were used to investigate the influence of demographic variables on RA VLT performance. To verify convergent-type construct validity, a correlation between RA VLT and DR-BCSB scores was calculated. Significant correlations were found between age and $\sum A_1-A_5 (r = -0.383, p < 0.01)$, $A_5-A_1 (r = -0.054, p < 0.01)$, $A_7 (r = -0.361, p < 0.01)$, and $\text{REC A} (r = -0.288, p < 0.01)$. Educational level correlated with $\sum A_1-A_5 (r = 0.405, p < 0.01)$, $A_5-A_1 (r = 0.184, p < 0.01)$, $A_7 (r = 0.334, p < 0.01)$, and $\text{REC A} (r = 0.329, p < 0.01)$. No correlation was found between gender and RA VLT performance. A significant correlation was also found between RA VLT and DR-BCSB performance ($r = 0.5, p < 0.01$). These results revealed some normative data and convergent-type construct validity of the RA VLT for a Brazilian elderly population. Keywords: RA VLT, demographic variables, construct validity, normative data.

Received 6 November 2009; received in revised form 19 March 2010; accepted 19 March 2010. Available on line 26 June 2010

Introduction

Memory assessment is essential to neuropsychological evaluation in elderly populations with cognitive dysfunction. Memory decline is a marker of mild cognitive impairment (Mattos et al., 2003; Meyer, Xu, Thornby, Chowdhury, & Quach, 2002; Petersen et al., 2001) and is an early predictor of dementia, especially Alzheimer’s disease (AD) (Almeida, 1998). In addition to the importance of diagnostic issues, demographic variables influence memory performance (Van Der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005). Episodic memory is the long-term memory of facts and
events in our life, delimited by time and space (Sternberg, 2000). The scientific literature contains several tests that assess episodic memory, such as the delayed recall test from the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) neuropsychological battery (DR-CERAD) (Bertolucci et al., 2001; Morris et al., 1989), California Verbal Learning Test (CVLT) (Elderkin-Thompson, Mintz, Haroon, Lavretsky, & Kumar, 2006), Visual Reproduction Test (VRT), Free and Cued Selective Reminding Test (FCSRT) (Buschke & Fuld, 1974; Minett, da Silva, Ortiz, & Bertolucci, 2007), Episodic Memory Test (Jorm et al., 1997), Wechsler Memory Scale (WMS-R) (Janowsky, Carper, & Kaye, 1996; Wechsler, 1987), and the delayed recall test from the Brief Cognitive Screening Battery (DR-BCSB) (Nitrini et al., 2004; Takada et al., 2006).

The present study focused on the Rey Auditory Verbal Learning Test (RAVLT) (Spreen & Strauss, 1998), which is a widely used task to assess episodic memory (Mattox et al., 2003). The RAVLT was developed by André Rey from the Edouard Claparède list (Lezak, Howieson, & Loring, 2004) and is an easily administered (Knight, McMahon, Green, & Skeaff, 2006) episodic learning memory test (Knight, McMahon, Skeaff, & Green, 2007; Malloy-Diniz, Lasmar, Gazinelli, Fuentes, & Salgado, 2007; Minett et al., 2007; Spreen & Strauss, 1998).

The RAVLT presents some international (Lezak, 1995) and only one Brazilian normative date for the elderly population (Malloy-Diniz et al., 2007). Studies have emphasized its psychometric properties, suggesting a correlation between the internal content of the RAVLT and that of other memory tests (Spreen & Strauss, 1998). Some authors have reported the reliability of the test (ranging from 0.60 to 0.77) from the test-retest method after a period of time and use of parallel forms. Van Den Burg & Kingma (1999), for example, noted that the most reliable measure was the total number of words, in addition to showing the contribution of the effect of age on the results. The greater reliability of the total number of recalled words was also reported by Geffen, Butterworth and Geffen (1994) in a sample of normal 20- to 67-year-old adults. Strauss, Sherrman and Spreen (2006) reported levels of temporal stability for the RAVLT. In a recent study by Knight et al. (2007), significant results concerning test-retest reliability in healthy, elderly, 65-year-old subjects were found.

These results suggest some important findings regarding the reliability of this test, despite a few studies showing temporal instability and a lack of construct validity for the RAVLT. Most studies have used the RAVLT for diagnostic purposes (Schoenberg et al., 2006). In qualitative studies, people with frontotemporal dementia (FTD) have more intrusion words in RAVLT recall than AD and Parkinson’s disease (PD) patients. Moreover, more words are recalled in PD compared with AD and FTD.

Moreover, the RAVLT has been previously shown to help differentiate patients with dementia with Lewy bodies from those with AD (Ferman et al., 2006) and assess memory function in people with mild cognitive impairment with vascular risk factors (Siuda, Gorzkowska, Opala, & Ochudlo, 2007).

Similar to the majority of cognitive and memory tasks, the RAVLT is influenced by demographic variables, such as age (Malloy-Diniz et al., 2007; Messinis, Tsakona, Malefaki, & Papathanasopoulos, 2007; Lezak et al., 2004; Schoenberg et al., 2006; Spreen & Strauss, 1998; Van Der Elst et al., 2005) and education (Foss, Vale, & Speciali, 2005; Lezak et al., 2004; Malloy-Diniz et al., 2007; Messinis et al., 2007; Schoenberg et al., 2006; Spreen & Strauss, 1998; Van Der Elst et al., 2005). Malloy-Diniz et al. (2007) obtained data on education, age, and gender in an elderly population sample. Education positively correlated with performance on the RAVLT, whereas age had a negative correlation. With regard to gender, they found that women had better performance than men. However, some studies are controversial about the influence of gender (Knight, et al., 2006; Messinis et al., 2007; Van Der Elst et al., 2005). In another study, Malloy-Diniz, Cruz, Torres and Cosenza (2000) applied the test in a sample ranging in age from 16 to 93 years and showed differences in mean and standard deviations for adolescents, adults, and the elderly, demonstrating not only differences among the elderly but also between different age groups.

As mentioned above, the RAVLT is a well documented and accepted memory test. Furthermore, word list learning has been a classic paradigm of memory assessment since the Hermann Ebbinghaus experimental studies about forgetfulness in 1885. However, only one normative study of the RAVLT has been conducted in a Brazilian elderly sample (Malloy-Diniz et al., 2007), and no study has assessed the construct validity of this test within this setting. To achieve this goal, a comparison with other memory tests is required. For the present study, we chose the memory test from the BCSB because it has been previously validated in Brazil and is not influenced by education (Nitrini et al., 2004).

The goal of the present study was to analyze normative data of the RAVLT in a sample of Brazilian healthy elderly subjects and investigate the effects of age, education, and gender on test performance. The study also verified the construct validity of the RAVLT by comparing it with DR-BCSB performance.

Methods

Participants

The sample was composed of 183 normal elderly (65 female, 118 male) from outpatient care units of two university reference centers in Rio de Janeiro and
São Paulo, Brazil (150 subjects from the Behavioral and Cognitive Neurology Unit, University of São Paulo; 33 subjects from the Geriatrics Unit “Mario A. Sayeg,” Policlinica Piquet Carneiro, State University of Rio de Janeiro). All subjects were normal elderly without a dementia diagnosis and free of neurological and psychiatric drugs. The primary sociodemographic characteristics and global cognitive performance of the participants are depicted in Table 1.

**Instruments**

The RAVLT is divided into two lists (A and B), each with 15 unrelated nouns. List A is first presented, followed by free recall. This procedure is repeated five times (A1, A2, A3, A4, A5). List B (interference) is then presented, followed by free recall (B). Afterward, the subject is asked to recall List A (A6). The next step is a 30 minute delayed recall task (A7). The last part of the test is the recognition task, in which 50 words are presented (15 List A, REC A; 15 List B, REC B; plus 20 others). In the latter task, each word list must be followed by the subject indicating whether the word was in List A or List B or in either of the lists (Malloy-Diniz et al., 2007). The following scores were calculated: total number of correct words across trials (∑A1-A5), learning (A5 minus A1 score), proactive interference (B1/A1), retroactive interference (A6/A5), and forgetting speed (A7/A6) (Lezak et al., 2004; Spreen & Strauss, 1998).

The BCSB, 10 line drawings of common objects are presented to the subject, who is asked to name them. The first step is incidental free recall. In the second step, the paper is shown again for 30 seconds two more times, and the subject is asked to memorize and then recall (immediate memory and learning tests). Finally, 5 minutes later, the individual is asked to recall as many figures as possible, followed by the recognition task, in which the 10 pictures previously seen are presented together with 10 other figures not previously seen (Takada et al., 2006).

**Procedure**

All participants were subjected to physical and neurological evaluation and to comprehensive neuropsychological assessment. Evaluation was conducted in a quiet room, one subject at a time.

The participants demonstrated preservation of global cognitive functioning in the Mini-Mental State Examination (MMSE) (Brucki, Nitrini, Caramelli, Bertolucci, & Okamoto, 2003; Folstein, Folstein, & McHugh, 1975) and independence in daily living activities determined by the Pfeffer Functional Activities Questionnaire (Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982) or Lawton Scale (Lawton & Brody, 1969) and did not meet the criteria for dementia according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 1994).

The study was approved by the Ethics Committee from both university centers, and all participants provided written informed consent.

**Statistical analysis**

RAVLT scores were composed of total (∑A1-A5), learning (A5-A1), 30 minute delayed recall (A7), and recognition List A (REC A). A descriptive analysis was performed for age, education, gender, MMSE, and RAVLT scores for all individuals. Pearson’s

<table>
<thead>
<tr>
<th>Table 1. Main characteristics of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td>≤ 75 years</td>
</tr>
<tr>
<td>&gt; 75 years</td>
</tr>
<tr>
<td><strong>Education</strong></td>
</tr>
<tr>
<td>&lt; 8 years</td>
</tr>
<tr>
<td>≥ 8 years</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
</tbody>
</table>

MMSE, Mini-Mental State Examination; **p < .01, Student’s t-test.
correlation coefficients were calculated between age, education, and RAVLT scores. Student’s *t*-test was used to analyze the effects of age, education, and gender on RAVLT performance. Correlation analysis between the delayed recall and recognition tasks from the RAVLT and BCSB were performed to verify convergent-type construct validity. Data were analyzed using SPSS version 16.0 software. The significance level was set at *p < .01.*

**Results**

Effects of age, education, and gender

According to the Pearson analysis, significant negative correlations were found between age and the following scores: ∑A1-A5 (*r* = -.383, *p < .01), A5-A1 (*r* = -.054, *p < .01), A7 (*r* = -.361, *p < .01), and REC A (*r* = -.288, *p < .01). Significant positive correlations emerged between education and the following scores: ∑A1-A5 (*r* = .405, *p < .01), A5-A1 (*r* = .184, *p < .01), A7 (*r* = .334, *p < .01), and REC A (*r* = .329, *p < .01).

No significant difference was found between gender and the following scores: ∑A1-A5 (*t* = 1.9, *p > .05), A5-A1 (*t* = .2, *p > .05), A7 (*t* = 1.5, *p > .05), and REC A (*t* = .7, *p > .05). The results of the descriptive analysis, means, standard deviations, and *t*-test are shown in Table 2.

Construct validity

Positive correlations were found between A7-RAVLT and DR-BCSB (*r* = .528, *p < .01) and between REC A-RAVLT and REC-BCSB (*r* = .197, *p < .01).

**Discussion**

In the RAVLT, processes such as codification, storage, and recall (Atkinson, 1995) are assessed by a learning curve (A1-A5) and recognition and delayed recall (A7-RAVLT) aspects. In the present study, RAVLT performance in healthy elderly subjects was associated with age and education variables, but not with gender. RAVLT performance improved with increasing education and declined with increasing age.

In healthy elderly subjects, RAVLT performance has been previously shown to be affected by age, an effect attributable to functional and anatomical changes (Brockway et al., 1998; Janowsky et al., 1996). However, Janowsky et al. (1996) showed that this decline is not linear with age and that it predominates within the verbal memory domain. Moreover, aging promotes deterioration of other cognitive domains, such as executive function (Lin, Chan, Zheng, Yang, & Wang, 2007; Scuteri, Palmieri, Lonoce, & Giampaoli, 2005), which can be associated with learning performance possibly through its influence with search strategies (Ratcliff et al., 1998). Another hypothesis is that highly educated elderly people have more possibilities to exercise this function because they are more exposed to intellectual stimuli than minimally educated people.

A few studies have shown a gender effect on RAVLT performance, with women presenting higher scores than men (Knight et al., 2006; Malloy-Diniz et al., 2007; Messinis et al., 2007; Van Der Elst et al., 2005). Despite these results, however, we did not find significant differences in performance between genders, consistent with a study by Foss et al. (2005).

**Table 2. Mean (SD) of age, education, and gender groups**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>∑A1-A5</th>
<th>A5-A1</th>
<th>A7</th>
<th>Rec A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 75 years</td>
<td>120</td>
<td>37.1 (11.5)**</td>
<td>4.6 (2.6)</td>
<td>5.7 (4.0)**</td>
<td>10.2 (4.7)</td>
</tr>
<tr>
<td>&gt; 75 years</td>
<td>63</td>
<td>30.7 (8.3)</td>
<td>4.2 (2.6)</td>
<td>4.2 (3.6)</td>
<td>8.1 (4.5)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 8 years</td>
<td>88</td>
<td>31.6 (9.8)**</td>
<td>4.1 (2.4)</td>
<td>4.1 (3.5)**</td>
<td>8.5 (4.8)</td>
</tr>
<tr>
<td>≥ 8 years</td>
<td>89</td>
<td>38.8 (10.6)</td>
<td>4.9 (2.5)</td>
<td>6.2 (3.9)</td>
<td>10.8 (4.2)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>118</td>
<td>36.01 (10.63)</td>
<td>4.51 (2.58)</td>
<td>5.34 (3.98)</td>
<td>9.71 (4.65)</td>
</tr>
<tr>
<td>Male</td>
<td>65</td>
<td>32.81 (11.08)</td>
<td>4.43 (2.26)</td>
<td>4.45 (3.51)</td>
<td>9.16 (4.88)</td>
</tr>
</tbody>
</table>

**p < .01.**
Few validation studies on these tests have been conducted. One such report was a case-control study by Takada et al. (2006), in which the accuracy of two dementia diagnostic tests (DR-BCSB and DR-CERAD) was compared. The DR-BCSB showed higher accuracy than the DR-CERAD among illiterate subjects, similar accuracy among literate subjects, and a statistical trend toward higher accuracy in the entire population.

We also found a positive correlation between performance on the RAVLT and BCSB on both the delayed recall and recognition tasks. The two tests evaluate memory function, but they have some differences. The RAVLT is a verbal episodic memory test and involves auditory-verbal coding, which requires more elaborate encoding and search information strategies than the BCSB. People use a hearing input, and semantic clues are not available in the task; such associations must be established by the subject. The BCSB test, in contrast, has a visual input, and semantic encoding is possible. The DR-BCSB is also a task that shows less educational and age influences than the CERAD (Nitrini et al., 2004). For these reasons, the correlation observed in the present study to investigate construct validity was significant, although the level of significance was not robust.

The present findings represent an important contribution for clinicians because it expands the data on the effects of education on RAVLT performance, demonstrating an association between episodic memory and sociodemographic variables. This feature is particularly relevant in a country like Brazil, which has wide variability in education levels. Moreover, the confirmation of construct validity of the RAVLT in our setting may also be important to neuropsychologists and cognitive neurologists.

Further research is necessary to investigate the effects of age and education in greater detail by considering different groups according to these variables. Other aspects of the test can also be evaluated, such as different trials. Additional studies can verify the effects of education and age on RAVLT performance in people with dementia and determine the test’s clinical validity by comparing healthy elderly subjects with elderly subjects with dementia.

References


Knight, R.G., McMahan, J., Skeaff, C.M., & Green, T.J. (2007). Reliable change index scores for persons over the age of 65 tested on alternate forms of the Rey AVL. Archives of Clinical Neuropsychology, 22, 513-518.


