Color Trails Test: Normative Data and Criterion Validity for the Greek Adult Population

Lambros Messinis1,*, Amaryllis-Chryssi Malegiannaki1,2, Tessa Christodoulou1, Vassillis Panagiotopoulos1,3, Panagiotis Papathanasopoulos1

1Department of Neurology, Neuropsychology Section, University of Patras Medical School, Rion, Patras, Greece
2Department of Psychology, Cognitive and Neuropsychology Division, University of Thessaloniki, Greece
3Department of Neurosurgery, University of Patras Medical School, Rion, Patras, Greece

*Corresponding author at: Department of Neurology, Neuropsychology Section, University of Patras Medical School, Rion, Patras, Greece.
Tel: +30-2610-999-348; Fax: +30-2610-455-209.
E-mail address: lmessinis@upatras.gr (L. Messinis).
Accepted 31 March 2011

Abstract

The Color Trails Test (CTT) was developed as a culturally fair analog of the Trail Making Test. In the present study, normative data for the CTT were developed for the Greek adult population and further the criterion validity of the CTT was examined in two clinical groups (29 Parkinson’s disease [PD] and 25 acute stroke patients). The instrument was applied to 163 healthy participants, aged 19–75. Stepwise linear regression analyses revealed a significant influence of age and education level on completion time in both parts of the CTT (increased age and decreased educational level contributed to slower completion times for both parts), whereas gender did not influence time to completion of part B. Further, the CTT appears to discriminate adequately between the performance of PD and acute stroke patients and matched healthy controls.

Keywords: Executive functions; Attention; Norms/normative studies; Parkinson’s disease; Cerebrovascular disease/accident and stroke; Validity

Introduction

Due to their ease of administration and sensitivity to brain damage, Trail Making Tests (TMTs) have long been among the most widely used measures in the practice of neuropsychology (Lezak, Howieson, & Loring, 2004). The original TMT developed in 1944, due to its linguistic and phonologic nature is significantly restricted within cross-cultural settings. The Color Trails Test (CTT; D’Elia, Satz, Uchiyama, & White, 1996) was developed as a culturally fair analog of the TMT (Armitage, 1946); however, it was not designed to produce equivalent or even similar time to completion scores when compared with the TMT, but rather to tap similar frontal/executive cognitive abilities and to allow broader cross-cultural applications. In order to minimize cultural and language bias, the CTT does not use any letters (of any language), and instructions are given either orally or non-verbally (visual indications; D’Elia et al., 1996; Vlahou & Kosmidis, 2002). For the CTT part 1 (CTT1), the respondent uses a pencil to rapidly connect circles numbered 1 through 25 in sequence. For the CTT part 2 (CTT2), the respondent rapidly connects numbered circles in sequence, but alternates between pink and yellow colors. The examiner records the length of time to complete each trial along with qualitative features of performance indicative of brain dysfunction, such as near-misses, prompts, number sequence errors, and color sequence errors (see test manual, D’Elia et al., 1996). Respondents aged ≥ 18 must be able to recognize Arabic numerals from 1 to 25 and to distinguish between the colors pink and yellow (D’Elia et al., 1996; Strauss, Sherman, & Spreen, 2006).

Although the physical designs of the CTT and TMT are similar, they are not, nor were they intended to be, identical. The CTT1 and TMT part A generally have similar periods of time to completion; however, the CTT2 generally takes slightly longer to complete than the TMT part B (Mitrushina, Broone, Razani, & D’Elia, 2005). Longer completion time for the CTT2
compared with the TMT part B has been demonstrated in several studies and in different cultures (see D’Elia et al., 1996, for English-speaking individuals, Dugbarty, Townes, & Mahurin, 2000, for Turkish samples, and Lee & Chan, 2000, for Chinese (Mandarian-)speaking samples). This difference in longer time to completion for the CTT2 most probably can be attributed to the difference in test demands, with the CTT2 requiring almost twice as many stimuli to scan and to consider compared with the TMT part B (Mitruschina et al., 2005).

To be more specific, Lee and Chan (2000) studied 108 adults between the ages of 17 and 54 in Hong Kong and found high correlations (.72) between part B of the TMT and CTT2 only in individuals of an older age (35–54) and higher education (12–22 years). The authors concluded that these two tests estimate corresponding cognitive fields only within the specific education and age limits. Furthermore, Lee, Cheung, Chan, J & Chan, C (2000) reported high correlations between CTT1 and CTT2 and the TMT (.72 and .75, respectively) in English-speaking participants, but low correlations (.25) between part A of the TMT and CTT1 in Chinese participants. These findings demonstrate an effect of native language on CTT1 performance. Maj and colleagues (1993) reported moderate correlations between CTT1 and CTT2 and parts A and B of the TMT (.41 and .50, respectively) and a factor analysis of the CTT, TMT, Stroop test, and the Color figure mazes test revealed that the CTT loaded on the same factors with parts A and B of the TMT in geriatric and non-geriatric samples (Strauss et al., 2006). In a study with 64 Turkish students (mean age 22.67 years), it was found that TMT part A and the CTT1 actually estimate corresponding cognitive fields (Dugbarty, Townes, & Mahurin, 2000). However, the significant differences in the completion time of TMT part B and CTT2 led the authors to the conclusion that these are two qualitatively different tests. The authors’ general conclusion was that the cognitive alternation, required in the CTT2 between numbers and colors, called for a greater information processing and visual-cognitive perceptivity than the alternation between numbers and letters required in part B of the TMT (Strauss et al., 2006).

Regarding the impact of demographic factors on CTT performance, data obtained from the U.S. standardization sample (1,528 individuals, including 182 African Americans and 292 Hispanic Americans, between 18 and 89 years of age, with mean age = 57.72 years, SD = 20.06) revealed that increasing age and lower education adversely affects performance on both the CTT 1 and CTT2. Participants who had higher education levels and were aged between 18 and 29 years had better performance on the CTT1 (M = 26 s, SD = 8.47) and CTT2 (M = 56 s, SD = 16.23). Participants aged >75 with minimal educational levels had the longest times to completion on both CTT1 (M = 73 s, SD = 24.74) and CTT2 (M = 144 s, SD = 45.51). Preliminary normative data for the Hispanic Americans are provided separately since all participants in this subsample were educated outside the USA and were primarily Spanish speaking. Although normative data for the Hispanic Americans are provided separately, this study does not actually report whether ethnicity affects test scores. Further, the influence of gender in this study was not significant on CTT performance after removing the effects of age (D’Elia et al., 1996), although it should be noted that the sample was primarily men as women compromised only 12% of the sample.

In another study providing normative data for Spanish-speaking individuals (Ponton et al., 1996), the authors report that both age and education influence performance on the CTT. The norms provided in this study, however, due to their age group intervals are more adequate for younger individuals under the age of 50. La Rue, Romero, Ortiz, Liang, and Lindeman (1999) provide normative data for 797 community dwelling senior adults aged 65–97 of either non-Hispanic white or Hispanic ethnic origin. They note the contribution of age and education on CTT performance and stratify their norms based on various age and education groups in addition to the ethnic group. In a recent study that evaluated the CTT in a Brazilian healthy sample (1,942 participants aged 18–86), the authors reported significant contributions of age and education on completion time, with significant differences noted between the various age groups and the worst results for ages >60. Further, the authors note that for both genders, an increase in completion time was noticed with increasing age, and better performance was associated with increasing levels of education (Sant’Ana Rabelo et al., 2010).

In general, the findings of the above studies showed that increased age negatively affects CTT performance in both test parts, whereas higher educational levels are associated with better performance, especially on CTT2 (Mitruschina et al., 2005). Gender on the other hand does not seem to particularly affect CTT performance (D’Elia et al., 1996). Regarding the influence of ethnicity on CTT performance, the manual (D’Elia et al., 1996) does not report whether ethnicity affects test scores, although it does provide some information that performance varies by ethnicity/geographical location (e.g., Caucasian, Hispanic, Cantonese). In this respect, the use of norms imported from one foreign context to another is discouraged (Lee & Chan, 2000). Despite the absence of information regarding the influence of ethnicity on CTT performance in the CTT manual, other research to date suggests that the CTT holds promise for cross-cultural and clinical assessment of sequencing, visual scanning, and speed of mental processing in non-English-speaking adults and adults with limited education, English as a second language, and reading disorders (Mitruschina et al., 2005).

The TMT has also proven very sensitive in detecting cerebral dysfunction in different clinical populations (Armitage, 1946; Cicerone & Azulay, 2002; Reitan & Wolfson, 2004; Vlahou & Kosmidis, 2002). Respectively, significant delays in the
completion time of CTT1 and CTT2 have been reported in patients who have suffered traumatic brain injuries (D’Elia et al., 1996) and HIV seropositives (Maj et al., 1993). More recently, Lee and Cheung (2005) noted a relationship between perceptual-visual function and attention in schizophrenia patients tested with the CTT. Other research in individuals with acquired brain injury who sought to obtain a drivers license (Hartman-Maeir, Erez, Ratzon, Mattatia, & Weiss, 2008) and were assessed with the CTT indicated that approved participants on the traffic test completed the CTT1 and CTT2 more quickly than individuals who failed the traffic test, although it should be noted that the difference between groups was significant only for the CTT1. In a recent study that examined the effects of rivastigmine in patients with cognitive impairment due to cerebrovascular disease, the CTT showed a nonsignificant trend toward improvement in these patients when compared with baseline performance, potentially indicating the sensitivity of tests in detecting executive functioning and attention deficits (Narasimhalu et al., 2010).

In the present study, considering the need for standardized neuropsychological tests in Greece, the relatively good psychometric properties of the CTT and its potential cross-cultural applicability, we decided that such a tool would be useful to Greek clinicians and researchers and therefore developed norms for the Greek adult population. Additionally, we examined the criterion validity of the test, administering it to two groups of patients (Parkinson’s disease [PD] and patients who had suffered a recent ischemic stroke). These two clinical samples were chosen, first due to the fact that they usually present deficits in the cognitive domains examined by the CTT (Dubois & Pillon, 1997; Messinis, Lyros, & Papathanasopoulos, 2007; Nys et al., 2005; Sachdev, Brodaty, Valenzuela, Lorentz, & Koschera, 2004) and secondly because there is insufficient data available in the literature for the criterion validity of the CTT based on these two clinical groups. Finally, we estimated the convergent validity between the Greek TMT and CTT in the healthy participants.

Method

Participants

There were three groups of Greek participants in the study (one nonclinical and two clinical) including: (i) a nonclinical group (n = 163; mean age = 45.71; mean educational level = 12.85 years), (ii) a PD group (n = 29; mean age = 63.76; mean educational level = 10.10 years), (iii) a stroke group (n = 25; mean age = 65.08; mean educational level = 8.92 years). The demographic data for these three groups shown in Table 1. The nonclinical group was recruited primarily from southwestern Greece (sample of convenience) and participated in the present study voluntarily, and after providing written informed consent for their participation. Potential nonclinical participants were approached by the experimenters with the goal of including a broad range of adult ages and education levels. These participants were screened, based on a clinical interview that was conducted by the researchers. Exclusion criteria were (a) recent diagnosis of drug abuse or any history of psychiatric diagnosis based on DSM IV-TR, (b) participants suffering from any neurological disease that might affect neuropsychological performance, (c) participants suffering from any other medical condition or those that were undergoing a pharmaceutical regimen that might affect neuropsychological performance, and (d) participants that were non-native Greek speakers. Likewise, nonclinical participants above the age of 60 that scored below 27 in the Mini-Mental State Examination (MMSE; Fountoulakis, Tsolaki, Chantzi, & Kazis, 2000), which is used as a brief screening measure for global cognitive deficits, were also excluded.

In order to establish whether the CTT could discriminate between groups with established brain dysfunction from their healthy counterparts, we administered it to two clinical samples, suffering from neurological diseases. We specifically examined a group of PD patients that had visited the outpatient neurological unit of Patras University Hospital for examination and renewal of their pharmaceutical regimen (see Table 1 for demographic characteristics of PD patients). The average age of manifestation of illness was 57.07 years (SD = 10.13). Duration of the disease was 6.8 years (SD = 5.23) on average.

Table 1. Demographic data for nonclinical and two clinical groups

<table>
<thead>
<tr>
<th></th>
<th>Nonclinical group (N = 163)</th>
<th>Stroke patients (N = 25)</th>
<th>PD patients (N = 29)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>45.71 ± 17.40</td>
<td>65.08 ± 7.50</td>
<td>63.76 ± 9.23</td>
</tr>
<tr>
<td><strong>Education level (years)</strong></td>
<td>12.85 ± 2.90</td>
<td>8.92 ± 3.68</td>
<td>10.10 ± 4.28</td>
</tr>
<tr>
<td><strong>Gender (men:women)</strong></td>
<td>79:84</td>
<td>14:11</td>
<td>16:13</td>
</tr>
</tbody>
</table>

Note: PD = Parkinson’s disease.
Their daily medicine dosage L-Dopa type was 423.28 mg on average (SD = 303.83), whereas their UPDRS score (Soukup & Adams, 1996) was 37.54 (SD = 12.23) on average. Their mean score on the MMSE scale (Fountoulakis et al., 2000) was 27.55 (SD = 1.90). In the short form of the Beck Depression Inventory (Beck, Steer, & Brown, 2000), their mean score was 3.62 (SD = 3.38). All participants of this group had a normal premorbid intellectual ability (premorbid IQ >75) as determined by the Wechsler Abbreviated Scale of Intelligence (WASI; Messinis & Papathanasopoulos, 2010; Wechsler, 1999). Patients with PD Dementia (PDD) according to DSM-IV-TR, 2000 criteria and modified PDD criteria by Emre and colleagues (2007), history of severe psychiatric disorder or other coexisting neurological conditions were excluded, due to the possibility of complicating neuropsychological assessment.

We further examined a group of patients who suffered a recent ischemic stroke (see Table 1 for demographic characteristics of stroke patients). For 17 patients, this was their first stroke, while the other 8 patients had suffered at least two more ischemic strokes. Mean duration of assessment from the day of the stroke was 8.5 days (SD = 2.8). According to the Oxford Stroke Classification System (Bamford, Sandercock, Dennis, Burn, & Warlow, 1991), 4 individuals had suffered a TAC-Total Anterior Circulation type stroke, 11 a PAC-Partial Anterior Circulation stroke, 6 an LAC-Lacunar stroke, and finally 4 individuals suffered a POC-Posterior Circulation stroke. The diagnosis and classification of the stroke type were inferred by an experienced neurologist specializing in stroke patients. Moreover, neuroimaging data (CT scan or MRI) were available, in order to make a distinction between ischemic and hemorrhagic strokes. According to the Rankin scale (van Swieten, Koudstaal, Visser, Schouten, & van Gijn, 1988), which measures the degree of disability of patients who have suffered a stroke, the grading of the group on this scale was (median = 3, range = 1–4). All participants had a normal premorbid intelligence ability (premorbid IQ >75). Patients with serious communication disabilities, confusion, Alzheimer’s, or vascular dementia according to DSM-IV-TR, 2000 criteria, a history of severe psychiatric disorder or other coexisting neurological conditions, which might affect neuropsychological performance, were excluded.

Procedure

Nonclinical participants were assessed in the neuropsychology section, Neurological Department of Patras University Hospital with the Greek version of the TMT (Greek letters are utilized in the Greek TMT and a standardized version with normative data is available, see Vlahou & Kosmidis, 2002) and the CTT. In order to ensure that the order in which the CTT and the Greek TMT were administered would not affect the results, we initially administered the CTT to half of the healthy participants, followed by the Greek TMT, and to the remainder, we initially administered the Greek TMT and then the CTT. It should be noted, however, that there is no evidence of such effects in the literature (Llorente et al., 2003; Strauss et al., 2006).

Regarding the method of administration and scoring of the two tests, we followed the most common methods, as suggested in the test manuals and the relevant publications (D’Elia et al., 1996; Vlahou & Kosmidis, 2002). More specifically, the examiner administered a short demonstrative trial for each part of the two tests, and afterward administered the actual test, once convinced that the participants understood the directions in order to complete the test.

The two clinical groups were assessed in the neuropsychology section, Neurological Department of Patras University Hospital. The CTT was administered within the framework of a short battery of tests at the beginning of the testing session. These included the Vocabulary subtest and the Block Design subtest of the WASI battery (Messinis & Papathanasopoulos, 2010; Wechsler, 1999), the Boston Naming Test (Tsolaki, Tsantali, Lekkas, Kiosseoglou, & Kazis, 2003), the Verbal Fluency Test (Kosmidis, Vlahou, Panagiotaki, & Kiosseoglou, 2004), the Rey Auditory-Verbal Learning Test (Messinis et al., 2007), and the Beck Depression Inventory-short form (Beck et al., 2000).

The dependent variable for both tests (CTT and Greek TMT) used in the research was the completion time for each part of the test. Potential qualitative features of performance, such as near-misses, prompts, number sequence errors, and color sequence errors, were recorded, but not included in the analyses of the data.

Statistical Analysis

Nonclinical group. We initially calculated the demographic data (means, SDs) for this group. Stepwise multiple linear regression analyses were then used to examine the potential contribution of demographic variables (age, gender, and years of formal education) to completion time for the CTT. Given the significant contribution of age and education to the completion time for CTT, and in order to obtain normative data for the Greek adult population, we grouped our sample into demographic categories. Graphs illustrating changes over the age range yielded three age groups: 19–39, 40–59, and 60–75 years old. We also grouped our sample based on the level of education so as to reflect school requirements in Greece (compulsory education is 9 years): 1–9, 10–12 (high school), and 13 years and above (higher education including technological and other university level education). Finally, we calculated means, standard deviations, and percentile rankings by age and education level. In
order to estimate the convergent validity between the Greek TMT and the CTT in this group, we estimated Pearson’s correlation coefficient between the completion time of the two tests for both parts. The level of statistical significance was set at \( p = .05 \), and all analyses were conducted using the SPSS 15.0 software.

Clinical groups. We initially observed our data visually to determine whether the distributions met normality requirements. All data points that were considered outliers or extreme outliers were excluded from analyses (scores were considered outliers if they exceeded 2 SDs from the mean and extreme outliers if they exceeded 3 SDs from the mean). The normality assumption of our data was further investigated using the Kolmogorov–Smirnov test for normality. Most of our variables were normally distributed, so parametric tests were mainly employed. In order to investigate the equality of means, we used independent-samples \( t \)-tests for normally distributed variables, and the Mann–Whitney \( U \)-test for variables not normally distributed. In cases where statistically significant differences were found between the variances of groups, the \( t \)-test of unequal variances was used and the degree of freedom was estimated using the Welch–Satterthwaite approximation. Levene’s test was employed in order to investigate the equality of variances. The level of statistical significance was set at \( p = .05 \), and all analyses were conducted using the SPSS 15.0 software.

Results

Influence of Demographic Variables on the Completion Time of the CTT for the Nonclinical Group

In order to examine the potential contribution of demographic variables to performance (completion time of CTT1 and CTT2), we conducted a linear regression analysis. Results showed a main effect of all three demographic variables (gender, age, and educational level) on the completion time of CTT1—\( F(3, 159) = 275.016, p < .001 \), although the influence of gender was small (women performed relatively poorer compared with men). However, only age and educational level had a significant effect on the completion time of CTT2—\( F(2, 160) = 434.460, p < .001 \) (Table 2). Consequently, increased age and low educational level have a strong influence on the completion time for both parts of the CTT.

Given the significant contribution of age and education to the completion time for CTT, and in order to obtain normative data for the Greek adult population, we grouped our sample into demographic categories. Graphs illustrating changes over the age range yielded three age groups: 19–39, 40–59, and 60–75 years old. We also grouped our sample based on the level of education so as to reflect school requirements in Greece (compulsory education is 9 years): 1–9, 10–12 (high school), and 13 years and above (higher education including technological and other university level education). Table 3 presents means, standard deviations, and percentile rankings stratified by age and education level.

Convergent Validity in the Nonclinical Group

In order to estimate the convergent validity between the Greek TMT and the CTT in nonclinical participants, we estimated Pearson’s correlation coefficient between the completion time of the two tests for both parts. We found positive correlations of high significance between the completion time of part A of the Greek TMT and the CTT1 (\( r = .95, p < .001 \)) and the completion time of part B of the Greek TMT and the CTT2 (\( r = .96, p < .001 \)).

Table 2. Linear regression analyses: Contributions of age, education, and gender on completion time for the CTT (parts 1 and 2)

<table>
<thead>
<tr>
<th>Test</th>
<th>Variable</th>
<th>( B )-value</th>
<th>( t )-value</th>
<th>( p )-value</th>
<th>( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1</td>
<td>(Constant)</td>
<td>46.516</td>
<td>10.399</td>
<td>&lt;.001</td>
<td>.838</td>
</tr>
<tr>
<td>Completion time</td>
<td>Age</td>
<td>0.643</td>
<td>17.449</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>-2.273</td>
<td>-10.492</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>3.012</td>
<td>2.745</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>Part 2</td>
<td>(Constant)</td>
<td>108.130</td>
<td>14.817</td>
<td>&lt;.001</td>
<td>.844</td>
</tr>
<tr>
<td>Completion time</td>
<td>Age</td>
<td>1.135</td>
<td>16.948</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>-4.782</td>
<td>-11.913</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.053</td>
<td>1.675</td>
<td>.104</td>
<td></td>
</tr>
</tbody>
</table>

Note: CTT = Color Trails Test.
Criterion Validity in the Clinical Groups

In order to determine the ability of the CTT to distinguish clinical from normal samples, we compared the performance of 29 patients with PD to a group of 35 normal participants matched in age, educational level and gender. In order to examine whether there was a significant difference between the performance of the two groups in the completion time for CTT1 and CTT2, we used the parametric \( t \)-test for independent samples, since the data were normally distributed. Our results showed that patients with PD needed significantly more time to complete both CTT1 and CTT2 compared with the healthy participants (Table 4).

Similarly, we compared the performance of 25 patients who had recently suffered a stroke with a group of 26 healthy participants, matched in age, educational level, and gender. In order to examine performance of the two groups regarding completion time for CTT1 and CTT2, we used the parametric \( t \)-test for independent samples, as our variables were normally distributed. Our results indicated that patients required significantly more time to complete both CTT1 and CTT2 compared with the healthy participants (Table 5).

Discussion

Despite the widespread use of neuropsychological measures in clinical and research settings in Greece in recent years, normative data for commonly used neuropsychological tests remain largely unavailable. In an attempt to contribute toward filling this gap, we developed normative data for the Greek adult population of a useful and relatively simple paper-and-pencil neuropsychological test of visual attention, executive functions, and processing speed, stratified by those demographic characteristics that contributed significantly to performance on this test. We further provided data on the test’s validity in discriminating adult patients by comparing the time of completion scores of PD patients and acute stroke patients to separate age-, gender-, and education-matched healthy control groups. To our knowledge, there have been no attempts to date to develop normative data for the CTT in the adult population in Greece.
For the completion of part B of the standard TMT, participants need to know the English alphabets and low performance in part B of the TMT has often been connected to a lack of familiarity with the English alphabet, which is the case with illiterate individuals or individuals with a low educational level, as well as people who speak English as their second language (Strauss et al., 2006). Moreover, adults with language and reading disorders find it more difficult to complete the standard TMT (D’Elia et al., 1996). Thus, the CTT was originally created in order to avoid dependence from any alphabet (in any language), so as to release the test from any language barriers.

In this study, we noted a significant contribution of age and educational level in both parts of the CTT (older participants with a lower educational level completed both parts of the test slower than their counterparts). This finding is in agreement with the relatively few publications in the CTT literature (D’Elia et al., 1996; La Rue et al., 1999; Maj et al., 1993; Ponton et al., 1996; Sant’Ana Rabelo et al., 2010). Contrary to the literature, we found a weak influence of gender in part 1 of the test, however, not sufficient to justify the construction of norms based on gender. Further, we noted that healthy participants between the ages of 19 and 39 with university education had better completion times on CTT1 ($M = 34.80$ s, $SD = 4.92$) and CTT2 ($M = 67.54$ s, $SD = 10.21$) compared with the other stratified groups (Table 3). This finding is in accordance with the D’Elia and colleagues (1996) study who found that participants with higher education levels and aged between 18 and 29 had better performance on both the CTT1 and CTT2 compared with their other stratified groups. In contrast to the present study, the influence of gender in the D’Elia and colleagues (1996) study was not significant. A possible explanation for this finding was that their sample was primarily men as women compromised only 12% compared with 51.5% of women included in our study. In addition, this contradictory finding may be attributed to the fact that the women of our sample had a lower educational level compared with the male participants. Our findings are also in agreement with the more recent Sant’Ana Rabelo and colleagues (2010) study that noted differences between the various age groups and the worst results for ages >60.

Despite its frequent application for clinical and/or research purposes in different cultures, no information exists, comparing the original normative data provided in the CTT manual (D’Elia et al., 1996) with a new set of normative data. In this respect, our study provides evidence that the norms for healthy Greek adults are not entirely consistent with the manual norms for healthy adults residing in the USA. To be more specific, an example will be provided on this issue. For a healthy adult individual aged between 45 and 59 with an education level of between 9 and 11 years, a raw completion time of 55 s for the CTT1 and 122 s for the CTT2 ranks the performance on the 34th percentile in the original U.S. normative data (D’Elia et al., 1996). In contrast, a raw completion time of 55 s for the CTT1 and 122 s for the CTT2 in the same individual utilizing our normative data would place his performance on the 50th percentile (i.e., relatively better performance). Although our normative data cannot be directly compared with other available non-Greek norms, as the age and education stratification groups available in the literature are not directly comparable with our demographic group stratifications, this example provides some evidence that the CTT is not entirely culture fair. Therefore, the use of the CTT norms provided in the U.S. manual should be used with caution in non-U.S. samples and preferably should not be imported from one foreign context to another. We further suggest that the normative data provided here are utilized only within the context they were developed for, that is, adult native Greek speakers living in Greece with an age range 19–75 and an educational level of between 3 and 21 years.

With respect to the test’s validity in discriminating specific patient groups from nonclinical individuals, our data demonstrated significant differences in completion time for both parts of the CTT between a group of PD patients and a group of recent stroke patients compared with matched healthy controls. In both cases, the nonclinical participants had faster completion times compared with the clinical groups. These findings supplement the relevant literature, in which it has been reported that there is a significant delay in the completion time of CTT1 and CTT2 by patients with traumatic brain injuries (D’Elia et al., 1996), HIV-seropositive patients (Maj et al., 1993) and individuals with acquired brain injury who sought to obtain a drivers license (Hartman-Maeir et al., 2008). In general, TMTs and more specifically the TMT have proven very sensitive in detecting

### Table 5. Performance data for a nonclinical group and ischemic stroke patients demographically matched

<table>
<thead>
<tr>
<th></th>
<th>Nonclinical group ($N = 26$)</th>
<th>Stroke patients ($N = 25$)</th>
<th>Statistical testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean $\pm SD$</td>
<td>Mean $\pm SD$</td>
<td>Z-score</td>
</tr>
<tr>
<td>Age (years)</td>
<td>62.04 $\pm$ 8.83</td>
<td>65.08 $\pm$ 7.50</td>
<td>$Z = 1.692$</td>
</tr>
<tr>
<td>Education level (years)</td>
<td>9.77 $\pm$ 3.46</td>
<td>8.92 $\pm$ 3.68</td>
<td>$Z = 0.911$</td>
</tr>
<tr>
<td>Gender (men:women)</td>
<td>12:14</td>
<td>14:11</td>
<td>$\chi^2(1) = 0.494$</td>
</tr>
<tr>
<td>Completion time CTT1</td>
<td>71.12 $\pm$ 19.79</td>
<td>106.20 $\pm$ 22.33</td>
<td>$t(49) = 5.943$</td>
</tr>
<tr>
<td>Completion time CTT2</td>
<td>137.04 $\pm$ 43.29</td>
<td>226.44 $\pm$ 32.68</td>
<td>$t(49) = 8.298$</td>
</tr>
</tbody>
</table>

Note: CTT = Color Trails Test.
brain damaged patients in various clinical populations (Cicerone & Azulay, 2002; Reitan & Wolfson, 2004; Vlahou & Kosmidis, 2002). However, this study does not provide comparative data between the performance on the CTT and TMT in our clinical groups; therefore, the question of whether the CTT is more sensitive to central nervous system disturbance than the TMT in PD and ischemic stroke patients remains to be investigated.

Regarding the convergent validity between the Greek TMT and CTT, we found positive significant high correlations between the completion time of the two tests for both parts. This finding is partly in accordance with the Maj and colleagues (1993) report that noted significant but moderate correlations between the completion time of the TMT and CTT for both parts in a sample of 30 healthy German-speaking participants. Our findings are only partly in accordance with the Maj and colleagues (1993) report due to the significantly higher coefficients that were noted in our study, that is, .95 and .96, respectively, for parts A and B compared with .41 and .49 noted in the Maj and colleagues report. Further, our results are partly consistent with the Lee and Chan (2000) report that found high correlations (.72) between part B of the TMT and CTT2 in individuals of an older age (35–54) and higher education (12–22 years). Our healthy sample had an average age and education level within the ranges noted by this group; however, our study found higher correlations for both parts between the Greek TMT and CTT relative to the Lee and Chan (2000) report. We therefore also support the notion that these two tests assess corresponding cognitive fields mainly within specific education and age limits as previously concluded by Lee and Chan (2000). A possible explanation for the higher correlations noted between the two tests in our study is related to the contribution of language. It appears from our results that the impact of the Greek language on TMT performance (TMT B uses the Greek alphabets) is less significant than in English-speaking or Chinese-speaking individuals (Lee, Cheung, Chan, J & Chan, C, 2000).

In contrary, our findings regarding the convergent validity between the Greek TMT part B and CTT2 do not support those reported by Dugbarty and colleagues (2000) in a Turkish sample who concluded that the CTT and TMT are two qualitatively different due to the cognitive alternation, required in the CTT2 between numbers and colors.

In evaluating the generalizability of the present results, several limiting factors need to be considered. First, the age- and education-stratified subgroups utilized in the study were not balanced in size. This is usually the case, however, in normative studies involving participants from a broad age range who were recruited from the community in a sample of convenience (Mitrushina et al., 2005; Strauss et al., 2006). Indeed, if one could achieve random sampling from the community, one would expect certain population trends to be reflected in the sample, that is, a decrease in the level of education as a function of increasing age. This is especially true for the Greek population, as many elderly people educated 40–50 years ago have low levels of education, because they either did not attend school, left school early, or did not attend a university or other higher educational institution, mainly for socioeconomic reasons (i.e., war, poverty). As a result, the sample sizes for certain subgroups are small (e.g., for elderly, highly educated individuals and younger adults with limited education). Indeed, the absence of normative data for young adult participants in the 19–39-age group, with 1–9 years of education, is evident in our data. This caveat could not be avoided as we were unable to recruit participants with very low levels of education, due mainly to formal schooling requirements in Greece, making a minimum of 9 years of education compulsory.

A possible solution to small subgroup sizes is the use of broader age categories to increase the number of participants per subgroup. This strategy, however, may in turn cause problems related to the boundary values of the subgroups. Second, the relatively broad age range of our stratified subgroups may be a limiting factor, and this is especially true for the elderly participants. Elderly individuals typically show a more distinguishable pattern of performance decline with advancing age, and it would have been preferable to have used narrower groupings, for example, 60–65, 66–69, 70–75, and so on. Third, a lack of familiarity with neuropsychological assessment procedures, which differ from traditional medical procedures to which elderly individuals in Greece have become accustomed, may have also influenced our findings. Examiners were, however, well trained in the administration of the CTT and had previous experience with elderly research participants. Significant efforts were made in order to ensure that these participants understood all administration procedures, therefore minimizing this possible limitation. Another limitation concerns the risk of sampling bias associated with motivation to participate in this study. It would appear that healthy individuals willing to participate in the study are more motivated and possibly more curious about what a neuropsychological examination involves. It is also worth noting that Greek participants were not paid for their participation in this study, therefore enhancing the possibility that our sample was skewed in the direction of more high functioning individuals, and is a limitation that could have improved by payment of volunteers.

Additional investigations to establish the relationship of the CTT with other measures of attention, for example, Symbol Digit Modalities Test (Smith, 2002) and other verbal and non-verbal core neuropsychological measures is also important. Finally, the diagnostic utility of the test warrants further exploration in other clinical populations in Greece.

Conflict of Interest

None declared.