



Evidence of criterion validity for the Benton Visual Retention Test: comparison between older adults with and without a possible diagnosis of Alzheimer's disease

Adriana M. Zanini¹, Gabriela P. Wagner^{2,3}, Maxciel Zortea¹, Joice D. Segabinazi¹, Jerusa F. Salles¹, Denise R. Bandeira¹, and Clarissa M. Trentini¹

1. Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

2. Federal University of Health Sciences of Porto Alegre, Porto Alegre, RS, Brazil

3. FADERS- Laureate International Universities, Porto Alegre, RS, Brazil

Abstract

The present study sought to provide evidence of criterion validity for the Benton Visual Retention Test by making comparisons between older adults with and without a possible diagnosis of Alzheimer's disease. The control group was composed of 50 older adults, and the clinical group was composed by 16 subjects. Descriptive and inferential statistics were performed, including χ^2 test, F and Wald statistics, t -test, analyses of covariance with $\alpha = .05$, and effect size calculations. We used a sociodemographic data form, the Geriatric Depression Scale-15, and Mini Mental State Examination. Despite the small clinical sample size, the results pointed to evidence of validity for the Benton Visual Retention Test for Administration A (Memory) and Administration C (Copy). The clinical group had significantly poorer performance on most scores. These results also indicate important deficits in other neuropsychological functions in Alzheimer's disease. **Keywords:** Benton Visual Retention Test (BVRT), memory, visuoconstructive abilities, criterion validity, Alzheimer's disease.

Received 12 June 2013; received in revised form 30 January 2014; accepted 05 February 2014. Available online 20 June 2014.

Introduction

Neuropsychological instruments must undergo the same process of validation and adaptation as psychological instruments. Authors consider this important, especially when an instrument is used in another country, because of differences in culture that can generate differences in scores (Hogan, 2006;

Lezak, Howieson, & Loring, 2004). The relative shortage of neuropsychological instruments adapted to Brazil makes research in this area extremely necessary. Because of this, several studies have been conducted to test the validity of the Benton Visual Retention Test (BVRT) in Brazilian samples (Benton Sivan, 1992; Salles, Bandeira, Trentini, Segabinazi, & Hutz, 2014).

Adriana Mokwa Zanini, Maxciel Zortea, Joice Dickel Segabinazi, Jerusa Fumagalli de Salles, Denise Ruschel Bandeira, and Clarissa Marceli Trentini, Psychology Institute, Federal University of Rio Grande do Sul. Adriana Mokwa Zanini and Clarissa Marceli Trentini, Psychological Evaluation and Psychopathology Study Group (NEAPP/Federal University of Rio Grande do Sul). Maxciel Zortea and Jerusa Fumagalli de Salles, Group for Studies in Cognitive Neuropsychology (NEUROCOG/Federal University of Rio Grande do Sul). Joice Dickel Segabinazi and Denise Ruschel Bandeira, Group of Study, Application and Research in Psychological Assessment (GEAPAP/Federal University of Rio Grande do Sul). Gabriela Peretti Wagner, Psychology Department, Federal University of Health Sciences of Porto Alegre and FADERS Laureate International Universities. Correspondence regarding this article should be directed to: Adriana Mokwa Zanini, Instituto de Psicologia - Universidade Federal do Rio Grande do Sul - Rua Ramiro Barcelos, 2600 - Bairro Santa Cecília - Porto Alegre - RS - Brasil - CEP: 90035-003. E-mail: adrizanini@gmail.com

The original version of the BVRT (Benton Sivan, 1992) is composed of three Forms (C, D, and E) that can be administered in four different ways that were standardized by the author (Administrations A, B, C, and D). All of the Forms and Administrations have normative data, including studies with children, adults, old adults, and clinical samples with a history or evidence of neurological damage from the United States. According to Lezak et al. (2004), the BVRT has the sensitivity to detect neurological diseases, thus helping to identify cases of visual agnosia, heminegligence, and visual memory and praxic deficits. Additionally, the instrument has been used internationally to detect and monitor neurodegenerative diseases such as Alzheimer's disease (AD; Robinson-Whelen, 1992).

According to Urbina (2004), validity evidence for the scores of a test can be obtained through any systematic research that confirms or adds to the evidence, regardless of when the research occurs or who performs it. Several

sources of validity evidence are necessary to consider an instrument as valid. Among the different aspects of test validity is criterion validity, which consists of the degree of effectiveness of a test to predict a subject's specific performance (Pasquali, 2001).

Previous Brazilian standardization (Salles et al., 2014) used Forms C and D through Administrations A (Memory) and C (Copy), which are detailed in the methodological description of the BVRT. These combinations of Administrations and Forms were chosen by considering that they were the most often used in research worldwide (Benton Sivan, 1992).

To find evidence of validity, knowing the cognitive functions that the instrument evaluates is essential. According to neuropsychology test manuals, execution of the BVRT requires good visual memory function, visuoconstructive ability, and visual perception (Burin, Drake, & Harris, 2007; Strauss, Sherman, & Spreen, 2006). Visual memory can be characterized within the multicomponential models of working memory of Baddeley and Hitch (1974) and Baddeley, Allen, and Hitch (2011, for an updated version). This model assumes that verbal information is processed in a component (i.e., the phonological loop) separately from the component that processes visuospatial information (i.e., the visuospatial sketchpad). Both components assume short-term storage. Accordingly, one can infer that Administration A (Memory) of the BVRT more specifically assesses the visuospatial loop through visual input. However, the phonological loop and long-term memory information can be recruited once the stimulus can be named and some items can be easily recognized, such as circles and squares.

Most tasks that assess visual memory also involve the use of visuoconstructive skills or constructive praxias, which form expressive functions (Lezak et al., 2004) and are related to the capacity to join parts or stimuli in an organized way. To execute visuoconstructive activities such as copying a drawing, the individual also needs to have preserved visuoceptive function because the process depends on visual information that reaches peripheral visual organs (Zuccolo, Rzezak, & Góis, 2010). In Administration A (Memory) and Administration C (Copy) of the BVRT, respondents use their visuoconstructive and visuoceptive skills. However, in Administration C (Copy), these functions are isolated, allowing the evaluator to observe distinctions between performance in both administrations. Therefore, memory impairment may be reflected by worse performance on Administration A (Memory) but better performance on Administration C (Copy; Caramazza, & Coltheart, 2006).

A recent study by Lockwood, Mansoor, Homer-Smith, and Moses (2011) explored the dimensionalization of the BVRT using factor analysis. The experiments revealed a four-component model that explained 81.04% of the shared variance among three different forms of the BVRT and the Wechsler Adult Intelligence Scale-

Revised (WAIS-R). The components indicated that the moderately difficult items from the BVRT loaded with Perceptual Organization on the WAIS-R. The easiest items of the BVRT loaded separately with both Verbal Comprehension and Freedom from Distractibility on the WAIS-R. The most difficult items from the BVRT loaded weakly with Perceptual Organization on the WAIS-R. These results indicated the importance of other basic neuropsychological and intellectual variables for BVRT performance such as attentional capacity, perceptual organization, and general verbal ability.

Researchers also pointed out the influence of age (Coman, Moses, Kraemer, Friedman, Benton, & Yasavage, 1999) and level of education (Carret, Rainville, Lechevallier, Lafont, Letenneur, & Fabrigoule, 2003; Seo et al., 2007) on BVRT scores, considering this information is important when assessing healthy or clinical samples. In addition to changes in cognitive function in healthy ageing, the emergence of diseases such as dementia also characterizes this age group. McKhann et al. (2011) and McKhann, Drachman, Folstein, Katzman, Price, and Stadlan (1984) defined dementia as the decline of memory and other cognitive functions compared with the individual's previous functioning.

The most common cause of dementia is Alzheimer's disease (AD), with symptom-onset usually after around 65 years of age (late-onset AD). Alzheimer's disease is a degenerative neurological disorder that is clinically manifested as damage to cognitive function, attention, and visuospatial skills (American Psychiatry Association, 2002). These cognitive deficits and decline in the performance of daily activities are the focus of the disease's clinical evaluation (Salmon & Bondi, 2008). The final diagnosis of AD is made based on histological evidence, which is obtained only through biopsy or autopsy, as well as behavioral changes (McKhann et al., 1984, 2011). For this reason, "possible AD dementia" is the term recommended for academic practice. Accordingly, the individual presents diagnostic criteria for AD dementia, but with an evolution pattern that is different from a slow, progressive onset, evidence of other etiologies, or a lack of data about the disease's evolution (McKhann et al., 2011; Brucki & Schultz, 2011).

The BVRT has shown favorable results with regard to criterion validity in the international literature, especially in relation to the differentiation of groups of healthy older adults and patients with dementia (Eslinger, Damasio, Benton, & Van Allen, 1985). Additionally, there are indications of the capacity of this instrument to detect cognitive alterations related to the performance of immediate memory in patients who develop AD years later (Zonderman, Giamba, Arenberg, Resnick, Costa, & Kawas, 1995).

Kawas et al. (2003) assessed 1,425 older adults (> 60 years of age) using the BVRT and Vocabulary subtest of the WAIS. This longitudinal study design

showed that the increase in errors on the BVRT was associated with an increased risk of developing AD up to 14 years after the assessment. This finding suggests that visual memory is affected in the earlier years of dementia. Therefore, an assessment of visual memory might be relevant to the early diagnosis of AD, thus reducing its functional impact.

Resnick, Troutman, Kawas, and Zonderman (1995) published a study that considered specific types of errors in performing the BVRT. The authors examined changes in these scores related to aging. Comparisons between groups indicated that cerebral aging impacts all categories of errors, and intraindividual comparisons showed increases in distortion, omission, and rotation. Therefore, changes in these specific errors may be predictive of brain changes linked to cerebral aging.

Robinson-Whelen (1992) reported a study with no psychometric focus. Instead, this study had the main aim of comparing BVRT scores between AD (very mild, mild, and moderate stages) and control groups. The results showed that participants with a diagnosis of AD had poorer performance compared with healthy older adults. Several deficits in visual memory were detected in the very early stages of dementia, whereas other visuoconstructive abilities remained preserved, even in the later stages of the disease.

Evidence of the validity of the BVRT can also be found in studies that focused on demonstrating how test scores were related to scores on other instruments that supposedly assess the same constructs (Golski, Zonderman, Malamut, & Resnick, 1998; Messinis, Tsakona, Mlefaki, & Papatanasopoulos, 2007; Pino & Bassi, 2011). Comparisons have mainly been made with neuropsychological tests because of the wide use of the BVRT for clinical purposes. For example, McKay, Casey, Wertheimer, and Fichtenberg (2007) searched for evidence of convergent validity for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in a sample of patients with brain trauma. The authors correlated the battery of subtests that assesses visuoconstructive capacity (Figure Copy and Line Orientation) with Number Correct and Error Scores on the BVRT, finding favorable results for most of the correlations. Uc, Rizzo, Anderson, Sparks, Rodnitzky, and Dawson (2007) also indicated evidence of validity for the BVRT when they compared the performance of drivers with and without Parkinson's disease who were all licensed to drive. The battery had visual, cognitive, and motor ability tests. Each driver also performed a supervised task, making a route in an instrumented vehicle. Among other memory instruments, the BVRT showed significant correlations with the measures of the abilities mentioned.

The Brazilian version of the BVRT (Salles et al., 2014; Segabinazi, Barcelos, Souza, Castro, & Bandeira, 2014) maintains the original scoring system, but some aspects of the definitions of the errors were modified (Segabinazi, Duarte Junior, Salles, Bandeira, Trentini, & Hutz, 2013). Several investigations of

this version have evaluated the test's psychometric parameters. For example, studies have investigated agreement among evaluators (Segabinazi et al., 2014) and correlations with other tests that assess similar constructs (Zanini, Wagner, Salles, Bandeira, & Trentini, 2011). Additionally, performance norms for children, adolescents, university students, and older adults and comparisons of test performance between healthy samples and several clinical samples have been presented. Because of the importance of the BVRT for neuropsychologically assessing older adults, the present study sought to verify evidence of the test's criterion validity. To accomplish this, we compared performance in a group of older adults with possible AD with controls. Our hypothesis was that the clinical group would have poorer performance on the BVRT than the control group, especially in Administration A (Memory).

Method

Design and Participants

The present study used a cross-sectional design to compare clinical and control groups using a convenience sampling. The clinical group was composed of 16 older adults with a possible diagnosis of AD. This group had fewer participants than the control group because of the late diagnosis, in addition to a rigorous sample selection process. Therefore, older adults with other types of dementia did not participate in the study. More than 10 people with a possible diagnosis of AD were also excluded, once they could not understand or execute the tasks because of the advanced stage of their disease. All 16 participants in the clinical group lived in Porto Alegre or a metropolitan area, and most of them lived in nursing homes (68.75%). This group was predominantly composed of women (68.8%). The control group was composed of 50 older adults from the BVRT's normative sample (Salles et al., 2014). All of them lived in the community in the same area as the clinical group, and 41 were women (82%). Both groups had individuals with low, middle, and high levels of education (Kochhann, Varela, de Macedo Lisboa, & Chaves, 2010).

Table 1 shows the sociodemographic data and Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Kochhann et al., 2010) scores for both groups. Additional analyses were performed that compared age, years of formal education (*t*-test), and gender (χ^2 test). A significant difference ($p < .001$) in age was found between groups, with no significant differences in years of formal education ($p = .604$) or gender ($p = .259$). The clinical MMSE scores revealed a significant difference ($p < .001$) between the control and clinical groups ($t_{16,3} = .88$).

Selection Criteria

The inclusion criteria for the control group included the following: literate older adult, ≥ 1 year of formal education, >60 years old, absence of self-reported

diagnoses of neurological and psychiatric diseases, no signs of depression or dementia assessed by exclusion instruments, and a subject in the BVRT's normative sample (Salles et al., 2014).

Table 1. Sociodemographic data of the control and clinical groups.

		Control Group (n = 50)	Clinical Group (n = 16)
Age (years) ^a	M (SD)	69.26 (5.53)*	80.15 (6.05)*
	Range	60 to 85	72 to 90
Gender ^b	Male	9	5
	Female	41	11
Formal Education (years) ^a	Low (1 to 5)	15 (30%)	6 (37.5%)
	Middle (6 to 11)	19 (38%)	5 (31.25%)
	High (≥ 12)	16 (32%)	5 (31.25%)
	M (SD)	9.14 (4.98)	8.38 (5.51)
	Range	1 to 19	1 to 16
MMSE Score ^a	M (SD)	28.52 (1.61)	19.75 (4.36)*
	Range	24 to 30	15 to 26
Habitation	Nursing Homes	—	11 (68.75%)
	Community	50 (100%)	5 (31.25%)

^a t-test; ^b χ^2 test. * $p < .001$.

The selection criteria for the clinical group included the following: literate older adult, ≥1 year of formal education, diagnosis of possible AD (McKhann et al., 2011; Brucki, & Schultz, 2011) by a physician (geriatrician, neurologist, or psychiatrist), and able to understand and execute the tasks (which is not possible in the advanced stages of dementia).

Instruments and procedures for data collection

The participants in both groups also comprised the sample of other studies, which diversifies certain aspects of data collection, but all the studies were designed to avoid interference from one instrument in another, in addition to fatigue. In the control group, all of the participants were originally from the BVRT's normative sample. Participants with a high level of education completed a sociodemographic form and the Geriatric Depression Scale (GDS; Yesavage et al., 1983; Almeida, & Almeida, 1999) in a group session because both instruments are self-administered. In a second session, the BVRT and MMSE were administered individually. The participants were assessed in classrooms, and the total average time of administration was 25 min per session. Older adults with low and middle levels of education were evaluated individually, even for the self-administered instruments, to avoid errors caused by a lack of understanding. The evaluation occurred in two 25-min sessions. The sociodemographic data form, BVRT, and MMSE were administered in the first session, and the GDS-15 was administered in the second session.

The participants in the clinical group were assessed individually at their homes or nursing homes after the researchers obtained their caregivers' signed consent and the subjects' own consent (verbal and/or signed) to participate in the study. In cases in which the participants lived in nursing homes, consent was obtained from the responsible parties for the institution and also the participant. Each assessment occurred in one to three sessions according to the willingness and performance of the participants in the following tasks: sociodemographic data form, MMSE, and BVRT (in this order).

Sociodemographic data form

This form collected information on gender, age, and years of education. Medical background data were also obtained.

Benton Visual Retention Test

The BVRT (Benton Sivan, 1992) assessed visual memory and visuoconstructive abilities. The Brazilian standardization, composed of Administration A (Form C) and Administration C (Form D), was used. Both Administrations had 10 designs with increasing complexity. Most of the designs contained three geometric figures: two main figures and one peripheral lower figure. In Administration A (Form C) or Administration A (Memory), the examinee viewed a design for 10 s and immediately reproduced it from memory. In Administration C (Form D) or Administration C (Copy), the examinee reproduced a design in the presence of stimuli, with an unlimited exposure time. In both Administrations, all 10 designs were reproduced. The administration time for the entire test was approximately 10-20 min. Two main scores were obtained: Number Correct Score (i.e., the number of correctly reproduced designs) and Number Error Score (i.e., the sum of all errors in all of the designs). The errors were also subdivided into six categories: Omissions, Distortions, Perseverations, Rotations, Misplacements, and Size Errors, which might occur in the main and/or peripheral figures (right or left).

Mini Mental State Examination

The MMSE (Folstein et al., 1975) was used to exclude participants from the control group who had cognitive impairment and provided information about cognition in the clinical group. It is composed of items that assess several cognitive functions such as orientation in time and space, immediate and delayed recall, attention and calculation, language, and visuoconstructive capacity. Cut-off points were used only for the control group: 22 for a low level of education (1-5 years of formal education), 23 for a middle level of education (6-11 years of formal education), and 24 for a high level of education (≥12 years of formal education; Kochhann et al., 2010).

Geriatric Depression Scale

The GDS (Yesavage et al., 1983) was administered to exclude participants from the control group with evidence of depression. It is a questionnaire that assesses the intensity of depressive symptoms in older adults, and its questions must be answered objectively (yes or no). We used the short form of the Brazilian version, which contains 15 items (Almeida & Almeida, 1999). The total score is obtained by summing the scores attributed to the answers checked, and the cutting point proposed by the authors is 5/6 (not case/case).

Procedures for data analysis

Descriptive and inferential data analyses were performed. Frequencies, means, and standard deviations were calculated. Comparisons among gender, age, and years of formal education were made using the χ^2 test and *t*-test. The analyses revealed a statistically significant difference between the clinical and control groups only for age. Nonlinear changes in memory performance mainly at older ages were investigated. The Shapiro-Wilk test of residual distributions for the analyses of covariance (ANCOVAs) showed a normal (Gaussian) distribution only for the following variables: Number Correct Score, Number Errors Score, and Distortion and Perseveration errors for memory (Administration A). For all of the remaining variables, the sample did not originate from a population with a normal distribution. In fact, most of the variables fit a Poisson distribution. Considering this, the results using both *F* and Wald statistics (for Poisson distributions) were compared. These comparisons showed that the result was the same for all of the variables using both types of analysis (i.e., significant differences were found for the same measures, and vice versa). ANCOVAs were conducted for BVRT scores, controlling for age between groups, and effect sizes (*d*) were calculated (Cohen, 1988). Associations with *p* < .05 were considered significant.

Results

Comparisons of BVRT scores in the control and clinical groups are shown in Tables 2 and 3.

Table 2. Performance on Administration A (Memory) in the control and clinical groups and mean comparison test (ANCOVA).

Administration A- (Memory)	Control Group (n = 50)	Clinical Group (n = 16)	F	p	d
	M (SD)	M (SD)			
Number Correct Score	5.04 (2.02)	1.38 (1.78)	14.082	< .001*	1.9
Number Error Score	8.10 (3.67)	20.31 (6.46)	28.180	< .001*	2.4
Omissions	.90 (1.50)	5.25 (3.72)	14.959	< .001*	1.7
Distortions	4.02 (2.03)	9.38 (4.0)	18.188	< .001*	1.8
Perseverations	.88 (.92)	1.06 (1.06)	.584	.448	.2
Rotations	1.50 (1.49)	1.38 (1.46)	.112	.739	.1
Misplacements	.72 (1.01)	2.13 (2.45)	.342	.561	.8
Size Errors	.20 (.50)	.81 (1.17)	4.326	.042*	.7

**p* < .05.

According to Table 2, in Administration A (Memory), significant differences were found between groups in the main scores of the BVRT (i.e., Number Correct Score and Number Error Score). Significant differences were also found between groups in Omissions, Distortions, and Size Errors, and the effect sizes ranged from medium to large. Therefore, the clinical group showed poorer performance than the control group in most BVRT scores.

Table 3. Performance on Administration C (Copy) of the BVRT in the control and clinical groups and mean comparison test (ANCOVA).

Administration C- (Copy)	Control Group (n = 50)	Clinical Group (n = 16)	F	p	d
	M (SD)	M (SD)			
Number Correct Score	9.34 (1.15)	6.19 (2.56)	23.965	< .001*	1.7
Number Error Score	.72 (1.23)	5.63 (5.45)	16.600	< .001*	1.5
Omissions	.02 (.14)	1.13 (2.00)	3.544	.064	1.0
Distortions	.32 (.65)	2.88 (4.02)	9.392	.003*	1.1
Perseverations	#	.06 (.25)	—	—	—
Rotations	.04 (.20)	.38 (.62)	11.441	.001*	.8
Misplacement	.28 (.70)	1.00 (1.03)	5.386	.024*	.83
Size Errors	.06 (.31)	.13 (.34)	1.146	.289	.2

**p* < .05; #No errors of this type so no test was performed.

As shown in Table 3, in Administration C (Copy), significant differences were found between groups in Number Correct Score, Number Error Score, Distortions, Rotations, and Misplacement, with medium and medium-to-large effect sizes. A trend toward significance (*p* < .10) was found for Omissions and Perseverations, and the effect sizes for these differences ranged from small to large. Therefore, the clinical group showed poorer performance than the control group, even when memory was not specifically assessed. Qualitatively, the clinical group had a higher Number Error Score for Administration A (Memory) and a lower Number Correct Score compared with Administration C (Copy).

Discussion

The present results provide evidence of criterion validity for the BVRT, revealing differences between the control and clinical groups in test performance. A higher Number Correct Score and lower Number Error Score were observed in the control group in both Administrations. With regard to specific error categories, significant differences were found in both Administrations for the Distortions category. Moreover, in Administration A (Memory), significant differences were found in the number of Omission, Distortion, and Size errors. In Administration C (Copy), significant differences were found in Rotation, Distortion, and Misplacement errors. The effect sizes among the variables ranged from medium to large, indicating the

existence of effective differences among subjects in the skills assessed (Cohen, 1988).

According to the hypothesis of the present study, the clinical group actually presented poorer BVRT scores. According to previous studies (Coman et al., 1999; Seo et al., 2007; Messinis, Lyros, Georgiou, & Papathanasopoulos, 2009; Strauss et al., 2006; Carret et al., 2003), the variables age and level of education also interfere with BVRT performance. We assessed differences in age and years of education between groups. Because of the difference in age, ANCOVAs were performed. This analysis showed that the differences obtained could be attributed to the actual diagnosis rather than to other variables. With regard to level of education, no significant differences were found between groups. This was consistent with Kawas et al. (2003) in which education was shown not to interfere with their regression models.

Although most studies do not perform their data analyses differently for Administration A or C or specific types of errors, it is important to note that each Administration assesses different functions. Additionally, the error distribution may supply information about the performance of groups with specific diagnoses. Qualitatively, all types of errors were more frequent for Administration A (Memory) compared with Administration C (Copy) in both groups. This result was expected because of the absence of stimuli in the first administration. The frequency of Omission, Distortion, and Size errors in Administration A (Memory) and Distortion, Rotation, and Misplacement errors in Administration C (Copy) was higher in the clinical group, indicating the higher sensitivity of this type of error in the evaluation of patients with AD. The study by Robinson-Whelen (1992) was one of the few to specify the types of BVRT administrations and errors, thus allowing more detailed comparisons with the present data.

In Administration A (Memory), Robinson-Whelen (1992) found significant differences between groups in Omission, Distortion, and Rotation errors. The same differences were found in the present study, except that we did not find Rotation errors but did observe Size errors. In Administration C (Copy), Robinson-Whelen found a higher number of Omission, Distortion, Rotation, Misplacement, and Size errors in the AD group, a result that is similar to the present study, with the exception of Omission and Size errors. Importantly, Robinson-Whelen (1992) worked with a larger sample than the one used in the present study (122 healthy older adults and 191 AD participants). Nevertheless, even with a smaller sample size, most of the group effects were consistent with Robinson-Whelen (1992). Some types of errors did not show any significant difference between groups in either study, indicating that some skills were preserved in AD participants. Indeed, some abilities must be minimally preserved to perform the test, and understanding and executing the instructions were selection criteria in both studies.

The expectation of differences between Administration A (Memory) and Administration C (Copy) in the present study was confirmed, but statistically significant differences were found between groups in both administrations. This finding and those of Robinson-Whelen (1992) indicate other important impairments in AD that are related to visual perception, visual attention, and visuoconstructive skills. A review of visuospatial scores as biomarkers of AD (Mandal, Joshi, & Saharan, 2012) reported some important findings, but it did not present any study that used the same copying task that the BVRT provides.

The original manual for the BVRT considers that the test assesses only visual memory in Administration A (Memory). Although, other types of memory are also involved in this task such as working memory, attention (Lockwood et al., 2011), verbal short-term memory, and long-term memory). Moreover, during the execution of Administration A (Memory), aspects of episodic memory are also likely required. According to Tulving (1972), one of the aspects of episodic memory is the capacity to store events that have time and space information.

Therefore, each BVRT stimulus can be considered an episode composed of microepisodes. The literature shows that AD is associated with episodic memory damage (Salthouse & Becker, 1998), and the findings of the present study corroborate this statement. Episodic memory deficits are central to a variety of limitations in the AD patient's daily life. Together with lower performance in Administration C (Copy), which was expected because of the relationship between AD and visuospatial impairment, a global cognitive decline could be identified in our clinical sample. This decline has also been reported in previous studies (e.g., Mandal et al., 2012).

Some limitations of the present study should be noted. The level of dementia is usually determined with the help of an informant, but the present study did not have informants with the participants throughout the evaluation, mainly when the participants lived in a nursing home. Therefore, obtaining more detailed scores was not possible. Most of the AD participants were likely in the moderate stage of the disease because of the late diagnosis, and we excluded participants who did not understand or execute the instructions. This exclusion of participants may have prevented verification of the evolution of abilities through the stages of the disease the way Robinson-Whelen (1992) performed her study.

The fact that most of the AD participants lived in institutions did not allow an investigation of socioeconomic status. Depression symptoms also were not evaluated because of the participant's limited understanding of the items of the instrument.

Despite these limitations, the findings of the present study provide new evidence of the psychometric properties of the BVRT in assessing visual memory and visuoconstructive abilities. The data provide evidence

of criterion validity for the BVRT and corroborate the literature with regard to the legitimacy of the instrument (Golski et al., 1998; Kawas et al., 2003; Messinis et al., 2007), indicating that it can be used to help with the early diagnosis of AD in the Brazilian population.

Final considerations

The present study provided additional criterion validity for the BVRT in its main measures—Number Correct Score and Number Error Score—and some specific types of errors. As expected, performance in the control group was better than in the clinical group in all of the measures. In Administration A (Memory), differences in Number Correct Score, Number Error Score, Omission errors, Distortion errors, and Size errors were statistically significant. In Administration C (Copy), differences in Number Correct Score, Number Error Score, Distortion errors, Rotation errors, and Misplacement errors were statistically significant. Omissions and Perseverations also showed a tendency toward significance in Administration C. These differences were observed despite the reduced sample size. Therefore, the use of the BVRT is suggested to help diagnose suspected AD and other neurological diseases among the Brazilian population. Most dementias cannot be diagnosed only using laboratory tests, so clinical and neuropsychological evaluations may be valuable in such cases for the early diagnosis and assessment of the evolution of the disease.

Acknowledgements

Adriana Mokwa Zanini was supported by a doctoral grant from CNPq. Maxciel Zortea and Joice Dickel Segabinazi were supported by doctoral grants from CAPES.

References

- American Psychiatric Association (2002). *Manual Diagnóstico e Estatístico dos Transtornos Mentais*, IV-TR. YEAS, p. 6: Artmed.
- Almeida, O.P., & Almeida, S.A. (1999). Confiabilidade da versão brasileira da Escala de Depressão em Geriatria (GDS) versão reduzida. *Arquivos de Neuropsiquiatria*, 57(2B): 421-426.
- Baddeley, A.D., & Hitch, G. (1974). Working memory. In: G.H. Bower (Ed.), *The psychology of learning and motivation: Volume 8. Advances in research and theory* (pp. 47-89). New York: Academic Press.
- Baddeley, A. D., Allen, R. J., & Hitch, G. (2011). Binding in visual working memory: the role of the episodic buffer. *Neuropsychologia*, 49, 1393-1400.
- Benton Sivan, A. (1992). *Benton Visual Retention Test*. San Antonio, TX: Psychological Corporation.
- Brucki, S., & Schultz, R. (2011). Recomendações em Alzheimer. *Dementia & Neuropsychologia*. YEAS, p. 6 and 10]: Academia Brasileira de Neurologia.
- Burin, D., Drake, M., & Harris, P. (2007). *Evaluación neuropsicológica en adultos*. Buenos Aires: Paidós.
- Caramazza, A., & Coltheart, M. (2006). Cognitive neuropsychology twenty years on. *Cognitive Neuropsychology*, 23(1), 3-12.
- Carret, N., Rainville, C., Lechevallier, N., Lafont, S., Letenneur, L., & Fabrigoule, C. (2003). Influence of education on the Benton Visual Retention Test performance as mediated by a strategic search component. *Brain and Cognition*, 53, 408-411.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale, NJ: Erlbaum.
- Coman, E., Moses, J.A., Jr., Kraemer, H.C., Friedman, L., Benton, A.L., & Yasavage, J. (1999). Geriatric performance on the Benton Visual Retention Test: demographic and diagnostic considerations. *Clinical Neuropsychologist*, 13(1), 66-77.
- Eslinger, P.J., Damasio, A.R., Benton, A.L., & Van Allen, M. (1985). Neuropsychologic detection of abnormal mental decline in older persons. *Journal of the American Medical Association*, 253, 670-674.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Golski, S., Zonderman, A.B., Malamut, B.L., & Resnick, S.M. (1998). Verbal and figural recognition memory: task development and age associations. *Experimental Aging Research*, 24(4), 359-385.
- Hogan, T. (2006). *Introdução à prática de testes psicológicos*. Rio de Janeiro: LTC.
- Kawas, C.H., Corrada, M.M., Brookemeyer, R. Morrison, A., Resnick, S.M., Zonderman, A.B., & Aremberg, D. (2003). Visual memory predicts Alzheimer's disease more than a decade before diagnosis. *Neurology*, 60(7), 1089-1093.
- Kochhann, R., Varela, J.S., de Macedo Lisboa, C.S., & Chaves, M.L.F. (2010). The Mini Mental State Examination: review of cutoff points adjusted for schooling in a large Southern Brazilian sample. *Dementia and Neuropsychology*, 4(1), 35-41.
- Lezak, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological assessment*. New York: Oxford University Press.
- Lockwood, C.A., Mansoor, Y., Homer-Smith, E., & Moses, J.A., Jr. (2011). Factor structure of the Benton Visual Retention Tests: dimensionalization of the Benton Visual Retention Test, Benton Visual Retention Test-Multiple Choice, and the Visual Form Discrimination Test. *Clinical Neuropsychologist*, 25(1), 90-107.
- McKay, C., Casey, J.E., Wertheimer, J., & Fichtenberg, N.L. (2007). Reliability and validity of the RBANS in a traumatic brain injured sample. *Archives of Clinical Neuropsychology*, 22, 91-98.
- Mandal, P.K., Joshi, J., & Saharan, S. (2012). Visuospatial perception: an emerging biomarker for Alzheimer's disease. *Journal of Alzheimer's Disease*, 31(Suppl. 3), S117-S135.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34, 939-944.
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Jr., Kawas, C.H. ... & Phelps, C.H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia*, 7, 263-269.
- Messinis, L., Lyros, E., Georgiou, V., & Papathanasopoulos, P. (2009). Benton Visual Retention Test performance in normal adults and acute stroke patients: demographic considerations, discriminant validity, and test-retest reliability. *Clinical Neuropsychologist*, 23, 962-977.
- Messinis, L., Tsakona, I., Mlefaki, S., & Papathanasopoulos, P. (2007). Normative data and discriminant validity of Rey's Verbal Learning Test for the Greek adult population. *Archives of Clinical Neuropsychology*, 22, 739-752.
- Pasquali, L. (2001). Parâmetros psicométricos dos testes psicológicos. In: L. Pasquali (Ed.), *Técnicas de exame psicológico - TEP - manual: fundamentos das técnicas psicológicas* (pp. 111-136). São Paulo: Casa do Psicólogo.
- Pino, J., & Bassi, N. (2011). Predictive ability of the MMSE in detecting impairments of higher cerebral functions in Hispanic patients that report memory difficulties. Annual Meeting of the National Academy of Neuropsychology Marco Island. November, 16-19.
- Resnick, S.M., Trotman, K.M., Kawas, C., & Zonderman, A.B. (1995). Age-associated changes in specific errors on the Benton Visual Retention Test. *Journal of Gerontology: Series B. Psychological Sciences and Social Sciences*, 50, P171-P178.

- Robinson-Whelen, S. (1992). Benton Visual Retention Test performance among normal and demented older adults. *Neuropsychology, 6*(3), 261-269.
- Salles, J.F., Bandeira, D.R., Trentini, C.M., Segabinazi, J.D., & Hutz, C.S. (2014). *Manual do Teste de Retenção Visual de Benton*, in press.
- Salmon, D.P., & Bondi, M.W. (2008). Neuropsychological assesment of dementia. *Annual Review of Psychology, 60*, 257-282.
- Salthouse, T.A., & Becker, J.T. (1998). Independent effects of Alzheimer's disease on neuropsychological functioning. *Neuropsychology, 12*(2), 242-252.
- Segabinazi, J.D., Duarte Junior, S., Salles, J.F., Bandeira, D.R., Trentini, C.M., & Hutz, C.S. (2013). Teste de Retenção Visual de Benton: apresentação do manual brasileiro. *Avaliação Psicológica, 12*(3), 429-432.
- Segabinazi, J.D., Barcelos, E.D., Souza, T.S., Castro, S.M., & Bandeira, D.R. (2014). Estudos de fidedignidade do Teste de Retenção Visual de Benton. In: *Manual do Teste de Retenção Visual de Benton*, in press.
- Seo, E.H., Lee, D.Y., Choo, I.H., Youn, J.C., Kim, K.W., Jhoo, J.H. ... & Woo, J.I. (2007). Performance on the Benton Visual Retention Test in an educationally diverse elderly population. *Journal of Gerontology: Series B. Psychological Sciences and Social Sciences, 62*(3), P191-P193.
- Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary*, 3rd edition. New York: Oxford University Press.
- Tulving, E. (1972). Episodic and semantic memory. In: E. Tulving, & W. Donaldson (Eds.). *Organization of memory* (pp. 381-403). New York: Academic Press.
- Urbina, S. (2004). *Essentials of psychological testing*. Hoboken, NJ: Wiley.
- Uc, E.Y., Rizzo, M., Anderson, S.W., Sparks, J.D., Rodnitzky, R.L., & Dawson, J.D. (2007). Impaired navigation in drivers with Parkinson's disease. *Brain, 130*, 2433-2440.
- Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., & Leirer, V.O. (1983). Development and validation of a geriatric depression rating scale: a preliminary report. *Journal of Psychiatric Research, 17*, 37-49.
- Zanini, A.M., Wagner, G.P., Salles, J.F., Bandeira, D.R., & Trentini, C.M. (2011). Teste de Retenção Visual de Benton (BVRT): evidências de validade para idosos. *Avaliação Psicológica, 11*(2), 287-296.
- Zonderman, A.B., Giamba, L.M., Arenberg, D., Resnick, S.M., Costa, P.T., Jr., & Kawas, C.H. (1995). Changes in immediate visual memory predict cognitive impairment. *Archives of Clinical Neuropsychology, 10*, 111-123.
- Zuccolo, P., Rzezak, P., & Góis, J. (2010). Praxia e visuoconstrução. In: L. Malloy-Diniz, D. Fuentes, P. Mattos, & N. Abreu (Eds.), *Avaliação neuropsicológica* (pp. 114-122). Porto Alegre: Artmed.