

# Specific algorithm method of scoring the Clock Drawing Test applied in cognitively normal elderly

Liana Chaves Mendes-Santos<sup>1</sup>, Daniel Mograbi<sup>1,2</sup>, Bárbara Spenciere<sup>3</sup>, Helenice Charchat-Fichman<sup>1</sup>

**ABSTRACT.** The Clock Drawing Test (CDT) is an inexpensive, fast and easily administered measure of cognitive function, especially in the elderly. This instrument is a popular clinical tool widely used in screening for cognitive disorders and dementia. The CDT can be applied in different ways and scoring procedures also vary. **Objective:** The aims of this study were to analyze the performance of elderly on the CDT and evaluate inter-rater reliability of the CDT scored by using a specific algorithm method adapted from Sunderland et al. (1989). **Methods:** We analyzed the CDT of 100 cognitively normal elderly aged 60 years or older. The CDT (“free-drawn”) and Mini-Mental State Examination (MMSE) were administered to all participants. Six independent examiners scored the CDT of 30 participants to evaluate inter-rater reliability. **Results and Conclusion:** A score of 5 on the proposed algorithm (“Numbers in reverse order or concentrated”), equivalent to 5 points on the original Sunderland scale, was the most frequent (53.5%). The CDT specific algorithm method used had high inter-rater reliability ( $p < 0.01$ ), and mean score ranged from 5.06 to 5.96. The high frequency of an overall score of 5 points may suggest the need to create more nuanced evaluation criteria, which are sensitive to differences in levels of impairment in visuoconstructive and executive abilities during aging.

**Key words:** Clock Drawing Test, inter-rater reliability, elderly, neuropsychology.

## MÉTODO COM ALGORITMO ESPECÍFICO PARA PONTUAÇÃO DO TESTE DO DESENHO DO RELÓGIO EM IDOSOS COGNITIVAMENTE NORMAIS

**RESUMO.** O Teste do Desenho do Relógio (TDR) é uma barata e rápida medida de função cognitiva, de fácil aplicação, especialmente em idosos. Este instrumento é uma ferramenta clínica muito conhecida, amplamente utilizada no rastreamento de transtornos cognitivos e demência. O TDR pode ser aplicado de diferentes formas e a sua pontuação também varia. **Objetivo:** Os objetivos deste estudo foram analisar o desempenho dos idosos no TDR e avaliar a confiabilidade inter-examinadores do TDR pontuado por um método com algoritmo específico, adaptado a partir dos critérios estabelecidos por Sunderland et al. (1989). **Métodos:** Analisamos o TDR de 100 idosos cognitivamente saudáveis com 60 anos de idade ou mais. O TDR (“desenho livre”) e o Mini-Exame do Estado Mental (MEEM) foram administrados em todos os participantes. Seis avaliadores independentes pontuaram 30 TDR para avaliar a confiabilidade inter-examinadores. **Resultados e Conclusão:** A pontuação 5 do algoritmo proposto (“Os números em ordem inversa ou concentrados”) equivalente a 5 pontos na escala original de Sunderland foi a mais frequente (53,5%). O método com algoritmo específico do TDR utilizado teve alta confiabilidade entre avaliadores ( $p < 0,01$ ), e a média da pontuação variou entre 5,06 e 5,96. A alta frequência de 5 pontos na pontuação geral pode sugerir a necessidade da elaboração de critérios de avaliação mais sutis, que sejam sensíveis às diferenças entre indícios de comprometimento nas habilidades visuoconstrutivas e executivas durante o envelhecimento. **Palavras-chave:** Teste do Desenho do Relógio, confiabilidade inter-examinadores, idosos, neuropsicologia.

## INTRODUCTION

The Clock Drawing Test (CDT) is a simple and ecological neuropsychological instru-

ment that covers a wide range of cognitive functions, including selective and sustained attention, auditory comprehension, verbal

This study was conducted at the Department of Psychology, Pontifical Catholic University of Rio de Janeiro, Brazil.

<sup>1</sup>PhD, Department of Psychology, Pontifical Catholic University of Rio de Janeiro RJ, Brazil. <sup>2</sup>Department of Psychology, Institute of Psychiatry, King's College London, UK. <sup>3</sup>BSc, Department of Psychology, Pontifical Catholic University of Rio de Janeiro RJ, Brazil.

**Helenice Charchat-Fichmann.** Rua Marquês de São Vicente 255 / 2º andar – Departamento de Psicologia – 22543-900 Rio de Janeiro RJ – Brasil. E-mail: hcfichman@puc-rio.br; hcharchat@uol.com.br

Disclosure: The authors report no conflicts of interest.

Received February 12, 2015. Accepted in final form April 25, 2015.

working memory, numerical knowledge, visual memory and reconstruction, visuospatial abilities, on-demand motor execution (praxis) and executive function.<sup>1</sup> Some studies have demonstrated the robust psychometric properties of the CDT.<sup>2-4</sup>

The CDT has been used as a cognitive screening tool, particularly in the elderly population, to differentiate cognitively normal individuals from individuals with cognitive impairment and dementia.<sup>5-7</sup> This test has the capacity to evaluate multi-domain impairments that may go undetected by other cognitive screening instruments, such as the Mini-Mental State Examination (MMSE).<sup>2,8</sup> The relative independence of verbal abilities<sup>9,10</sup> makes it especially useful in patients presenting marked verbal impairment or aphasia. In addition, the CDT has also shown strong associations with other cognitive measures, such as the Cambridge Cognitive Examination (CAMCOG),<sup>6,11,12</sup> justifying the inclusion of the CDT in several neuropsychological cognitive screening batteries.<sup>1,10,12</sup>

Although there is great interest in CDT as a screening instrument, a wide variety of CDTs have been developed, each relying on different systems of administration and quantitative or qualitative error scoring, with no consensus on which system produces the most valid results.<sup>3,5,13</sup> The currently used methods are Shulman et al.,<sup>14</sup> Sunderland et al.<sup>10</sup> and Mendez et al.<sup>1-3,15</sup> These different applications and systems of scoring are somewhat similar and highly correlated in some aspects, but their diagnostic accuracy, and the cognitive processes involved in their performance are different.<sup>16</sup>

CDT performance is associated with several brain regions, including the bilateral parietal lobes, right and left posterior and middle temporal lobes, right middle frontal gyrus, and the right occipital lobe.<sup>16,17</sup> These areas can also be associated with a broad spectrum of pathologies. A number of studies have shown the potential of the CDT for investigating cognitive performance in patients with schizophrenia, Alzheimer's disease, Parkinson's disease, depression and other disorders.<sup>9,18,19</sup>

Previous studies have investigated the test-retest reliability,<sup>1,9</sup> and determined inter-rater reliability, of the CDT.<sup>6,10,20-24</sup> These studies compared the different application systems and showed that the systems were well correlated, took little time and had high inter-rater reliability. On the other hand, CDT reliability has rarely been assessed in population-based studies, particularly in developing countries. Three studies determining inter-rater reliabilities of the CDT in elderly in Brazil were found: one scored the tests with Shulman's method,<sup>20</sup> while the others compared the accuracy of scales (Men-

dez, Shulman and Sunderland;<sup>6</sup> Sunderland, Shulman, Manos & Wu and Wolf-Klein<sup>24</sup>) and determined the inter-rater reliability of CDT performance. These investigations showed good inter-rater reliabilities.

One of the most used methods of CDT scoring is Sunderland et al.<sup>10</sup> This method of scoring is well established in the literature<sup>10,25-27</sup> and widely used in Brazil, being part of cognitive screening batteries for the elderly.<sup>28,29</sup>

With the aim of providing a more detailed, specific and quantitative analysis of the different aspects of CDT assessment, the present study proposed an algorithm method for scoring the CDT adapted from Sunderland et al.<sup>10</sup> To this end, the performance of 100 elderly was analyzed using the new algorithm, and its inter-rater reliability was evaluated.

## METHODS

**Participants.** The sample was part of a larger study involving 350 elderly from community centers, known as "Casas de Convivência", belonging to the Rio de Janeiro municipal administration. One hundred elderly took part in this study (93 females and seven males). The inclusion criteria were: [1] to be literate (a person who can read and write; mean=9.8 years of education, SD=4.2), [2] to be aged 60 years or older (mean age=72.6 years old, SD=5.9), and [3] to be cognitively healthy (MMSE mean score=25.3, SD=3.4). Cut-off scores for the MMSE were defined according to educational level. MMSE scores range from 0 to 30, with higher scores indicating better cognitive function; the cut-off for cognitive impairment was 18 in individuals with fewer than four years of formal education and 24 for participants with more than four years of education.<sup>8,30</sup> Exclusion criteria were: to be visually and/or hearing impaired or have uncorrected deficits, presence of endocrine and metabolic abnormalities, neurological and psychiatric disorders, or difficulty executing hand movements due to rheumatic diseases.

Before entry to the study all participants received an explanation on the objectives of the research, and signed an informed consent form. The Research Ethics Committee of the State University of Rio de Janeiro approved this study.

**Materials and procedures.** Subjects were first submitted to a standardized questionnaire, which collected data on sociodemographic variables (i.e., gender, age and education), on subjective memory impairment (i.e., "Do you feel like your memory has gotten worse?"), and on comorbidities. All participants then completed the same

protocol of cognitive screening tests. The tests were applied in the following sequence (based on Nitrini et al.<sup>31</sup>): [1] MMSE;<sup>8,30</sup> [2] Memory Test Figures;<sup>31</sup> [3] Verbal Fluency – Animals;<sup>9,32</sup> [4] CDT (described below); [5] The Lawton Instrumental Activities of Daily Living<sup>33,34</sup> (for further details see Charchat-Fichman et al.<sup>35</sup>). Besides the cognitive and functional evaluations, participants completed the Geriatric Depression Scale (GDS).<sup>36</sup>

The CDT was applied to all participants in the spontaneous modality that uses a pencil and blank sheet of paper. The patients were asked to draw a clock without a model. Trained examiners issued a standardized instruction: “Draw a clock, put in all the numbers, and set the hands to 2 hours and 45 minutes.” There was no time limit.

Table 1 shows the original CDT scoring scale by Sunderland et al.,<sup>10</sup> which forms the basis of the new algorithm (Table 2). Both Tables 1 and 2 present the correspondence of higher scores indicating better performance. Examples of the CDT scoring scale by Sunderland et al.<sup>10</sup> are given in Figure 1. According to the new algorithm (Table 2), the examiner must first mark with an “X” all the items present in the clock drawing. The list of items has increasing complexity.

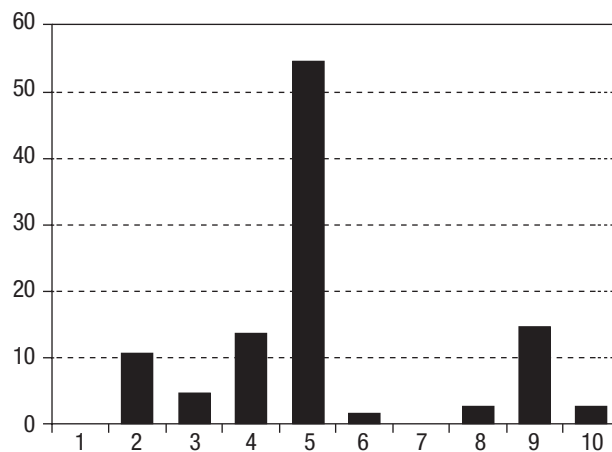
Inter-rater reliability was assessed by comparing CDT scores from six independent examiners, who each scored the same 30 clocks from subjects sampled randomly.

## RESULTS

A summary of the participants’ sociodemographic characteristics, performance on cognitive screening tests, as well as cognitive function and depression scales is given in Table 3. Table 4 shows performance on the CDT.



**Figure 1.** Examples of CDT score in accordance with the specific algorithm method based on Sunderland et al.<sup>10</sup>: 9, 5 and 2 (right to left), respectively.



**Figure 2.** Histogram showing the frequency of CDT scores according to the scoring system developed by Sunderland et al.<sup>10</sup>.

According to the histogram shown in Figure 2, regarding the performance of the elderly on the CDT, the frequency of score “5” was 53.5%, and scores “1” and “7” were not present in the current sample.

Pearson’s correlation was used to evaluate the relationship between schooling, age and MMSE with CDT scores. No significant correlation was found between schooling and CDT ( $r=0.014$ ,  $p>0.05$ ) or age and CDT

**Table 1.** The original Sunderland method for scoring the CDT.<sup>10</sup>

10-6	Drawing of clock face with circle and numbers is generally intact.
10	Hands are in correct position.
9	Slight errors in placement of hands.
8	More noticeable errors in placement of hour and minute hands.
7	Placement of hands is significantly off course.
6	Inappropriate use of clock hands (i.e., use of digital display or circling of numbers despite repeated instructions).
5-1	Drawing of a clock face with circle and numbers is not intact.
5	Crowding of numbers at one end of the clock or reversal of numbers. Hands may still be present in some fashion.
4	Further distortion of number sequence. Integrity of clock face is now gone (i.e., numbers missing or placed at outside of the boundaries of the clock face).
3	Numbers and clock face no longer obviously connected in drawing. Hands are not present.
2	Drawing reveals some evidence of instructions being received but only a vague representation of a clock.
1	Either no attempt or an uninterpretable effort is made.

**Table 2.** New algorithm method for CDT scoring based on the original criteria of Sunderland et al.<sup>10</sup>

<b>You should mark with an "X" all the items present in the clock drawn</b>	
(a) Presence of circle.	(j) Presence of hour hand.
(b) Presence of 12 numbers.	(k) Presence of minute hand.
(c) Numbers entered in the internal limit of the clock.	(l) Minute hand proportionally larger than the hour hand.
(d) Number in the correct ascending order.	(m) One of the hands between 2 and 3.
(e) Numbers in correct spatial position.	(n) One of the hands on exactly 9.
(f) Can you draw a straight vertical line between 12 and 6.	(o) Wrong use of hands (digital or circling the numbers).
(g) Can you draw a straight horizontal line between 3 and 9.	(p) Some evidence of having understood that it is a clock.
(h) Numbers not concentrated in one part of the clock.	(q) Did not try or did not represent a clock.
(i) Presence of two pointers.	
<b>Follow the algorithm for the score, but consider these three points initially</b>	
1. If the item "o" is checked, the score is 6 points.	
2. If the item "p" is checked, the score is 2 points.	
3. If the item "q" is checked, the score is 1 point.	
<b>The score will be 10-6 if the clock and the numbers are drawn correctly</b>	
10	Correct time (no "X" in the items: "o", "p", "q").
9	Very mild disorder of hands (absence of "X" in at least one item: "l", "m" or "n").
8	Mild disorder of hands (absence of "X" in at least 2 items: "l", "m", "n").
7	Severe disorder of hands (absence of "X" in the items: "l", "m", "n").
6	Wrong use of hands (presence of "X" in item "o").
<b>The score will be 5-1 if the drawing of the clock and the numbers are incorrect</b>	
5	Numbers in reverse order or concentrated (no "X" in the items: "d" or "h").
4	Numbers missing and located outside the boundary of the clock (no "X" in items: "b" and "c").
3	Absence of hands (no "X" in the items: "i", "j", "k").
2	Only some evidence of having understood that it is a clock (presence of "X" in item p).
1	Not tried or did not represent a clock (presence of "X" item in q).

**Table 3.** Participants' sociodemographic characteristics, and performance on cognitive screening tests, as well as cognitive function and depression scales.

<b>Sociodemographic characteristics</b>	<b>Mean</b>	<b>SD*</b>	<b>Minimum value</b>	<b>Maximum value</b>	
Age	72.6	5.9	60	84	
Years of education	9.8	4.2	3	24	
Instruments and scales	MMSE (Memory Figures Test)	25.2	3.3	18	30
	• Incidental Memory	25.4	1.1	2	8
	• Immediate Memory 1	7.9	1.3	4	10
	• Immediate Memory 2	8.6	1.1	5	10
	• 5 Minutes - Delayed Memory	7.7	1.5	4	10
• Recognition	9.9	0.3	8	10	
Verbal Fluency	15	4.8	5	27	
Lawton's Scale	20.1	1.4	18	21	
GDS	1.9	2.1	0	8	

\*SD: standard deviation.

**Table 4.** Participants' performance on CDT: mean, median, standard deviation, minimum and maximum score.

<b>N</b>	<b>Mean</b>	<b>Median</b>	<b>Standard deviation</b>	<b>Minimum score</b>	<b>Maximum score</b>
100	5.22	5	2.02	2	10

( $r=0.04$ ,  $p>0.05$ ), but a significant positive correlation was found between MMSE and CDT ( $r=0.22$ ,  $p<0.05$ ).

On the other hand, the investigation of inter-rater reliability of the CDT, scored by criteria based on Sunderland et al.,<sup>10</sup> showed that the mean ranged from 5.06 to 5.96 (Table 5).

Pearson's correlation analysis was performed between the scores found by the six independent raters for 30 tests. A significant positive correlation was found between the examiners ( $p<0.01$ ): 1 and 2 ( $r=0.79$ ), 1 and 3 ( $r=0.7$ ); 1 and 4 ( $r=0.75$ ); 1 and 5 ( $r=0.84$ ), 1 and 6 ( $r=0.71$ ), 2 and 3 ( $r=0.87$ ), 2 and 4 ( $r=0.79$ ), 2 and 5 ( $r=0.79$ ), 2 and 6 ( $r=0.79$ ), 3 and 4 ( $r=0.79$ ), 3 and 5 ( $r=0.69$ ), 3 and 6 ( $r=0.8$ ), 4 and 5 ( $r=0.79$ ), 4 and 6 ( $r=0.88$ ) and 5 and 6 ( $r=0.74$ ).

The agreement between raters was high, consistently remaining statistically significantly above expected chance agreement (in all cases,  $p<0.001$ ). The combined kappa for all six examiners was 0.60, with pairwise analyses between evaluators indicating an average level of agreement of 90.2% and an average weighted kappa of 0.69.

## DISCUSSION

The current study analysed the performance of a cognitively normal elderly community sample on the CDT using a specific algorithm score method based on the Sunderland et al.<sup>10</sup> system. The mean score of participants was 5.22, and the standard deviation 2.02. The score 5 ("Numbers in reverse order or concentrated") was observed in 53.5% of clock drawings.

In general, studies with the CDT compare the performance of patients and controls in different applications and scoring systems<sup>2,25,37</sup> or verify the clinical validity of the test,<sup>21,23,38</sup> or investigate the translation and adaptation of the CDT model for a particular population.<sup>39,40</sup> There are few studies in community-dwelling samples or cognitively normal elderly.<sup>22,41-43</sup>

Five Brazilian studies using Sunderland's scoring method found higher scores than the present study (5.22, and standard deviation 2.02). Two of these studies did not mention CDT scores,<sup>6,24</sup> while the other results were: 9.7 ( $\pm 1.07$ ),<sup>41</sup> 7.8 ( $\pm 2.2$ ),<sup>28</sup> and 9.1 ( $\pm 1.8$ ).<sup>11</sup> However, comparison of the current findings with results of these studies is hampered because of a number of differences in study design. The most important difference was related to the intrinsic characteristics of the sample. The cited studies used small clinical samples recruited in hospital settings, in contrast to the present study which used a large sample of normal elderly from community centers with heterogeneous age and edu-

**Table 5.** Mean and SD of CDT scores rated by the six examiners.

Examiners	Mean	SD
1	5.06	2.24
2	5.66	2.57
3	5.96	2.74
4	5.73	2.55
5	5.23	1.95
6	5.6	2.71

SD: standard deviation.

cational levels.<sup>6,11,24,28,31,41</sup> The objectives of the studies also varied. Some compared different methods of CDT scoring,<sup>24,41</sup> others compared the instrument with other tests and finally there was a study that evaluated the profile of the elderly subjects on the CDT<sup>28</sup> based on a selected group of normal elderly as a control group compared to Alzheimer's disease patients.

Studies in the international literature that used the same method as Sunderland to score the CDT found the following results: 7.5 ( $\pm 1.9$ ),<sup>25</sup> 8.4 ( $\pm 1.6$ ),<sup>27</sup> 8.7 ( $\pm 1.1$ ),<sup>10</sup> and 8.9 ( $\pm 1.4$ ).<sup>26</sup> Similar to the Brazilian studies, all of these found higher scores for normal elderly individuals<sup>10,26,27</sup> than in the present study, except Kirby et al.<sup>25</sup> who found lower scores compared to the other international studies. Some studies failed to mention all important information, for example, the educational level<sup>10,26</sup> or did not use formal cognitive testing for normal controls<sup>10</sup> (including the MMSE<sup>10,27</sup>) while another did not separate the clinical group when describing the sample characteristics,<sup>27</sup> hindering comparisons among the studies. The aim of the present study differs from the main objective of the previous studies in that its aim was to evaluate the performance of the elderly with and without cognitive impairment.<sup>10,25-27</sup>

An important outcome regarding the performance of the elderly is the high percentage (53.5%) of the sample with scores of "5". The criterion for a score of "5" in Sunderland's original method is "Crowding of numbers at one end of the clock or reversal of numbers. Clock Hands may still be present in some fashion" and in the new algorithm denoted: "Numbers in reverse order or concentrated". The lower mean scores on the CDT compared to other studies, and the high frequency of elderly that scored at this level could be explained by the fact that strict correction was used to score the CDT in this study. Sunderland's method in its original version had a more subjective approach, for example, very high CDT scores, even with numbers slightly concentrated, could be found in Sunderland et al.<sup>10</sup> (Figure 1, p. 727). Accord-

ing to Sunderland's method, item 5 should be scored only when there is a drastic concentration, and in the present research this item included people with slight and severe difficulty in planning. Thus, when strict criteria are used, different results are obtained compared to the literature.

In this sense, it would be necessary to develop more specific scoring criteria that may be sensitive to planning strategy and visual-constructive execution of the CDT, and which could better differentiate specifically those elderly with possible executive dysfunction. Other methods of scoring the CDT, including semi-quantitative and qualitative scoring systems, attempt to discriminate the level of executive planning in clock drawings,<sup>42,44,45</sup> and emphasize the evaluation of executive components in CDT.<sup>42-44</sup> For example, Royall et al.<sup>45</sup> developed the Executive Clock Drawing Task (CLOX) in order to discriminate these components and allow a more specific analysis of how the executive functions can be tested in the CDT.

No significant correlation was found between education or aging and CDT scores. The relationship between education, aging and CDT performance is controversial in the literature.<sup>22,24,38,41,43</sup> This finding may also be related to the existence of various application methods and different scoring scales. For example, Brodaty and Moore found a correlation of CDT score with years of education for the Shulman and Sunderland, but not for the Wolf-Klein scoring system.<sup>2</sup> Sunderland et al.<sup>10</sup> did not report the educational level of control subjects in the original study.

On the other hand, a significant positive correlation was found between the CDT and MMSE, confirming previous findings.<sup>6,7,15</sup> A high correlation has been found for the scales of Shulman,<sup>14</sup> Mendez<sup>1</sup> and the CLOX scale.<sup>45</sup> The association between MMSE score and several CDTs was also high in the study by Schramm et al.<sup>7</sup>

These various systems of application and scoring are an obstacle to establishing direct comparisons and drawing conclusions. The different forms of application include differences in the clock time requested (2:45, 11:10, 8:05) and presence of drawing assistance (e.g. some have a pre-drawn circle). In addition, the various scoring systems include: 10 hierarchical patterns (0-10), scale based on errors each scored 0/1 (0-20), clock divided into eighths, points given for numbers and hands in right place (0-10) and others.<sup>3,14,37,43</sup>

In this study, an algorithm with more specific scores based on Sunderland et al.<sup>10</sup> criteria was devised to increase inter-rater reliability. The examination of the inter-rater reliability showed that the criteria developed

for the present study were reliable and a significant positive correlation was found between the six independent examiners. These results are similar to those found in previous studies, also indicating high inter-rater reliability of CDT scores.<sup>10,21-23</sup> Again, the various ways of presenting the test and the different principles involved in scoring, make comparisons difficult. Another aspect that hampers comparisons is the use of several different study designs. Some studies examined inter-rater reliabilities of the CDT scored by one scoring system in cognitively normal elderly<sup>20</sup> or in differentiating between cognitively normal and individuals with different types of pathologies,<sup>2</sup> while others examined inter-rater reliability using different scoring systems among cognitively normal elderly<sup>22,37</sup> or cognitively normal and individuals with different types of diseases.<sup>21</sup> Two other studies that evaluated the inter-rater reliability using various score systems, including the method of Sunderland et al.,<sup>10</sup> compared subjects with and without pathologies (fibromyalgia and mild cognitive impairment, MCI)<sup>37,46</sup> and showed good inter-rater reliability.

The idea of systematic scoring of the CDT has focused on the development and standardization of simple and easy-to-interpret scoring methods.<sup>21,22</sup> There are two general CDT scoring approaches, including qualitative and quantitative approaches. The Sunderland et al.<sup>10</sup> is a semi-quantitative scoring system that focuses on scoring the whole clock.<sup>37</sup> Other quantitative scoring systems focus on different aspects of the clocks (such as clock face, numbers or hands) and score them separately (i.e., the Clock Drawing Interpretation Scale by Mendez et al.<sup>1</sup> and Rouleau et al.<sup>12</sup>). Furthermore, the scoring systems differ regarding scoring procedures.

One limitation of this study is the non-stratification of participants by age for comparison. Perhaps the advanced age of some participants may have influenced the low average scores. Another question to be considered centers on the intrinsic characteristics of the sample and on the volunteers that participated in the activities of the Casas de Convivência. For example, the sample comprises mostly women (93%), with few health conditions. However, considering this is a convenience sample, it was not possible to limit recruitment on the basis of personal characteristics. In addition, other Brazilian studies also feature a higher percentage of women,<sup>11,20,24,41</sup> making it unlikely that this represents a major bias in results. These subjects were normal elderly (criterion for inclusion in the sample was to score above the cut-off point on the MMSE), but some older adults with MCI might have been included in the sample; a number of conditions associated with aging

could be present, and some comorbidities not directly related with cognition may have influenced the results. Another limitation to be considered is associated with the method of sample selection. To adequately address selection bias, a randomized sample would have been better than the convenience sample used in the present study. Moreover, other limitations were the absence of other measures of executive functions to compare with the CDT and no functional literacy examination.

The present findings represent an important contribution to the discussion on which CDT administration and scoring system produces the most valid results. The results confirmed the consistency of the scoring criteria of Sunderland et al.<sup>10</sup>. Furthermore, the findings contribute to the discussion about the lack of consensus on the different scoring criteria developed for the CDT and on which would produce more valid results. On the other hand, they may further suggest the need for creating more subtle evaluation criteria, which are sensitive to the differences between impairment in visuoconstructive and executive abilities during aging.

Future research should replicate these findings in elderly with higher and lower formal education to com-

pare the impact of educational level on the CDT. Additional studies could explore more qualitative aspects of the CDT, including strategies implemented, as well as comparing it to other scoring criteria, and clinical validation in the case of Alzheimer's disease, MCI and depression.

**Author contributions.** Liana Chaves Mendes-Santos wrote the paper and participated in statistical analysis. Daniel Mograbi participated in statistical analysis and revised the paper. Bárbara Spenciere wrote the paper. Helenice Charchat-Fichman designed the study, wrote the paper, participated in statistical analysis and revised the paper.

**Grant Support.** CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), FAPERJ (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Rio de Janeiro).

**Acknowledgement.** Christina Martins Borges Lima, Ana Lara Soares Blum Malak, Amanda Buhler Ricciari, Marina Zaitune Baumgratz Lopes Bueno, Maria Fernanda Fernandes de Castro Barbosa, Eduarda Naidel Barbosa e Barbosa.

## REFERENCES

- Mendez MF, Ala T, Underwood KL. Development of scoring criteria for the clock drawing task in Alzheimer's disease. *J Am Geriatr Soc* 1992; 40:1095-1099.
- Brodsky H, Moore CM. The clock drawing test for dementia of the Alzheimer's type: a comparison of three scoring methods in a memory disorders clinic. *Int J Geriatr Psychiatry* 1997;12:619-627.
- Shulman, KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* 2000;15:548-561.
- Todd ME, Dammers PM, Adams SG Jr, Todd HM, Morrison M. An examination of a proposed scoring procedure for the clock drawing test reliability and predictive validity of the clock scoring system. *Am J Alzheimer's Dis* 1995;4:22-26.
- Aprahamian I, Martinelli JE, Neri AL, Yassuda MS. The Clock Drawing Test: a review of its accuracy in screening for dementia. *Dement Neuropsychol* 2009;3:74-80.
- Aprahamian I, Martinelli JE, Neri AL, Yassuda MS. The accuracy of the Clock Drawing Test compared to that of standard screening tests for Alzheimer's disease: results from a study of Brazilian elderly with heterogeneous educational backgrounds. *Int Psychogeriatr* 2010;22:64-71.
- Schramm U, Berger G, Muller R, Kratzsch T, Peters J, Frolich L. Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population. *Int J Geriatr Psychiatry* 2002;17:254-260.
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the mental state of patients for clinician. *J Psychiatr Res* 1975;12:189-198.
- Strauss E, Sherman EMS, Spreen O. *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd ed.). New York: Oxford University Press; 2006.
- Sunderland T, Hill JL, Mellow AM, et al. Clock drawing in Alzheimer's disease: a novel measure of dementia severity. *J Am Geriatr Soc* 1989; 37:725-729.
- Nunes PV, Diniz BS, Radanovic M, et al. CAMCOG as a screening tool for diagnosis of mild cognitive impairment and dementia in a Brazilian clinical sample of moderate to high education. *Int J Geriatr Psychiatry* 2008;23:1127-1133.
- Rouleau I, Salmon DP, Butters N, Kennedy KC, McGuire K. Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. *Brain Cogn* 1992;18:70-87.
- Mainland BJ, Amodeo S, Shulman KI. Multiple clock drawing scoring systems: simpler is better. *Int J Geriatr Psychiatry* 2014;29:127-136.
- Shulman KI, Gold DP, Cohen CA, Zuccheri CA. Clock-drawing and dementia in the community: a longitudinal study. *Int J Geriatr Psychiatry* 1993;8:487-496.
- Richardson HE, Glass JN. A comparison of scoring protocols on the clock drawing test in relation to ease of use, diagnostic group, and correlations with Mini-Mental State Examination. *J Am Geriatr Soc* 2002;50:169-173.
- Matsuoka T, Narumoto J, Okamura A, et al. Neural correlates of the components of the clock drawing test. *Int Psychogeriatr* 2013;16:1-7.
- Ino T, Asada T, Ito J, Kimura T, Fukuyama H. Parieto-frontal networks for clock drawing revealed with fMRI. *Neurosci Res* 2003;45:71-77.
- Royall DR, Palmer RF, Chiodo LK, Polk MJ, Markides KS, Hazuda H. Clock-drawing potentially mediates the effect of depression on mortality: replication in three cohorts. *Int J Geriatr Psychiatry* 2008;23:821-829.
- Saur R, Maier C, Milian M, et al. Clock test deficits related to the global cognitive state in Alzheimer's and Parkinson's disease. *Dement Geriatr Cogn Disord* 2012;33:59-72.
- Fuzikawa C, Lima-Costa MF, Uchoa E, Barreto SM, Shulman KI. A population based study on the intra and inter-rater reliability of the clock drawing test in Brazil: the Bambuí Health and Ageing Study. *Int J Geriatr Psychiatry* 2003;18:450-456.
- García-Caballero A, Recimil MJ, García-Lado I, et al. ACE clock scoring: A comparison with eight standard correction methods in a population of low educational level. *J Geriatr Psychiatry Neurol* 2006;19:216-219.
- Hubbard EJ, Santini V, Blankevoort CG, et al. Clock drawing performance in cognitively normal elderly. *Arch Clin Neuropsychol* 2008;23: 295-327.
- Nair AK, Gavett BE, Damman M, et al. Clock Drawing Test Ratings by Dementia Specialists: Interrater Reliability and Diagnostic Accuracy. *J Neuropsychiatry Clin Neurosci* 2010;22:85-92.
- Lourenço RA, Ribeiro-Filho ST, Moreira IFH, Paradelo EMP, Miranda AS.

- The Clock Drawing Test: performance among elderly with low educational level. *Rev Bras Psiquiatr* 2008;30:309-315.
25. Kirby M, Denihan A, Bruce I, Coakley D, Lawlor BA. The Clock Drawing Test in primary care: sensitivity in dementia detection and specificity against normal and depressed elderly. *Int J Geriatr Psychiatry* 2001; 16:935-940.
  26. Lee H, Swanwick GRJ, Coen RF, Lawlor BA. Use of the clock drawing task in the diagnosis of mild and very mild Alzheimer's disease. *Int Psychogeriatr* 1996;8:469-476.
  27. Powlishta KK, Von Dras DD, Stanford A, et al. The Clock Drawing Test is a poor screen for very mild dementia. *Neurology* 2002;59:898-903.
  28. Hamdan AC, Hamdan EMLR. Teste do desenho do relógio: desempenho de idosos com doença de Alzheimer. *RBCEH* 2009;6:98-105.
  29. Hollveg P, Hamdan AC. Avaliação neuropsicológica em idosos. *RBCEH* 2008;5:110-123.
  30. Brucki SMD, Nitrini R, Caramelli P, Bertolucci PHF, Okamoto IH. Suggestions for utilization of the Mini-Mental State Examination in Brazil. *Arq Neuropsiquiatr* 2003;61:777-781.
  31. Nitrini R, Lefevre BH, Mathias SC, et al. Neuropsychological tests of simple application for diagnosing dementia. *Arq Neuropsiquiatr* 1994; 52:457-465.
  32. Mitrushina MN, Boone KB, D'Elia LF. *Handbook of Normative Data for Neuropsychological Assessment*. New York: Oxford University Press; 1999.
  33. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179-186.
  34. Santos RL, Virtuoso Júnior JS. Confiabilidade da versão brasileira da Escala de Atividades Instrumentais da Vida Diária. *RBPS* 2008;21: 290-296.
  35. Charchat-Fichman H, Fernandes C, Oliveira R, Caramelli P, Aguiar D, Novaes R. Predomínio de Comprometimento Cognitivo Leve Disexecutivo em idosos atendidos no ambulatório de geriatria de um hospital terciário na cidade do Rio de Janeiro. *Rev Neuropsi Latinoam* 2013; 5:31-40.
  36. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol* 1986;5: 165-173.
  37. Ehreke L, Luck T, Luppa M, König H, Villringer A, Riedel-Heller SG. Clock Drawing Test – screening utility for mild cognitive impairment according to different scoring systems: results of the Leipzig Longitudinal Study of the Aged (LEILA 75+). *Int Psychogeriatr* 2011;23:1592-1601.
  38. Bozikas VP, Giakoulidou A, Hatzigeorgiadou M, Karavatos A, Kosmidis MH. Do age and education contribute to performance on the clock drawing test? Normative data for the Greek population. *J Clin Exp Neuropsychol* 2008;30:199-203.
  39. Atalaia-Silva KC, Lourenço RA. Tradução, adaptação e validação de construto do Teste do Relógio aplicado entre idosos no Brasil. *Rev Saúde Pública* 2008;42:930-937.
  40. Royall D. The "Alzheimerization" of dementia research. *J Am Geriatr Soc* 2003;51:277-278.
  41. Cecato JF, Fiorese B, Montiel JM, Bartholomeu D, Martinelli JE. Clock Drawing Test in Elderly Individuals With Different Education Levels Correlation With Clinical Dementia Rating. *Am J Alzheimers Dis Other Dement* 2012;27:620-624.
  42. Crowe M, Allman RM, Triebel K, Sawyer P, Martin RC. Normative Performance on an Executive Clock Drawing Task (CLOX) in a Community-Dwelling Sample of Older Adults. *Arch Clin Neuropsychol* 2010;25: 610-617.
  43. Menon C, Hall J, Hobson V, Johnson L, O'Bryant SE. Normative performance on the executive clock drawing task in a multi-ethnic bilingual cohort: A project FRONTIER study. *Int J Geriatr Psychiatry* 2012;27: 959-966.
  44. Juby A, Tench S, Baker V. The value of clock drawing in identifying executive cognitive dysfunction in people with normal Mini-Mental State Examination Score. *CMAJ* 2002;167:859-864.
  45. Royall DR, Cordes JA, Polk M. CLOX: an executive clock drawing task. *J Neurol Neurosurg Psychiatry* 1998;64:588-594.
  46. Can SS, Gencay-Can A, Gunendi Z. Validity and reliability of the clock drawing test as a screening tool for cognitive impairment in patients with fibromyalgia. *Compr Psychiatry* 2012;53:81-86.