Mapping the clockworks: what does the Clock Drawing Test assess in normal and pathological aging?

Mapeando as engrenagens: o que o Teste do Desenho do Relógio avalia no envelhecimento normal e patológico?

Jonas Jardim de Paula1,2, Débora Marques de Miranda2, Edgar Nunes de Moraes3, Leandro Fernandes Malloy-Diniz1,2,4

Brief cognitive tests are widely used in neuropsychology and related areas, in order to assess and diagnose developmental, neurological, and psychiatric conditions. The Clock Drawing Test (CDT) is also adopted as a screening tool for cognitive impairment in several conditions such as dementia, stroke, schizophrenia, mild cognitive impairment and mood disorders¹. The CDT basically consists of the drawing of different elements of a clock, including the clock face, numbers, and arrow-hands. The instructions and scoring methods vary according to each version. Influences of aging and education are also related to test performance, where children and the elderly perform worse than adults and lower education associates with worse performance².

Different versions of the CDT were developed. Although very similar and highly correlated³, they differ in diagnosis accuracy⁴ and probably on the cognitive processes involved.
in test performance. Some versions\textsuperscript{1,5} require the patient to freely draw the clock face, demanding more pronounced visuoconstructional abilities, while other versions\textsuperscript{6,7} offer a pre-drawn circle for the clock face, focusing on the abilities to organize the clock elements on the empty space.

Different versions of the CDT are used in Brazil for the neuropsychological assessment of elderly patients, including the one adopted in this research, which seems a validity test for the distinction of patients with normal or impaired cognitive status\textsuperscript{8}. A study compared different versions of the CDT in terms of criterion validity and did not find significant differences between them\textsuperscript{9}. In addition, it showed good evidence of reliability, but reduced criterion-related validity for this population\textsuperscript{8}. Other studies\textsuperscript{10,11}, however, assessed CDT validity for cognitive assessment among individuals with heterogeneous educational backgrounds, including illiteracy, showing good results, but emphasizing the role of formal education on test performance.

Despite these factors, the CDT usually shows significant correlations with measures of visuospatial abilities\textsuperscript{2,12}, executive functions\textsuperscript{3}, semantic processing\textsuperscript{13}, or more global and diffuse cognitive abilities\textsuperscript{8}. In Brazil, it is primarily used for the assessment of executive functions\textsuperscript{4,14}. In this sense, since the score is related to several cognitive abilities impaired performance is less informative concerning which cognitive domain is more affected. According to a review of different scoring methods\textsuperscript{6} the CDT shows good evidence of psychometric properties, such as criterion validity and reliability. However, concerning construct validity, the studies are heterogeneous, using different neuropsychological measures as the gold standard for specific functions. Therefore, in this study we aim to verify what type of cognitive processes (executive functions, visuospatial processing, and semantics) is more related to the performance in the CDT. Our hypothesis is that the CDT is influenced by several cognitive domains, most importantly executive functions.

METHODS

Participants

We assessed 170 older adults. The groups were composed of patients diagnosed with mild cognitive impairment (MCI, $N=53$, 68\% women), mild Alzheimer’s dementia (AD, $N=60$, 55\% women) and normal controls matched by age, gender, and education (NC, $N=57$, 67\% women). The inclusion criteria were age equal to or greater than 60 years, at least 1 year of formal education, score on the Short Version of the Geriatric Depression Scale\textsuperscript{15} below the cutoff for depression, and lack of current use of typical or atypical antipsychotic drugs. MCI diagnosis was performed based on Petersen’s criteria\textsuperscript{16} and AD on the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer’s Disease and Related Disorders Association (ADRDA) criteria\textsuperscript{17}. All participants underwent a clinical and neuropsychological assessment at the Instituto Jenny Faria de Atenção à Saúde do Idoso, a secondary/tertiary health unit of the city of Belo Horizonte for elderly people. Diagnoses were performed in consensus by at least one geriatrician and one clinical neuropsychologist. The control group underwent neuropsychological assessment for the exclusion of cognitive impairment and answered a semi-structured interview for the exclusion of psychiatric, neurologic, or other disorders. All participants and their caregivers signed the written consent for participation. The study was approved by the Ethical Board of the Universidade Federal de Minas Gerais (COEP 334/06).

Neuropsychological assessment

For the present study Shulman’s version of the CDT\textsuperscript{8} was chosen. This is one of the most sensitive versions for cognitive impairment and dementia\textsuperscript{8}. On this CDT version the scoring system ranges from 0 (worst) to 5 (best). Based on previous CDT studies, we aim at the assessment of global cognitive functioning, executive functions, visuospatial abilities, and semantic knowledge on CDT performance, choosing four neuropsychological tests to represent these abilities.

Global cognitive functioning

The Brazilian version of the Mini-Mental State Examination (MMSE)\textsuperscript{18} was chosen as a global measure of cognitive functioning. The test consists of a brief assessment of temporal and spatial orientation, short-term memory, attention/calculation, language, motor, and constructional praxis. Scores range from 0 (worst) to 30 (best).

Executive functions

This complex and multidimensional cognitive domain was assessed by the Brazilian version of the Frontal Assessment Battery (FAB)\textsuperscript{19}, a bedside short screening battery for the assessment of frontal/executive functions. The battery has six subtests that have been developed for mapping specific components of executive functions. Scores range from 0 (worst) to 18 (best).

Visuospatial abilities

We chose a non-timed assembly model task for assessing this cognitive domain. The Stick Design Test (SDT)\textsuperscript{20} was developed for the constructional praxis examination in elderly individuals with little formal education, and has shown solid evidence of validity for the Brazilian population\textsuperscript{21}. The scores range from 0 (worst) to 12 (best).

Semantic knowledge

A naming test was chosen for the semantic assessment. Since cultural differences strongly influence naming
abilities, a Brazilian Test with regionalized stimuli was chosen: the Teste de Nomeação do Laboratório de Investigações Neuropsicológicas/Naming Test of the Laboratory of Neuropsychological Investigations (TN-LIN)².

**Statistical procedures**

Histogram analysis suggests non-normal distribution for most of the neuropsychological tests. In this sense, comparisons considering sociodemographic variables and neuropsychological assessment were performed by the Kruskal–Wallis test. Mann–Whitney tests (Bonferroni corrected) were used for specific group comparisons. An exploratory analysis of the association of the CDT with other cognitive measures was carried out by the Spearman rank-order correlation. After this preliminary procedure, the association between the different cognitive measures and CDT performance was investigated by ordinal regression models. Six models were tested. The first contains the CDT as dependent variable, and age and education as predictors. Models 2, 3, 4, and 5 contain the previous variables and each of the four neuropsychological tests independently. The final model (6) contains age, education, and all the neuropsychological measures together as predictors. We aim at the analysis of specific contributions of each predictor (Wald statistic) and the effect size of the model (N格尔kerke $R^2$ statistic).

**RESULTS**

The participants did not differ in terms of age, formal education, depressive symptoms, and proportion of men and women among groups (all $p>0.05$). On the neuropsychological assessment, differences were found in all cognitive measures, with NC, MCI, and AD patients performing progressively worse on the cognitive tests except for the SDT (Table 1).

The correlational coefficients between CDT scores and the other neuropsychological measures were all significant at $p<0.001$. The strongest correlation was found for the executive functions ($0.628$), followed by general cognitive functioning ($0.595$), semantic knowledge ($0.574$), and, lastly, visuospatial abilities ($0.518$). All proposed regression models were significant ($p<0.001$). Alone, the sociodemographic predictors account for 16% of the explained variance, while the models containing these predictors and one neuropsychological measure explain from 29% to 42% of the variance. When all the cognitive measures are taken together, the model reaches the largest effect size, with 53% of explained variance. In the last model, the executive functions, assessed by the FAB, were the best predictor ($Wald \chi^2 = 13.79, p<0.001$) (Table 2).

**DISCUSSION**

Our result suggests that different cognitive domains are related to CDT performance. The study uses different and well validated measures for each cognitive domain and they were all related to CDT performance. Together, the influence of the cognitive predictors was large, accounting for 53% of CDT variance.

The cognitive domain most related to CDT performance in the present study was the executive functioning, measured by the FAB total score. This cognitive construct has been related to CDT performance before, integrating a general factor of executive function in a participant’s sample similar to the one found in our study²³. The prefrontal cortex is related to CDT performance, especially to the time required for the drawing possibly due to more efficient planning²⁴. Functional magnetic resonance imaging studies also suggest bilateral frontal lobe activation during the CDT, although other regions are also involved²⁵. Studies with patients with focal lesions, circumscribed to the frontal lobes, have reported CDT impairment¹. Freedman et al.¹ argue that a different pattern of CDT impairment may be seen in right frontal hemisphere lesions (characterized by mild spatial deficits, misorientation of arrow-hands, and number inversion) when compared

Table 1. Participants' description and group comparisons of sociodemographic characteristics and neuropsychological tests.

<table>
<thead>
<tr>
<th>Sociodemographic and cognitive variables</th>
<th>NC (1) $N=57, F=38$</th>
<th>MCI (2) $N=53, F=36$</th>
<th>AD (3) $N=60, F=33$</th>
<th>Group comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>MED 74, SEM 1.02</td>
<td>MED 75, SEM 1.20</td>
<td>MED 76, SEM 0.79</td>
<td>KW 3.18, p=0.196</td>
</tr>
<tr>
<td>Education</td>
<td>MED 5, SEM 0.56</td>
<td>MED 4, SEM 0.53</td>
<td>MED 5, SEM 0.28</td>
<td>p=0.076</td>
</tr>
<tr>
<td>GDS-15 (Max=15)</td>
<td>MED 2, SEM 0.18</td>
<td>MED 2, SEM 0.22</td>
<td>MED 2, SEM 0.17</td>
<td>p=0.636</td>
</tr>
<tr>
<td>CDT (Max=5)</td>
<td>MED 5, SEM 0.14</td>
<td>MED 3, SEM 0.23</td>
<td>MED 2, SEM 0.20</td>
<td>p=0.001</td>
</tr>
<tr>
<td>MMSE (Max=30)</td>
<td>MED 28, SEM 0.34</td>
<td>MED 25, SEM 0.47</td>
<td>MED 20, SEM 0.52</td>
<td>p=0.001</td>
</tr>
<tr>
<td>FAB (Max=18)</td>
<td>MED 16, SEM 0.34</td>
<td>MED 12, SEM 0.38</td>
<td>MED 9, SEM 0.40</td>
<td>p=0.001</td>
</tr>
<tr>
<td>SDT (Max=12)</td>
<td>MED 12, SEM 0.04</td>
<td>MED 12, SEM 0.25</td>
<td>MED 10, SEM 0.31</td>
<td>p=0.001</td>
</tr>
<tr>
<td>TN-LIN (Max=65)</td>
<td>MED 63, SEM 0.46</td>
<td>MED 61, SEM 0.52</td>
<td>MED 54, SEM 1.10</td>
<td>p=0.001</td>
</tr>
</tbody>
</table>

NC: normal controls; MCI: mild cognitive impairment; AD: Alzheimer’s dementia; MED: Median; SEM: Standard Error of the Mean; GDS-15: Geriatric Depression Scale 15; CDT: Clock Drawing Test; MMSE: Mini-Mental State Examination; FAB: Frontal Assessment Battery; SDT: Stick Design Test; TN-LIN: Teste de Nomeação do Laboratório de Investigações Neuropsicológicas (Neuropsychological Investigations Laboratory Naming Test); KW: Kruskal-Wallis Test.

*Performed by group–group Mann–Whitney tests with Bonferroni correction.
with left frontal lobe lesions (difficulty in understanding the test instructions, plotting an inverse number pattern, and perseverations), although, as shown by another study\(^9\), this pattern may not be precisely clear. CDT performance may demand planning skills (for overall task goals), abstract reasoning (for number–space–arrow-hand comprehension), inhibitory control (for correctly sequencing the numbers and arrow-hand placing), and cognitive-shifting for adopting new strategies. All these specific cognitive processes are involved in executive functioning, the best single predictor of CDT performance according to our results.

Global cognitive impairment was the second predictor of CDT performance in our study and may be related to the integration of different cognitive domains necessary for task performance. As reported by previous studies\(^4,6,27\) the CDT is a powerful screening test for cognitive impairment, with evidence of sensitivity, specificity, and significant correlations with global cognitive measures, such as the MMSE. The multiple cognitive domains involved in test performance and the necessary integration of these independent aspects for accurate drawing may justify its role in cognitive screening. As a more general cognitive measure, the CDT can be used for detection of cognitive impairment\(^4,6,27\), staging of cognitive impairment\(^4,6,27\), differential diagnosis\(^4,6,28\), and for predicting subsequent cognitive decline\(^4,29\). In more discrete conditions, however, the results are inconsistent. Our results suggest worse performance of MCI patients than NC, but a review of similar studies found negative results\(^30\). The second predictor of CDT performance in our study may be related to the integration of different cognitive domains necessary for task performance.

The third significant predictor of CDT performance was semantic knowledge, assessed by a naming task. A previous study\(^9\) deconstructing the CDT into different components found moderate associations with tasks related to semantic processing, such as naming and category fluency. Drawing involves the search and recall of semantic representations of the clock itself, such as number disposition, hand-arrow meaning, and the spatial logics of its organization. In this sense, the degraded semantic knowledge, assessed by naming tests, may impair CDT performance. Other authors\(^31\), working with AD and frontotemporal dementia patients (including semantic dementia), found quantitative and qualitative impairments in CDT performance that may, at least partially, be related to semantic deficits. Functional neuroimaging studies with AD patients identified posterior left temporal lobe (areas related to language and semantic processing) blood flow to be a predictor of CDT performance\(^32,33\).

Finally, visuospatial abilities were related to CDT performance in our sample. Posterior brain regions are essential for drawing tasks, which involve the perception/semantic recall of an image, planning, and the subsequent motor programming for task execution. CDT usually presents moderate correlations with other measures of visuospatial processing, such as block assembly and figure drawing\(^31,33\). Patients with vascular lesions, especially those on the posterior right hemisphere\(^34\), usually show expressive impairment in the CDT. A large study of patients with focal brain injury suggests that CDT scores are associated with the right parietal cortex and the left inferior frontal-parietal opercular cortices\(^35\). Visuospatial abilities may contribute to different aspects of CDT performance, from visual perception of the ongoing

<table>
<thead>
<tr>
<th>Model</th>
<th>(\chi^2)</th>
<th>(R^2)</th>
<th>Predictors</th>
<th>B</th>
<th>Wald</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.00 (p&lt;0.001)</td>
<td>0.16</td>
<td>Age</td>
<td>-0.51</td>
<td>7.29</td>
<td>0.007</td>
</tr>
<tr>
<td>2</td>
<td>85.70 (p&lt;0.001)</td>
<td>0.41</td>
<td>MCI</td>
<td>-0.03</td>
<td>1.96</td>
<td>0.162</td>
</tr>
<tr>
<td>3</td>
<td>88.10 (p&lt;0.001)</td>
<td>0.42</td>
<td>Age</td>
<td>-0.04</td>
<td>5.16</td>
<td>0.023</td>
</tr>
<tr>
<td>4</td>
<td>56.81 (p&lt;0.001)</td>
<td>0.29</td>
<td>MMSE</td>
<td>0.27</td>
<td>50.54</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5</td>
<td>65.12 (p&lt;0.001)</td>
<td>0.33</td>
<td>MMSE</td>
<td>0.27</td>
<td>50.54</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6</td>
<td>120.73 (p&lt;0.001)</td>
<td>0.53</td>
<td>TN-LIN</td>
<td>-0.01</td>
<td>2.73</td>
<td>0.098</td>
</tr>
</tbody>
</table>

MMSE: Mini-Mental State Exam; FAB: Frontal Assessment Battery; SDT: Stick Design Test; TN-LIN: Teste de Nomeação do Laboratório de Investigações Neuropsicológicas (Neuropsychological Investigations Laboratory Naming Test).
copy to spatial planning and visual working memory. Even basic aspects of spatial reasoning, such as magnitude perception, may be related to CDT performance. Only a small fraction of the explained variance of CDT performance was explained by the adopted measure of visuospatial abilities. This may be secondary to the relative small variance of the SDT measure (the only significant differences were between non-demented and AD participants). Future studies should investigate the contributions of this cognitive domain to CDT performance on a more heterogeneous sample in order to maximize test variance.

Certainly, different versions of the CDT may present different cognitive demands. The selected version uses a pre-drawn circle where the patients must organize the clock, which may reduce the visuospatial demands for test performance, focusing on more executive components as planning and online monitoring. The procedures adopted here could be used in future research, examining different CDT versions and comparing the cognitive demands of each one. These studies may guide neuropsychologists and other clinicians in the choice of which CDT version may be more accurate for different hypotheses about cognitive functioning.

The influence of formal education on test performance must also be addressed in this discussion. Although education entered as a covariate on the ordinal linear regression models, the low variance between the groups (not significant) may reduce its influence on test performance. Although some studies reported good criterion validity for the CDT among poorly educated older adults\textsuperscript{8,10,11}, these findings are controversial\textsuperscript{19}. Concerning construct validity, for patients with very low formal education, the drawing processes may demand, besides more cognitive effort than better educated subjects, the recruitment of different cognitive abilities to improve task performance. In this sense, the greater association of executive functions with CDT performance may represent a compensatory process for drawing performance. Future studies should replicate our findings in elderly patients with higher formal education.

Our finding indicated that the CDT must be interpreted with caution in a clinical setting. For screening purposes of global cases of cognitive impairment, the test seems to be an adequate measure, with moderate correlations with other screening tests. However, in a comprehensive neuropsychological assessment, the CDT scores must be interpreted carefully, considering the influences of different specific cognitive domains on test performance.

CONCLUSION

The CDT in normal aging, MCI, and mild AD subjects was influenced by executive functions, semantic knowledge, visuospatial abilities, and general cognitive functioning. Together these components accounted for about half of the explained variance of the CDT. The results suggest that the CDT is a complex multidimensional measure, and should be interpreted along with other cognitive tests in the neuropsychological assessment.

References