Virtual Kitchen Test

Assessing Frontal Lobe Functions in Patients with Alcohol Dependence Syndrome

P. Gamito1,2; J. Oliveira1,2; C. Caires1; D. Morais1,2; R. Brito1,2; P. Lopes1,2; T. Saraiva1,2; F. Soares1; C. Sottomayor1; F. Barata1; F. Picarelli1; M. Prates1; C. Santos3

1School of Psychology and Life Sciences – Lusophone University of Humanities and Technologies, Lisbon, Portugal; 2COPELABS – Cognition and People-centric Computing Laboratories – Lisbon, Portugal; 3Lisbon Psychiatric Hospital Centre, Lisbon, Portugal

Keywords
Frontal functions, cognitive assessment, mHealth, alcohol dependence syndrome

Summary
Introduction: This article is part of the Focus Theme of Methods of Information in Medicine on "New Methodologies for Patients Rehabilitation".

Background: The ecological validity of paper-and-pencil neuropsychological tests is currently a matter of debate. Arguments in favor of alternatives indicate that paper-and-pencil forms are unable to account for both mental and functional aspects of cognitive functioning.

Objectives: In this study we developed a new neuropsychological evaluation test – the Virtual Kitchen Test (VKT) – devised to evaluate frontal brain functioning in cognitively impaired individuals. We designed this test according to the rationale of the Trail Making Test (TMT), in order to capture frontal lobe abilities during a more ecologically valid task.

Methods: Forty-nine participants, 25 from a clinical sample of patients diagnosed with Alcohol Dependence Syndrome, plus 24 healthy participants.

Results: Execution errors and task completion time were significantly higher in the clinical sample. Also, scores on the new VKT showed moderate to high positive correlations with scores on the TMT. Furthermore, the overall discriminant performance of the VKT was high for both of its indicators.

Conclusions: Overall results support the ability of the VKT to evaluate frontal lobe functions. The best cut-off scores based on this sample are discussed.

Alcohol dependence syndrome (ADS) has also been associated with brain dysfunction, particularly in the prefrontal cortex [3, 4]. Evidence from neuropsychological studies suggests a decrease in executive functioning in alcohol abusers, consisting of impairments in working memory, in attention, in sequencing and planning, in cognitive flexibility, and in inhibitory control [5–8]. These effects persist after discontinuation of consumption of alcohol, and even up to a year of abstinence [9].

Although functionality is usually evaluated with self-reports about everyday functioning in basic activities of daily living (or ADL, measuring self-care skills), or instrumental activities of daily living (or IADL, measuring independent living skills) [10], neuropsychological assessment does not always manage to serve its purpose due to lack of validity of the tests used. Many of the paper-and-pencil tests used to assess cognitive functions lack ecological validity because they propose tasks that do not effectively replicate patients’ impaired daily life activities [11, 12]. One example is the standard and well-known neuropsychological test used to assess frontal brain functioning: the Trail Making Test (TMT). The TMT was originally developed for the Army Individual Test Battery [13] in 1944, and was subsequently included in another neuropsychological battery, the Halstead-Reitan Battery [14]. This test consists of two forms: participants are asked to sequentially connect numbers (TMT-A) and to alternate between numbers and letters (TMT-B). Although the TMT is used to assess the ability to perform sequencing and visual search, cognitive flexibility, and motor and attentional functions [15], the assessment process (connecting numbers and letters) is quite different from a daily life activity context.

In order to fill this gap, the current study proposes a new approach to neuropsychological assessment that takes advantage of the benefits of new technologies, allowing for a more effective and reliable evaluation process by using a task similar to the tasks that are usually performed by the patient in their real-life settings. In-
Indeed, the interdisciplinary synergy of the medical sciences and informatics is a growing area of interest, and has been shown to increase the efficiency of assessment and intervention in health contexts [16]. Accordingly, our aim was to develop a test that combined all the characteristics of a well-established neuropsychological test for frontal functions with a context where these functions could be naturally assessed. Thus, the test that we have developed – the Virtual Kitchen Test – emulates the paper-and-pencil TMT, but with the advantages that new technologies have to offer. These include 3D interaction, which may be important to assess neuropsychological performance under ecologically valid contexts.

The test, by simulating the baking of a cake, reproduces food preparation tasks, which are an important domain of IADL that may be compromised in cases of mild to severe cognitive impairments. In fact, there is evidence suggesting that patients with mild cognitive impairments may experience changes that limit their ability to perform more complex instrumental tasks while maintaining relatively unimpaired ADLs [17].

In the current study, we recruited patients with alcohol dependence syndrome (ADS), because this population is known to have specific cognitive impairments associated to frontal lobe dysfunctions, namely in executive functioning and attention processes.

2. Objectives

Our main objective was to test the concurrent validity of the VKT by assessing the correlation of performance on this test with performance on the TMT. We expected that performance on the VKT would correlate with performance on the TMT, on which it was modeled. Thus, we also expected that ADS patients would perform worse than healthy controls not only on the TMT, but also on the VKT. A second aim was to test the discriminant performance of the VKT in distinguishing between the presence and absence of cognitive impairment through a quantitative approach based on the Receiver Operating Characteristic Curve (ROC).

3. Method

3.1 Participants

Twenty-five participants (17 male) diagnosed with Alcohol Dependence Syndrome (ADS) were recruited from alcoholism treatment units, and 24 healthy participants (10 male) were recruited from the general population (total n = 49). All participants were unpaid volunteers who gave their informed consent to neuropsychological assessment. Gender distribution between these groups did not differ significantly. However, the ADS group was significantly older (M-age = 45 yrs; SD-age = 8.2 yrs) than the healthy group (M-age = 35 yrs; SD-age = 15.6 yrs), t(47) = -2.84; p = .007, and significantly less educated (respectively, 7 vs. 19 participants with secondary or higher education, \( \chi^2(3) = 27.375; p < 0.001 \)). Because age and education are known to have an effect on neuropsychological test performance, these two variables were controlled for in all subsequent statistical analyses. Patients in the ADS group were receiving medication consisting of antidepressants and anxiolytics to minimize withdrawal symptoms; all patients in the ADS group were in the same medication regime.

The exclusion criteria were: a) a previous history of neurological or psychiatric disorders (other than alcohol dependence syndrome in the ADS group); b) a dependence (or history of dependence) on heroin, cocaine, or cannabis; c) scores below the cutoff values for their age and education groups in the Mini-Mental State Examination (MMSE ≤ 24) [18].

This study was approved by an ethics committee based on the statement of ethical principles of the Declaration of Helsinki [19].

3.2 Measures

The study protocol consisted of general cognitive screening tests and a specific test involving visuomotor ability, in order to test correlations with performance on the VKT application. The Mini Mental Examination Test (MMSE) was used to screen...
participants for minimal cognitive ability. Frontal brain functions were assessed with the Frontal Assessment Battery (FAB) [20], which evaluates six domains of frontal lobe functions: conceptualization, mental flexibility, motor programming, sensitivity to interference, inhibitory control, and environmental autonomy. The total score of the FAB (maximum 18 points) is estimated by adding each of the subtest scores. A more specific frontal lobe assessment was carried out using the Trail Making Test (TMT) [13]. Task performance indicators in the TMT are based on execution time and errors in each version of the test.

3.3 Virtual Kitchen Test

A pre-existing 3D kitchen scenario was modified and developed into two tasks or scenes. A kitchen cabinet was added to the scenario, as well as a list displaying a sequence of ingredients that the participants should use to bake a cake. The VKT was available through the web and it was designed with the rationale of the TMT as reference. It requires that the participants drag sequentially, from a kitchen cabinet, the ingredients for a cake displayed on a list. In the first scene of the scenario (Trial I – easier), participants had to drag the ingredients, in the order presented on the list, from the cabinet to the counter. During the second scene (Trial II – difficult), the reverse movement was requested. Thus, the second trial of the VKT was the reverse of VKT Trial I. The specific measures used to assess performance in the VKT were execution errors and time needed for task completion (for each trial), which are similar measures to those used in the TMT. Figure 2 Box-plots for TMT-A (top panel) and VKT Trial I (bottom panel).

3.4 Statistical Procedures

The results were analyzed using the dependent variables from the neuropsychological tests, namely, the total scores of the MMSE and FAB, and execution errors and time in the TMT and the VKT. We first analyzed the effect of ADS on cognitive ability by comparing the results of ADS patients and controls on the standard neuropsychological tests and on the VKT. We then analyzed the bivariate correlations between all dependent variables. Finally, we tested the discriminant performance of both indicators of the VKT in each trial using ROC analysis, as well as sensitivity (positive predictive value) and specificity (negative predictive value) associated with the best cut-off points. Given that there were age and education differences in distribution between cells, these two variables were also included in the analysis in an Analysis of Covariance design. We performed ANCOVAs with two between-subjects factors (i.e., group and education) and one covariate (i.e., age) for all neuropsychological and VKT-related variables.

4. Results

4.1 Comparisons between Patients and Controls in Neuropsychological Functioning

As expected, we found an effect of alcohol dependence on the total score of the MMSE ($F(1, 49) = 5.280; \eta^2 = .06, p = .027$) and on the hit rate of the TMT ($F(1, 49) = 13.858; \eta^2 = .16; p = .001$), with lower scores in ADS patients in comparison to controls, both on the MMSE ($M = 26.84; SE = .40$ vs. $M = 28.72; SE = .55$) and on the TMT ($M = 4.54; SE = .50$ vs. $M = .88; SE = .66$). In the FAB total score, there was not only the same effect of ADS ($F(1, 49) = 4.136; \eta^2 = .05; p = .048$), but also of education ($F(1, 49) = 3.103; \eta^2 = .12; p = .037$). The total score of the FAB was lower in ADS patients ($M = 13.72; SE = .48$) than controls ($M = 15.58; SE = .66$), but also in participants with basic or intermediate schooling ($M = 12.62; SE = .93$). There were no statistically significant effects of factors on the other neuropsychological variables (all $p’s > .05$).

Results showed that ADS patients performed worse than healthy participants on a variety of measures in the VKT: errors in trial I ($F(1, 49) = 8.038; \eta^2 = .12; p = .007$), execution time in trial I ($F(1, 49) = 11.218; \eta^2 = .11; p = .002$), errors in trial II ($F(1, 49) = 12.868; \eta^2 = .18; p = .001$) and execution time in trial II ($F(1, 49) = 14.817; \eta^2 = .15; p < .001$). There was, however, also an effect of the covariate (age) for execution time in trial I ($F(1, 49) = 11.801; \eta^2 = .12; p < .01$) and in trial II ($F(1, 49) = 7.785$;
\( \eta^2 = .08; p < .01 \), which indicates that age slows down performance in the VKT. Estimated marginal means for significant differences between ADS and controls on VKT measures showed lower performance of ADS patients in comparison to the controls (Execution errors Trial I: M = 3.34, SE = .64 vs. M = .53, SE = .88; Execution time Trial I: M = 343.78, SE = 25.40 vs. M = 182.10, SE = 35.03; Execution errors Trial II: M = 7.23, SE = .85 vs. M = 1.92, SE = 1.18; Execution time Trial II: M = 432.33, SE = 32.90 vs. M = 190.01, SE = 45.37, respectively for ADS patients and controls).

Although there was positive skewness in the distribution of the performance of healthy controls in the easiest trials (TMT-A and VKT-I) a ceiling effect in these data would put our analyses at risk of Type II error, which was not the case for the total sample (▶Figure 2).

4.2 Correlations between VKT and Neuropsychological Variables

In order to investigate the relations among the dependent variables related to the VKT and neuropsychological tests, bivariate correlations with Pearson’s \( r \) were computed. The results revealed negative, moderate to strong, significant associations between errors, as well as execution time, in the VKT on both trials with the total scores of the MMSE and the FAB, which indicates that poorer cognitive abilities were related to more errors and longer execution times in the VKT. Additionally, the correlations between the VKT and the TMT were significant only for the B version of the TMT. More precisely, errors and execution time in the VKT on both trials were positively correlated (ranging from \( r = 0.31 \) to \( r = 0.60 \)) with errors and execution time in the TMT B version (▶Table 1).

### 4.3 Discriminant Performance of the VKT

The discriminant performance of the VKT was also analyzed by observing the ROC curves. ROC analysis is a widely accepted method in medical literature to evaluate the accuracy of diagnostic tests. The ROC plots provide the discrimination between sensitivity (true-positive rate) and 1-specificity (false-positive rate) for a given measurement scale [21]. In the current study we performed a ROC analysis to test the discriminant performance of the VKT in distinguishing cognitive impairments. The ROC analysis was used for execution time and errors (absolute measure) for both trials of the test. The Area Under the Curve (AUC) revealed a good discriminant capacity for errors and execution time on both versions of the VKT, more precisely in errors for the direct VKT (AUC = .90; 95% confidence Interval ranged between .83 and .94) and execution time (AUC = .91; 95%). Confidence Interval ranged between .83 and .94); and errors for indirect VKT (AUC = 0.90; 95%. Confidence Interval ranged between .81 and .99) and execution time (AUC = 0.95; 95%. Confidence Interval ranged between .90 and 1.00).

The best cut-off scores were also estimated. The cut-off points for both versions of the VKT are acceptable to good (▶Table 2). They were estimated under the assumption of maximizing the sum of sensitivity and specificity.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cut-off point (1)</th>
<th>Sensitivity</th>
<th>1-Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>Trial I</td>
<td>1</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>Trial II</td>
<td>200 s</td>
<td>.82</td>
</tr>
<tr>
<td>Errors</td>
<td>Trial I</td>
<td>5</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>Trial II</td>
<td>280 s</td>
<td>.92</td>
</tr>
</tbody>
</table>

*Correlation is significant at the .01 level (2-tailed). **Correlation is significant at the .05 level (2-tailed). (1) Positive if greater or equal to 0.

5. Discussion

The project that supported this work was developed with the aim of offering a new neuropsychological tool to perform real life-like assessment of frontal functions. In order to assess this objective, we collected a clinical sample consisting of patients diagnosed with alcohol dependence syndrome (ADS) as well as a sample of healthy controls. The neuropsychological deficits in ADS patients are well documented in the literature (see [3] for a review), particularly regarding frontal lobe functions. Thus, our aim was use an ADS sample to cross-validate the proposed VR exercise-based measures with a traditional TMT. The neuropsychological evaluation carried out on both samples confirmed our predictions, revealing poorer cognitive ability in alcoholics in comparison to healthy volunteers on both traditional and VR-based measures of our Virtual Kitchen Test (VKT), thus supporting the assumption on which we based our test, which was that alcohol abuse has a negative effect on attention and cognitive flexibility, among other cognitive functions located in the frontal lobe.

Importantly, we showed that performance on the VKT is correlated to perform-
Widespread and substantial cognitive impairments and with a greater variety of measures.

We also attempted to study the ability of the VKT to discriminate between patients and controls. A ROC analysis was carried out on execution time and errors from the direct and indirect trials of the VKT. The results showed a good discriminant performance of the test, which was higher in the indirect (more difficult) trial of the VKT, as well as for the execution time in comparison to the number of errors in the task. The cut-off scores were also estimated under the same statistical procedure. Each of these values indicates an adequate level of sensitivity and specificity in discriminating the performance of patients with cognitive impairment vs. healthy individuals. The probability of discrimination a true positive for errors (cognitive impairment; i.e., sensitivity) was, on average, 84%, whereas the probability of discriminating a true negative (without cognitive impairment; i.e., specificity) was on average 89%. More particularly, in the direct trial of the VKT, which is the easier version of this test, one or more errors may be indicative of a cognitive impairment equivalent to the one revealed by alcoholics. The normative execution time of this test should be below 200 seconds (more than or equal to 200 seconds may be indicative of a cognitive deficit). In the indirect trial (more demanding version of the VKT), five or more errors or an execution time with more than or equal to 280 seconds may reveal cognitive deficits similar to the ones revealed by alcoholics.

Overall, the results from different statistical procedures seem to support the validity of the VKT in evaluating neuropsychological functions, particularly those related to the frontal lobes. However, further studies are needed with clinical samples of brain-injury patients in order to broaden the application of the VKT to other contexts of cognitive impairment. For example, it would be interesting to understand if the VKT is able to differentiate between different types of cognitive impairment, for example, those related to a focal brain injury, or a more global dysfunction resulting from a neurodegenerative condition.

There are, however, several limitations in the current study that need to be addressed. The main limitation is related to the differences found between the sample of patients and controls in potential confounders of neuropsychological testing: Although age and education were controlled statistically in the comparisons between patients and controls, the effects of age on execution time were not possible to isolate in the ROC analysis due to the small sample size. Thus, additional validation studies with age and education-matched control samples are needed in order to provide further support the use of the VKT in neuropsychological testing. Another limitation concerns the concurrent neuropsychological measures used in our study to assess frontal lobe abilities. The assessment of frontal functions was conducted with the TMT and a more global test, the FAB, but in subsequent studies, frontal lobe assessment should be extended to executive functioning as well. In sum, the results on the VKT open up the prospect of using VR tasks that replicate daily life activities as a viable alternative to the traditional paper and pencil tests, but require further testing on patients with a broader variety of cognitive impairments and with a greater variety of measures.

References