Psychology & Neuroscience

1983-3288/19/$12.00

2019, Vol. 12, No. 2, 169–179
http://dx.doi.org/10.1037/pne0000157

SPECIAL ISSUE: NEUROPSYCHOLOGY OF AGING

Evidence of Clinical, Criterion, and Convergent Validity of the Brazilian Version of the Picture Free and Cued Selective Reminding Test With Immediate Recall (pFCSRT-IR)

Murilo Ricardo Zibetti
Federal University of Rio Grande do Sul – Brazil and Tuiuti University of Paraná – Brazil

Andressa Hermes Pereira
Universidade Federal do Rio Grande do Sul

Ana Julia Lehnen
Federal University of Health Sciences of Porto Alegre – Brazil

Gabriela Zuñeda Peres Duarte
Federal University of Rio Grande do Sul – Brazil

Fabiula de Fátima Machado dos Santos
Faculty of Development of Rio Grande do Sul – Brazil

Renata Kochhann and Rochele Paz Fonseca
Pontifical Catholic University of Rio Grande do Sul – Brazil

Clarissa Marceli Trentini
Federal University of Rio Grande do Sul – Brazil

The picture Free and Cued Selective Reminding with Immediate Recall (pFCSRT-IR) has been widely useful for episodic memory assessment in patients with dementia due to Alzheimer's disease (AD). This study presents adaptation and evidence of criterion and concurrent validity of the pFCSRT-IR for use in Brazil. The study was conducted in two complementary steps. Step A: authorization, translation, back-translation, and selection of a set of the 16 most nameable pictures in Brazil. Step B: $n = 50$ elderly individuals, 61–86 years of age ($M = 73.26$, $SD = 6.70$) and 3–25 years of education ($M = 10.80$, $SD = 5.76$), were evaluated with the pFCSRT-IR. Participants were categorized into two paired groups (25 AD/25 controls) for comparison in the pFCSRT-IR. In addition, a correlational analysis was conducted among pFCSRT-IR and Rey Auditory-Verbal Learning Test (RAVLT) scores. Performance on the pFCSRT-IR was lower for the AD group. Moreover, considering all sample scores in the pFCSRT-IR and

This article was published Online First March 14, 2019.

Murilo Ricardo Zibetti, Postgraduate Psychology Program, Federal University of Rio Grande do Sul (UFRGS) and Postgraduate Psychology Program, Tuiuti University of Paraná (UTP); Andressa Hermes Pereira, Medicine Post-Graduation Program, Federal University of Rio Grande do Sul (UFRGS); Ana Julia Lehnen, Psychology Student, Federal University of Health Sciences of Porto Alegre (UFCSPA); Gabriela Zuñeda Peres Duarte, Psychology Student, Federal University of Rio Grande do Sul (UFRGS); Fabiula de Fátima Machado dos Santos, Psychology Student, Faculty of Development of Rio Grande do Sul (FADERGS); Renata Kochhann, Post-Doc position on Psychology Post-Graduate Program, Pontifical Catholic University of Rio Grande do Sul (PUC-RS); Rochele Paz Fonseca, Postgraduate Psychology Program, Pontifical Catholic University of Rio Grande do Sul (PUC-RS); Clarissa Marceli Trentini, Postgraduate Psychology Program, Federal University of Rio Grande do Sul (UFRGS).

Correspondence concerning this article should be addressed to Murilo Ricardo Zibetti, Tenente Ary Tarrago Street, 3095, Porto Alegre, RS, Brazil. E-mail: mrzibetti@gmail.com

the RAVLT showed a significantly positive correlation with moderate to strong size ($\rho >$ .40). The analysis showed evidence of convergent validity with external variables for the pFCSRT-IR. In conclusion, the Brazilian version of pFCSRT-IR has been shown to be a valid measure of episodic memory in the continuum of memory loss in aging.

Keywords: Free and Cued Selective Reminding Test, FCSRT, memory, Alzheimer's disease, development

The detection of impairments in episodic memory is the main criterion for the diagnosis of typical presentation of dementia due to Alzheimer's disease (AD), according to institutions such as the National Institute on Aging and the Alzheimer's Association (NIA/AA; McKhann et al., 2011), the American Psychiatric Association (APA; 2014), and independent researchers in the area of neurocognitive disorders (Dubois et al., 2007, 2014; Dubois, Picard, & Sarazin, 2009). For a more accurate diagnosis of AD, those institutions recommend that episodic memory impairments be measured by a standardized neuropsychological assessment and that test results be associated with impairments in the independence of patients' daily life activities (APA, 2014; McKhann et al., 2011).

If, on the one hand, the presence of deficits in episodic memory tasks may be considered an important biomarker of AD (Noel-Storr et al., 2013), on the other hand, even normal aging itself may reduce performance in episodic memory tests (Koen & Yonelinas, 2014). Additionally, other neuropsychiatric conditions, such as depression, lead to a lower performance in memory recall, complicating the differential diagnosis of AD (Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010). To address this problem, a consortium of researchers on AD (International Working Group [IWG]) included the Free and Cued Selective Reminding Test (hereafter, called the FCSRT) in the diagnostic algorithm of AD (Dubois et al., 2007, 2014).

The FCSRT recruits processes of controlled learning and cued recall, both guided by semantic cues, to minimize the effects of secondary cognitive impairments on the results of the episodic memory test (Buschke, 1984; Grober, 2014; Grober & Buschke, 1987; Grober, Buschke, Crystal, Bang, & Dresner, 1988). This kind of clinical application of the encoding specificity principle (Tulving & Thomson, 1973) allows the detection of hippocampal amnesia, improving the sensitivity for impairments of phenotypic memory in AD (Dubois et al., 2007, 2014). Considerable evidence has been accumulated over time corroborating the hypothesis that the FCSRT and its variations may identify patients in very early stages of AD (Grober, Sanders, Hall, & Lipton, 2010; Lemos, Simões, Santiago, & Santana, 2015), and predict the conversion from mild cognitive impairment to AD (Auriacombe et al., 2010; Grober, Mowrey, Katz, Derby, & Lipton, 2015; Mura et al., 2014). Additionally, it has been suggested that results from the FCSRT are related to cerebral metabolism (Caffarra et al., 2016; Wagner et al., 2012).

The traditional version of the FCSRT comprises 16 stimuli that are presented with their respective semantic cues (e.g., which one is the fruit? Grape). A distractor task is performed after encoding all items, followed by a free and a cued recall task using the same semantic cues. This process occurs throughout three consecutive trials and after 20 min (late recall; Grober & Buschke, 1987). Based on the original works of Buschke (1984) and Grober and Buschke (1987), some variations of the FCSRT have been developed. The four most widely employed versions differ from the original version in two main characteristics: the nature of the stimulus (words vs. pictures) and the presence of immediate recall (IR). The IR is an additional encoding process performed in the study phase when participants are presented a set of four items and make a cued recall immediately. This is a way of ensuring that participants learn the associations between item and cue before performing the free recall (Grober et al., 2010) and provides an opportunity for item transference from working memory to the medial temporal lobe system (Swerdlow & Jicha, 2012).

The few studies comparing these versions usually demonstrated that the pictoric FCSRT with IR (hereafter called pFCSRT-IR) have higher scores when compared with the version of FCSRT with words and without IR

(Slachevsky et al., 2014, 2018; Zimmerman et al., 2015). However, in a sample of healthy elderly individuals, the moderate and positive correlation between the scores of the two versions was highlighted (Zimmerman et al., 2015); as well as its similar diagnostic accuracy for AD patients (Delgado et al., 2016). Therefore, although they differ in some aspects, and in spite of some criticisms regarding the use of words/pictures and IR (Buschke, 2002; Grober et al., 2010; Lemos, Martins, Simões, & Santana, 2012), most versions exhibit good psychometric properties, so that the choice between them also depends on the characteristics of the target population.

In the Brazilian context, due to the low educational level of many elderly patients (Pilger, Menon, & de Freitas Mathias, 2011), the use of a less reading-dependent version would be adequate. In this case, the use of pictorial stimuli and IR (pFCSRT-IR) allows the assessment of episodic memory despite reading deficiencies and helps with detecting when the information is correctly encoded (Grober, 2014). However, the only version currently adapted to (European) Portuguese uses words and lacks measures of IR (Lemos et al., 2012). Considering the adequacy of the pFCSRT-IR to the Brazilian target population, the goal of this study is to present the process of adaptation of the pFCSRT-IR to the Brazilian context, as well as to present evidence of criterion and concurrent validity.

To better understand the processes involved in transcultural adaptation and to obtain evidence of validity for the pFCSRT-IR, the remainder of this article was divided into two sections, corresponding to the stages of the study. In Step A, we report the authorization, translation, back translation, standardization, and selection of pictorial stimuli. In step B, we report the the process of obtaining evidence of concurrent and criterion validity for the instrument, respectively, through correlation with external variables and through a comparative clinical analysis.

Step A: Authorization, Translation, Back Translation, Stimuli Selection, and Pilot Study

The cross-cultural adaptation was conducted following guidelines presented by Borsa, Damásio, and Bandeira (2012) and by Fonseca et al. (2011), with slight modifications due to the visual nature of the items. The cross-cultural adaptation process included initially: authorization, independent translation of item names and their respective semantic cues, and back translation and a consensus version of the Brazilian pFCSRT-IR (Buschke & Grober, 2002).

Three parallel forms of the original instrument were received along with the authorization, each comprising 16 picture stimuli, for a total of 48 pictures. As some pictures were judged by the authors as incompatible with the Brazilian reality (e.g., pretzel), we aimed to create a single new version composed of the 16 items deemed most appropriate to the Brazilian population. For this purpose, a brief standardization of naming of the picture stimuli was conducted, similar to what was done in previous studies (Snodgrass & Vanderwart, 1980; Zibetti, Bordignon, & Trentini, 2015). This step resulted in the selection of pictures with a high proportion of correct naming and a low number of competitors for the names given to the pictures selected.

Subsequently, a pilot study was conducted to verify the adequacy of the administration of the items regarding nameability. Difficulties in understanding categories were registered and raters proposed changes to the final version of the instrument (both for the items and for the administration protocol). Due to copyright issues, visual stimuli are not available in the current paper, but they can be accessed through contact with the authors and permission of the copyright owners (Copyright from the Albert Einstein College of Medicine at the Yeshiva University). The detailed process of cross-cultural adaptation was better described elsewhere (Zibetti & Trentini, 2017).

Step B: Evidence of Validity of the pFCSRT-IR

The goal of this step was to investigate the evidence of validity of the pFCSRT-IR through its relationship with variables external to the instrument (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). The sources of evidence were based on concurrent criteria with another instrument of memory assessment and on clinical/

criterion validity. To investigate evidence of criterion validity, mean performance in the pFCSRT-IR was compared between elderly individuals with mild impairments due to AD and healthy controls. To explore evidence of validity based on the relationship with other tests, we assessed the correlations of those participants' scores in the pFCSRT-IR with their performance in a classic task to evaluate episodic memory: the Rey Auditory-Verbal Learning Test (RAVLT; Rey, 1958). The hypothesis in both analyses was that the scores in the pFCSRT-IR in episodic memory would differentiate between performance of patients with AD and performance of controls. For the same reason, we expected a moderate and positive correlation with scores in the RAVLT.

Participants

The participants were 50 elderly individuals (aged above 60 years), speakers of Brazilian Portuguese, without uncorrected perceptual impairments or psychiatric disorders (bipolar mood disorders and psychotic disorders) or history of previous neurological disorders (stroke, Parkinson's disease) except for AD. The majority of the participants were female ($n = 30$). The participants' age ranged from 61–86 years of age ($M = 73.26$, $SD = 6.70$), and educational level varied between 3 and 25 years ($M = 10.80$, $SD = 5.76$). Participants were divided into two groups (25 patients with AD/25 controls); the selection criteria for each group are presented below. No statistically significant differences were observed between groups for gender, age, or educational level.

AD group. The AD group was comprised of $n = 25$ participants (14 women), with dementia due to probable AD, Aged 62–86 years ($M = 73.84$; $SD = 6.94$), with educational level between 3 and 18 years ($M = 9.16$, $SD = 5.30$). The research was presented among physicians with expertise in the care of dementia. The patients were referred for the research when they presented with memory complaints in outpatient consultation with their reference doctor. The first step of the study was a comprehensive neuropsychological assessment. Patients whose results of this evaluation indicated impairment of memory and other cognitive domains (scores below -1.5 standard deviation), as well as functional losses in daily life activities returned to the referring physician. Thus, the diagnosis of AD was confirmed by a reference physician when no laboratory or neuroimaging abnormalities were observed indicating another etiology of dementia or other pathology which could cause the symptoms. Participants with severe depression were not included in the study. Therefore, the diagnosis of AD followed the criteria of NIA/AA (McKhann et al., 2011), as recommended by the Brazilian Association of Neurology (Frota et al., 2011). The overall level of cognitive impairment of the patients was controlled by the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Patients' performance in the MMSE varied between 15 and 26 points ($M = 20.40$, $SD = 2.97$), and only patients with scores above 15 were included in the study. This cutoff point was used to avoid effects due to other cognitive impairments (e.g., language and attention) on the performance of the pFCSRT-IR, following a procedure employed in several other studies (Jongenelis et al., 2004; Smalbrugge et al., 2005). The stage of dementia was also inferred by the degree of impairment observed in the MMSE (Perneczky et al., 2006). This cutoff point also prevented the inclusion of patients in the severe stage of dementia in the sample. For patients with low education levels, only those with mild impairment (MMSE > 12) were included (Wajman, Oliveira, Schultz, Marin, & Bertolucci, 2014). The remaining participants could be classified as mild to moderate according to an international review (moderate: MMSE > 10; mild: MMSE > 20; Perneczky et al., 2006).

Control group. The control group was comprised of $n = 25$ elderly participants (16 women), Aged between 61 and 86 years of age ($M = 72.68$, $SD = 6.54$), with between 3 and 25 years of education ($M = 12.44$, $SD = 5.63$). They were matched by age to the clinical group (± 5 years). Only individuals with no (self-reported) cognitive complaints or functional losses in daily life activities were included. No participants were included who exhibited scores in the MMSE indicative of dementia (Kochhann, Varela, Lisboa, & Chaves, 2010), nor were individuals with results indicating impairments in memory in the comprehensive neuropsychological assessment (all subjects had scores above -1.5 standard deviation in the results).

Procedures and Instruments

Participants in the control group were contacted in the community, and data collection occurred in a single session in environments appropriate for the task. Participants in the AD group were indicated by specialist physicians and by the team that carried out the comprehensive neuropsychological evaluation. The following instruments were used in this step.

Sociodemographic and health conditions questionnaire. This questionnaire was used to describe the sample, as well as to investigate the exclusion criteria (according to group). The questionnaire is structured, with questions about demographic variables and inclusion criteria for the sample (neurological and psychiatric disorders, uncorrected sensory problems).

Brazilian version of the pFCSRT-IR. This instrument (Buschke, 1984; Grober & Buschke, 1987) was the focus of this study. It is composed of 16 nameable pictures coded semantically (e.g., which is the bird? Owl) followed by an IR task with semantic cues. Three trials were conducted with free recall and then cued recall of forgotten items. Between trials, participants performed a distractor task (counting). Finally, a late recall task was performed between 15 and 20 min after the initial stage (Buschke, 1984; Grober & Buschke, 1987). The main scores generated by the instrument are: immediate cued recall, free recall, total recall, free recall in the three trials, total recall in the three trials, delayed free recall, and delayed total recall. In all cases, total recall represents the sum of freely recalled items, plus those recalled with cues.

MMSE. The MMSE (Folstein et al., 1975) is a screening instrument for dementia composed of 30 questions and employed worldwide. We used the cutoff points, corrected for educational level, suggested by Kochhann et al. (2010) for the control group.

Activities of Daily Living Questionnaire (ADLQ). The ADLQ (Johnson, Barion, Rademaker, Rehkemper, & Weintraub, 2004) was administered only to the patient group, this questionnaire was answered by caregivers and indicates the functional impact of dementia in the patients' life. It is divided into 17 questions that assess six functional domains (self-care, interaction, intellectual activity, organization and planning, social participation, and nutrition). The version adapted for Brazil (Medeiros & Guerra, 2009) was used to detect functional losses.

RAVLT. The RAVLT (Rey, 1958) was used to assess episodic memory through repetition of word lists. The evaluation procedure and the control data employed were those for the Brazilian version, published by Malloy-Diniz, Lasmar, Gazinelli, Fuentes, and Salgado (2007). This instrument is considered an additional measure to assess episodic memory, and was included because it evaluates the same construct as the pFCSRT-IR. We used the four main RAVLT scores: Immediate Memory (A1); Learning (sum A1–A5); After Interference (A6) and Delayed Recall (A7).

Data Analysis

Data analysis showed that the data from the pFCSRT-IR, particularly the total recall score, showed differences to the normal curve according to the Shapiro-Wilk test ($p < .01$). Therefore, tests of criterion validity were conducted using Mann–Whitney mean ranks and a nonparametric receiver operating characteristic curve analysis. Cutoff points for pFCSRT-IR scores were defined by the Youden index. For evidence of concurrent validity, Spearman correlations were performed between the data obtained in the pFCSRT-IR and in the RAVLT.

Results

All scores from the pFCSRT-IR showed differences between the AD and control groups. Table 1 shows the mean data for each result in both groups.

The data in Table 1 corroborate the hypothesis that the Brazilian adaptation of the pFCSRT-IR generates scores that differentiate between mean performance in a group of patients with AD and a control group of elderly individuals without cognitive complaints. These data provide preliminary evidence of criterion validity for the adaptation of the instrument, with the most commonly used scores (sum of the three trials: free and total) showing high levels of sensitivity and specificity. In addition, considering the possibility that the inclusion of patients in moderate stages of AD could play a major role in the good sensitivity identified in the pFCSRT-IR, the analysis was repeated with

Table 1
Comparison Between the Control and AD Groups in the Mann-Whitney Test and Sensitivity and Specificity of the pFCSRT-IR for the AD Group

pFCSRT-IR score	Control group (n = 25) *M* (*SD*)	Control group *Mdn*	AD group (n = 25) *M* (*SD*)	AD group *Mdn*	Mann-Whitney *U*	Cutoff point	Sen (%)	Spe (%)	AUC	AUC (95% confidence interval)
Immediate cued recall	16.00 (.00)	16	12.36 (4.18)	14	112.50[a]	≤15	64	100	.820	[.696, .944]
Trial 1—free recall	9.08 (2.81)	9	3.32 (2.72)	3	45.000[a]	≤6	84	84	.928	[.862, .994]
Trial 1—total recall[b]	15.96 (.20)	16	10.08 (5.00)	11	41.500[a]	≤15	88	96	.934	[.855, 1.000]
Trial 2—free recall	11.36 (2.45)	11	3.16 (2.61)	3	14.000[a]	≤6	92	100	.978	[.943, 1.000]
Trial 2—total recall	16.00 (.00)	16	10.92 (4.73)	12	100.00[a]	≤15	68	100	.840	[.721, .959]
Trial 3— free recall	11.4 (2.33)	11	3.04 (3.35)	2	26.000[a]	≤5	80	100	.958	[.906, 1.00]
Trial 3— total recall	15.96 (.20)	16	10.84 (4.68)	12	92.000[a]	≤15	72	96	.853	[.739, .967]
Delayed free recall	12.96 (2.07)	13	3.40 (3.30)	3	5.000[a]	≤8	92	100	.992	[.976, 1.000]
Delayed total recall	15.92 (.28)	16	10.36 (5.07)	11	83.000[a]	≤15	72	96	.867	[.758, .976]
Sum of trials—free recall	31.84 (6.96)	31	9.52 (8.01)	7	18.500[a]	≤21	92	96	.970	[.878, .998]
Sum of trials—total recall	47.92 (.40)	48	31.84 (14.05)	33	31.000[a]	≤47	92	96	.950	[.859, .992]

Note. AD = dementia due to Alzheimer's disease; pFCSRT-IR = picture Free and Cued Selective Reminding with Immediate Recall Test; *M* = mean; *SD* = standard deviation; *Mdn* = median; *U* = value of Mann-Whitney test; Sen = sensitivity; Spe = specificity; AUC = area under the curve.
[a] Statistically significant ($p < .05$). [b] Total is the sum of free recall and selective cued recall.

a subsample of 14 AD patients with scores greater than 20 points on the MMSE. By using the same cutoff points there was no significant change in the sensitivity or specificity of the instrument. Another source of evidence of validity is the relationship with other instruments that presumably assess the same construct. Table 2 shows the results of correlations between the pFCSRT-IR and the main RAVLT scores. In all sample these results show moderate to high correlations between main scores in the RAVLT and those in the pFCSRT-IR.

Table 2 shows that in all trials, considering all samples, the correlations between the pFCSRT-IR and the RAVLT were higher for free recall than for total recall (sum of free recall and cued recall). Furthermore, delayed recall exhibits the most important correlations according to the measures provided by the pFCSRT-IR. When the data of the separate groups are considered the data present a smaller number of significant correlations and a lower intensity of the correlations. In the AD group, the observed pattern is like that of the all sample, that is, the later the evocation of the RAVLT, the higher the correlations. In turn, results from the control group indicate moderate and positive correlations between free evocation of pFCSRT-IR and RAVLT scores. It is no surprise that the total evocation (free more guided) does not present significant correlations to the RAVLT since there is a ceiling effect in this score.

Discussion and Conclusion

The goal of this study was to present the cross-cultural adaptation and, mainly, to obtain evidence of three sources of validation for the Brazilian version of the pFCSRT-IR. For the adaptation process, we reported the authorization, translation, and back-translation steps. Translation and back translation were performed following the International Test Commission Guidelines (Borsa et al., 2012). However, because this instrument is composed of nameable picture stimuli, some of the guidelines for adaptation of psychological instruments, particularly those concerning stimuli adequacy, could not be strictly followed. Despite this difficulty, the outcome of these processes are good nameability of the stimuli and the

Table 2
Spearman Correlation Between pFCSRT-IR and RAVLT

	RAVLT											
	Short memory Trial 1 (A1)			Learning (sum Trials A1–A5)			Interference (A6)			Late recall (A7)		
pFCSRT-IR	All (n = 50)	Control (n = 25)	AD (n = 25)	All (n = 50)	Control (n = 25)	AD (n = 25)	All (n = 50)	Control (n = 25)	AD (n = 25)	All (n = 50)	Control (n = 25)	AD (n = 25)
Imme. recall	.448	*ns*	*ns*	.678	*ns*	.559	.663	*ns*	.492	.722	*ns*	.550
Free 1	.569	.524	*ns*	.770	.656	*ns*	.816	.670	.424	.820	.470	.595
Total 1	.532	*ns*	*ns*	.732	*ns*	*ns*	.743	*ns*	.451	.759	*ns*	.470
Free 2	.605	.449	*ns*	.793	.522	*ns*	.828	.565	.663	.846	.478	.522
Total 2	.502	*ns*	*ns*	.693	*ns*	.460	.740	*ns*	.528	.738	*ns*	.450
Free 3	.591	.427	*ns*	.813	.611	.400	.829	.559	.620	.844	.485	.566
Total 3	.533	*ns*	*ns*	.689	*ns*	.463	.739	*ns*	.512	.706	*ns*	.462
Delayed free	.627	.441	*ns*	.803	.434	.541	.780	.411	.603	.854	.499	.503
Delayed total	.442	*ns*	*ns*	.653	*ns*	.510	.686	*ns*	.540	.721	*ns*	.466
Sum—free	.611	.517	*ns*	.818	.642	.419	.844	.658	.581	.862	.529	.602
Sum—total	.570	*ns*	*ns*	.769	*ns*	.470	.789	*ns*	.530	.786	*ns*	.480

Note. pFCSRT-IR = picture Free and Cued Selective Reminding with Immediate Recall Test; RAVLT = Rey Auditory-Verbal Learning Test; AD = dementia due to Alzheimer's disease; *ns* = not statistically significant; Total = the sum of free recall and selective cued recall; Imme. recall = immediate cued recall; Free = free recall; Numbers (1, 2, and 3) = trials; Delayed free = delayed free recall; Delayed total = delayed total recall; Sum = sum of three trials of study phase.

evidence of convergent and criterion validity of Brazilian pFCSRT-IR.

Once the final version of the pFCSRT-IR was fully adapted, the next step was to search for evidence of validity for the instrument. In this step, AD patients showed statistically lower performance compared with healthy participants (control group). Therefore, we conclude that the scores in the Brazilian version of pFCSRT-IR showed evidence of criterion validity for AD patients. Furthermore, both scores of Sum of trials (free and total) exhibited high levels of sensitivity and specificity for this sample, and are the most often scores employed of pFCSRT-IR. Taken together, these results show that this version is a useful tool to assess episodic memory in the Brazilian population. This study adds to the evidence, presented in several other studies, of validity for the pFCSRT-IR (Derby et al., 2013; Grober et al., 1988; Grober et al., 2010), as well as the use of controlled learning and cued recall for detection of AD (Lemos et al., 2015; Sarazin et al., 2007). This version also shows other results compatible with international findings, for example, the presence of a learning curve in free recall, and ceiling effect on total recall, for the control groups and absence of those effects for the AD group (Derby et al., 2013; Petersen, Smith, Kokmen, Ivnik, & Tangalos, 1992). An interesting result of this study was the statistical difference in IR between the AD group and the control group. This result is rarely presented in other studies, because it is considered as a measure of quality of encoding. However, it is consistent with the literature showing encoding deficits in patients with AD (Oltra-Cucarella, Pérez-Elvira, & Duque, 2014). Exploring this score might be useful to develop short versions of the instrument, or to create stop points based on encoding capacity.

Regarding evidence of concurrent validity based on correlations with other instruments, we adopted the hypothesis that moderate to high correlations would be observed between the pFCSRT-IR and the RAVLT scores, as both instruments are intended to assess episodic memory. This hypothesis was partially corroborated, and significant correlations were identified in the expected direction and size. That is, considering all samples, despite differences in input (images vs. words) and encoding (controlled vs. traditional learning), the correlations indicated that these instruments share a large amount of variance and probably assess the same construct (episodic memory). Another theoretically important result from this analysis is a tendency for RAVLT correlations to be stronger with free recall scores from the pFCSRT-IR than with other scores. This result was particularly predictable, considering the cognitive demands inherent to the response process of these scores. Both are based on free recall, which requires, in addition to storage, mental organization and access to information. The specific analysis of the control group showed that the strategies of information retrieval are particularly important to measure performance in this group since cued recall standardizes the performance obtained in cognitively healthy elderly individuals. That is, not only the presence of correlations, but also their size and relationship with information processing may be considered evidence of validity of the current version. This kind of evidence has been discussed by researchers like Reppold et al. (2015).

It is also noteworthy that the RAVLT score showing the highest correlation with the pFCSRT-IR scores was late recall in all sample. This is probably a consequence of the insertion of distractor tasks between the three pFCSRT-IR trials. This simple distractor task (reverse counting) interferes significantly with short-term (working) memory, minimizing the amount of information retrieved by primacy and recency effects in the pFCSRT-IR. In the RAVLT, the only score that does not suffer interference of primacy and recency effects is late recall, since, in learning trials, free recall is performed immediately after item presentation (Drolet et al., 2014). That is, the absence of recency and primacy effects justifies the high correlation between free recall in the pFCSRT-IR and delayed recall memory in the RAVLT. This effect is particularly important when we consider only the AD group. In this case, patients with AD, when subjected to interference, perform very differently from that obtained in the first RAVLT test, which depends on working memory. This is so important that no significant correlations are identified among these scores. Controlling the effects of working memory during test performance is pointed to as the greatest benefit of the pFCSRT-IR over the RAVLT (Drolet et al., 2014).

The step of obtaining evidence of validity for the pFCSRT-IR has some limitations: for example, use of a small sample. Although the levels of cognitive impairment in the clinical group were controlled, participants with mild cognitive impairment due to AD, mainly those with amnestic subtype, were not included. That is, the results were observed in patients with impairments in their daily life activities detected in previous diagnoses. For future studies, in addition to expanding the sample, the incremental validity of the pFCSRT-IR could be investigated for milder cases, as well as the sensitivity and specificity in other clinical groups and different educational levels, which allows for a differential diagnosis.

The present study presented the process of adaptation and obtained evidence of validity for the pFCSRT-IR. The data observed were compatible with the initial hypotheses, demonstrating that this adaptation shows evidence of validity with external variables (criterion and convergent). It is possible to consider the inclusion of this instrument in the context of neuropsychological assessment in Brazil.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. New York, NY: Author.

American Psychiatric Association. (2014). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Auriacombe, S., Helmer, C., Amieva, H., Berr, C., Dubois, B., & Dartigues, J.-F. (2010). Validity of the Free and Cued Selective Reminding Test in predicting dementia: The 3C study. *Neurology, 74,* 1760–1767. http://dx.doi.org/10.1212/WNL.0b013e3181df0959

Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e validação de instrumentos psicológicos entre culturas: Algumas considerações [Cross-cultural adaptation and validation of psychological instruments: some considerations]. *Paidéia, 22,* 423–432. http://dx.doi.org/10.1590/S0103-863X2012000300014

Buschke, H. (1984). Cued recall in amnesia. *Journal of Clinical Neuropsychology, 6,* 433–440. http://dx.doi.org/10.1080/01688638408401233

Buschke, H. (2002). *Buschke Free and Cued Selective Reminding Test*. New York, NY: Albert Einstein College of Medicine of Yeshiva University.

Buschke, H., & Grober, E. (2002). *Buschke and Grober's FCSRT-IR*. New York, NY: Albert Einstein College of Medicine of Yeshiva University.

Caffarra, P., Ghetti, C., Ruffini, L., Spallazzi, M., Spotti, A., Barocco, F., . . . Gardini, S. (2016). Brain Metabolism Correlates of the Free and Cued Selective Reminding Test in Mild Cognitive Impairment. *Journal of Alzheimer's Disease, 51,* 27–31. http://dx.doi.org/10.3233/JAD-150418

Delgado, C., Muñoz-Neira, C., Soto, A., Martínez, M., Henríquez, F., Flores, P., & Slachevsky, A. (2016). Comparison of the psychometric properties of the "word" and "picture" versions of the free and cued selective reminding test in a Spanish-speaking cohort of patients with mild Alzheimer's disease and cognitively healthy controls. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 31,* 165–175. http://dx.doi.org/10.1093/arclin/acv107

Derby, C. A., Burns, L. C., Wang, C., Katz, M. J., Zimmerman, M. E., L'italien, G., . . . Lipton, R. B. (2013). Screening for predementia AD: Time-dependent operating characteristics of episodic memory tests. *Neurology, 80,* 1307–1314. http://dx.doi.org/10.1212/WNL.0b013e31828ab2c9

Drolet, V., Vallet, G. T., Imbeault, H., Lecomte, S., Limoges, F., Joubert, S., & Rouleau, I. (2014). Comparaison des performances à l'épreuve des 15 mots de Rey et au RL/RI 16 dans le vieillissement normal et la démence de type Alzheimer. *Gériatrie et Psychologie Neuropsychiatrie du Vieillissement, 12,* 218–226.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., . . . Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurology, 6,* 734–746. http://dx.doi.org/10.1016/S1474-4422(07)70178-3

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., . . . Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *The Lancet Neurology, 13,* 614–629. http://dx.doi.org/10.1016/S1474-4422(14)70090-0

Dubois, B., Picard, G., & Sarazin, M. (2009). Early detection of Alzheimer's disease: New diagnostic criteria. *Dialogues in Clinical Neuroscience, 11,* 135–139.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Fonseca, R. P., Casarin, F. S., de Oliveira, C. R., Gindri, G., Ishigaki, E. C. S. S., Ortiz, K. Z., . . . Scherer, L. C. (2011). Adaptação de Instrumentos Neuropsicológicos Verbais: Um Fluxograma de

Procedimentos para Além da Tradução [Adaptation of Neuropsychological Verbal Instruments: A Fluxogram of Procedures far Beyond Translation]. *Interação Em Psicologia, 15,* 59–69.

Frota, N. A. F., Nitrini, R., Damasceno, B. P., Forlenza, O., Dias-Tosta, E., da Silva, A. B., . . . Magaldi, R. M. (2011). Criteria for the diagnosis of Alzheimer's disease [Critérios para o Diagnóstico de Doença de Alzheimer]. *Dementia e Neuropsychologia, 5*(Suppl. 1), 5–10. http://dx.doi.org/10.1590/S1980-57642011DN05030002

Grober, E. (2014). Controlled learning: From the bench to the clinic. In L.-G. Nilsson & N. Ohta (Eds.), *Dementia and memory* (pp. 72–92). New York, NY: Psychology Press.

Grober, E., & Buschke, H. (1987). Genuine memory deficits in dementia. *Developmental Neuropsychology, 3,* 13–36. http://dx.doi.org/10.1080/87565648709540361

Grober, E., Buschke, H., Crystal, H., Bang, S., & Dresner, R. (1988). Screening for dementia by memory testing. *Neurology, 38,* 900–903. http://dx.doi.org/10.1212/WNL.38.6.900

Grober, E., Mowrey, W., Katz, M., Derby, C., & Lipton, R. B. (2015). Conventional and robust norming in identifying preclinical dementia. *Journal of Clinical and Experimental Neuropsychology, 37,* 1098–1106. http://dx.doi.org/10.1080/13803395.2015.1078779

Grober, E., Sanders, A. E., Hall, C., & Lipton, R. B. (2010). Free and cued selective reminding identifies very mild dementia in primary care. *Alzheimer Disease and Associated Disorders, 24,* 284–290. http://dx.doi.org/10.1097/WAD.0b013e3181cfc78b

Johnson, N., Barion, A., Rademaker, A., Rehkemper, G., & Weintraub, S. (2004). The Activities of Daily Living Questionnaire: A validation study in patients with dementia. *Alzheimer Disease and Associated Disorders, 18,* 223–230.

Jongenelis, K., Pot, A. M., Eisses, A. M., Beekman, A. T., Kluiter, H., & Ribbe, M. W. (2004). Prevalence and risk indicators of depression in elderly nursing home patients: The AGED study. *Journal of Affective Disorders, 83,* 135–142. http://dx.doi.org/10.1016/j.jad.2004.06.001

Kochhann, R., Varela, J. S., Lisboa, C. S. M., & Chaves, M. L. F. (2010). The Mini Mental State Examination: Review of cutoff points adjusted for schooling in a large Southern Brazilian sample. *Dementia & Neuropsychologia, 4,* 35–41. http://dx.doi.org/10.1590/S1980-57642010DN40100006

Koen, J. D., & Yonelinas, A. P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A meta-analytic review. *Neuropsychology Review, 24,* 332–354. http://dx.doi.org/10.1007/s11065-014-9266-5

Lemos, R., Martins, C., Simões, M. R., & Santana, I. (2012). Estudo de adaptação do Teste de Recordação Selectiva Livre e Guiada para a população portuguesa [Adaptation study of the Free and Cued Selective Reminding Test to the Portuguese population]. *Avaliação Psicológica, 11,* 49–61.

Lemos, R., Simões, M. R., Santiago, B., & Santana, I. (2015). The Free and Cued Selective Reminding test: Validation for mild cognitive impairment and Alzheimer's disease. *Journal of Neuropsychology, 9,* 242–257. http://dx.doi.org/10.1111/jnp.12048

Malloy-Diniz, L. F., Lasmar, V. A. P., Gazinelli, L. S., Fuentes, D., & Salgado, J. V. (2007). The Rey Auditory-Verbal Learning Test: Applicability for the Brazilian elderly population. *Revista Brasileira de Psiquiatria, 29,* 324–329. http://dx.doi.org/10.1590/S1516-44462006005000053

Marazziti, D., Consoli, G., Picchetti, M., Carlini, M., & Faravelli, L. (2010). Cognitive impairment in major depression. *European Journal of Pharmacology, 626,* 83–86. http://dx.doi.org/10.1016/j.ejphar.2009.08.046

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 & Dementia: The Journal of the Alzheimer's Association, 7,* 263–269. http://dx.doi.org/10.1016/j.jalz.2011.03.005

Medeiros, M., & Guerra, R. (2009). Tradução, adaptação cultural e análise das propriedades psicométricas do Activities of Daily Living Questionnaire (ADLQ) para avaliação funcional de pacientes com a doença de Alzheimer [Translation, cultural adaptation and psychometric analysis of the Activities of Daily Living Questionnaire (ADLQ) for functional assessment of patients with Alzheimer's disease]. *Brazilian Journal of Physical Therapy, 13,* 257–266. http://dx.doi.org/10.1590/S1413-35552009005000027

Mura, T., Proust-Lima, C., Jacqmin-Gadda, H., Akbaraly, T. N., Touchon, J., Dubois, B., & Berr, C. (2014). Measuring cognitive change in subjects with prodromal Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 85,* 363–370. http://dx.doi.org/10.1136/jnnp-2013-305078

Noel-Storr, A. H., Flicker, L., Ritchie, C. W., Nguyen, G. H., Gupta, T., Wood, P., . . . McShane, R. (2013). Systematic review of the body of evidence for the use of biomarkers in the diagnosis of dementia. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 9,* e96–e105. http://dx.doi.org/10.1016/j.jalz.2012.01.014

Oltra-Cucarella, J., Pérez-Elvira, R., & Duque, P. (2014). Beneficios de la codificación profunda en

la enfermedad de Alzheimer. Análisis del rendimiento en una tarea de memoria mediante el item specific deficit approach [Benefits of deep encoding in Alzheimer disease. Analysis of performance on a memory task using the item specific deficit approach]. *Neurologia (Barcelona, Spain), 29,* 286–293. http://dx.doi.org/10.1016/j.nrl.2013.06.006

Perneczky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., & Kurz, A. (2006). Mapping scores onto stages: Mini-mental state examination and clinical dementia rating. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry, 14,* 139–144. http://dx.doi.org/10.1097/01.JGP.0000192478.82189.a8

Petersen, R. C., Smith, G., Kokmen, E., Ivnik, R. J., & Tangalos, E. G. (1992). Memory function in normal aging. *Neurology, 42,* 396–401. http://dx.doi.org/10.1212/WNL.42.2.396

Pilger, C., Menon, M. H., & de Freitas Mathias, T. A. (2011). Características sociodemográficas e de saúde de idosos: Contribuições para os serviços de saúde [Socio-demographic and health characteristics of elderly individuals: Support for health services]. *Revista Latino-Americana de Enfermagem, 19,* 1230–1238.

Reppold, C. T., Gomes, C. M. A., Seabra, A. G., Muniz, M., Valentini, F., & Laros, J. A. (2015). Contribuições da psicometria para os estudos em neuropsicologia cognitiva [Contributions of psychometrics to studies in cognitive neuropsychology]. *Psicologia: Teoria e Prática, 17,* 94–106. http://dx.doi.org/10.15348/1980-6906/psicologia.v17n2p94-106

Rey, A. (1958). *L'examenclinique en psychologie* [The psychological examination]. Paris, France: Presses Universitaires de France.

Sarazin, M., Berr, C., De Rotrou, J., Fabrigoule, C., Pasquier, F., Legrain, S., . . . Dubois, B. (2007). Amnestic syndrome of the medial temporal type identifies prodromal AD: A longitudinal study *Neurology, 69,* 1859–1867.

Slachevsky, A., Barraza, P., Hornberger, M., Muñoz-Neira, C., Flanagan, E., Henríquez, F., . . . Delgado, C. (2018). Neuroanatomical comparison of the "word" and "picture" versions of the Free and Cued Selective Reminding Test in Alzheimer's disease. *Journal of Alzheimer's Disease, 61,* 589–600. http://dx.doi.org/10.3233/JAD-160973

Slachevsky, A., Delgado-Derio, C., Muñoz-Neira, C., Soto, A., Martinez, M., Henríquez, F., & Flores, P. (2014). Discriminative validity of the FCSRT: Verbal and visual versions in a cohort of patients with mild alzheimer disease and controls. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 10,* P432–P433. http://dx.doi.org/10.1016/j.jalz.2014.05.569

Smalbrugge, M., Jongenelis, L., Pot, A. M., Beekman, A. T. F., & Eefsting, J. A. (2005). Comorbidity of depression and anxiety in nursing home patients. *International Journal of Geriatric Psychiatry, 20,* 218–226. http://dx.doi.org/10.1002/gps.1269

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215. http://dx.doi.org/10.1037/0278-7393.6.2.174

Swerdlow, R. H., & Jicha, G. A. (2012). Alzheimer disease: Can the exam predict the pathology? *Neurology, 78,* 374–375. http://dx.doi.org/10.1212/WNL.0b013e318245f48d

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373. http://dx.doi.org/10.1037/h0020071

Wagner, M., Wolf, S., Reischies, F. M., Daerr, M., Wolfsgruber, S., Jessen, F., . . . Wiltfang, J. (2012). Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology, 78,* 379–386. http://dx.doi.org/10.1212/WNL.0b013e318245f447

Wajman, J. R., Oliveira, F. F., Schultz, R. R., Marin, S. M., & Bertolucci, P. H. (2014). Educational bias in the assessment of severe dementia: Brazilian cutoffs for severe Mini-Mental State Examination. *Arquivos de Neuro-Psiquiatria, 72,* 273–277. http://dx.doi.org/10.1590/0004-282X20140002

Zibetti, M. R., Bordignon, S., & Trentini, C. M. (2015). Development and standardization of a Brazilian set of pictorial stimuli. *Revista Avaliação Psicológica, 14,* 9–21. http://dx.doi.org/10.15689/ap.2015.1401.02

Zibetti, M. R., & Trentini, C. M. (2017). *Aprendizagem controlada, evocação livre e guiada na avaliação da memória episódica em pacientes com doença de Alzheimer* [*Controlled learning, free and cued recall in the evaluation of episodic memory in patients with Alzheimer's disease*]. (Unpublished doctoral dissertation). Universidade Federal do Rio Grande do Sul, Brazil.

Zimmerman, M. E., Katz, M. J., Wang, C., Burns, L. C., Berman, R. M., Derby, C. A., . . . Lipton, R. B. (2015). Comparison of "word" vs. "picture" version of the Free and Cued Selective Reminding Test (FCSRT) in older adults. *Alzheimer's & Dementia (Amsterdam, Netherlands), 1,* 94–100. http://dx.doi.org/10.1016/j.dadm.2014.11.006

Received February 5, 2018
Revision received September 17, 2018
Accepted October 1, 2018 ■