American Psychological Association

Psychological Services

1541-1559/18/$12.00 http://dx.doi.org/10.1037/ser0000307

Getting Closer? Differences Remain in Neuropsychological Assessments Converted to Mobile Devices

Áine Maguire
University of Cambridge and Trinity College Dublin

Jennifer Martin
University of Cambridge

Hannes Jarke
University of Cambridge and University of Vienna

Kai Ruggeri
Columbia University and University of Cambridge

Dementia is an increasing concern in today's aging society. Despite the limited evidence for dementia screening at a population level, a push to improve diagnosis and the expansion of technology usage within health-care settings has led to the rising popularity of computerized neuropsychological assessment devices (CNADs). Some CNADs are completely new tests, others are direct translations of traditional pen-and-paper cognitive functioning tests. This study is an investigation of the equivalence between two existing pen-and-paper tests and their translated versions on mobile platforms. In this small-scale study ($N = 42$), the scores on two multidomain assessments—Saint Louis University Mental State Examination (SLUMS; Feliciano et al., 2013) and Cambridge University Pen to Digital Equivalence assessment (CUPDE; Ruggeri, Maguire, Andrews, Martin, & Menon, 2016)—were significantly different, even with multiple design iterations, when participants were matched by age and score on an independent screening tool, the Self-Administered Gerocognitive Exam (SAGE), $t(13) = 2.55$, $p < .05$, $d = .680$. There was no relationship between the Color Trails Task (CTT; D'Elia, Satz, Uchiyama, & White, 1996; Maj et al., 1993) and its mobile translation, the electronic CTT (eCTT), $\rho = -.144$, $n = 21$, $p = .533$. Though no difference was identified between the eCTT and the modified pen-and-paper CTT (pCTT) scores, $t(13) = .092$, $p = .928$, there was no relationship between eCTT and pCTT, $r = .139$, $n = 14$, $p = .635$. Outcome scores of mobile-based assessments appear to remain distinct from the established norms of traditional assessments, adding to existing concerns associated with population-screening programs via mobile applications.

Keywords: cognitive neuropsychology in dementia, assessment of cognitive disorders/dementia, neuropsychological assessment, memory, digital dementia assessment

Early detection of dementia enables access to the care and support needed by patients. It allows patients the time to make critical decisions about their wishes for now and the future while they have the capacity to do so (Ashford et al., 2007). However, there is a need for a better and more consistent quality of diagnosis (Department of Health, 2009). Timely diagnosis at the stage in which the symptoms of dementia begin to have an impact on daily life is also required (Brooker, La Fontaine, Evans, Bray, & Saad, 2014). It is estimated that 9.9 million new cases of dementia are diagnosed annually around the world (Prince et al., 2015), and the total estimated worldwide cost for dementia is nearly $480 billion (Alzheimer's Society, 2014). It is thought that the financial burden presented by dementia could be greatly decreased with earlier detection of conditions such as mild cognitive impairment (MCI), which may precede the onset of dementia (Robinson, Tang, & Taylor, 2015). At present, many countries—such as the UK—do not have a national screening program for dementia, and a professionally administered screening test is only given on request during a scheduled doctor's appointment.

Pressure to improve diagnostic rates for dementia continues to grow, in spite of the lack of conclusive evidence for dementia screening at the population level (Fox et al., 2013; Uflacker & Doraiswamy, 2017). Computerized neuropsychological assessment devices (CNADs) are increasingly popular within the health-care sector as a potential method for meeting this demand. CNADs comprise of any digital interfaces (as opposed to human examiners) used to administer, score, or interpret tests of brain function

Áine Maguire, Policy Research Group, Department of Psychology, University of Cambridge and Department of Psychiatry, Trinity College Dublin; Jennifer Martin, Department of Engineering, University of Cambridge; Hannes Jarke, Policy Research Group, Department of Psychology, University of Cambridge and Faculty of Psychology, University of Vienna; Kai Ruggeri, Department of Health Policy and Management in the Mailman School of Public Health, Columbia University and Policy Research Group, Department of Psychology, University of Cambridge.

We thank Janine Jakobs, Gillian Coughlan, Clara Buettner, Amel Benzerga, and Josie Ruggeri.

Correspondence concerning this article should be addressed to Áine Maguire, Department of Psychiatry, Trinity Centre for Health Sciences, St James's Hospital, James's Street, Dublin 8, D08 W9RT, Ireland. E-mail: afmaguir@tcd.ie

and factors relating to neurological health or illness (Bauer et al., 2012). A number of empirically validated CNADs already exist and have been approved for clinical use, such as the Cambridge Neuropsychological Test Automated Battery (CANTAB, 2017) a mobile screening-assessment software for cognitive impairment, and HeadMinder Cognitive Stability Index (Erlanger et al., 2002), a web-based tool used for monitoring neurocognitive function. Some of these CNADs are novel and purpose-built, presenting some advantages over clinician-administered measures (Di Rosa et al., 2014). However, one systematic review of 11 computerized platforms for testing age-related changes in cognition or early symptoms of dementia highlighted a paucity of normative data for over 50% of the tests selected, with only three tests meeting the highest standards of reliability (Wild, Howieson, Webbe, Seelye, & Kaye, 2008). This absence of reliability data remains, and the increasing availability of measures, coupled with the nonstandardized reporting of outcomes, has prevented comparisons across studies (Aslam et al., 2018). Other screening tools are direct translations of traditional pen-and-paper cognitive function assessments, and are further cause for concern, considering that presently there are insufficient normative data and accompanying research on the validity of these translated tests (Zygouris & Tsolaki, 2015). For example, MOBI-COG (Nirjon, Emi, Mondol, Salekin, & Stankovic, 2014), a mobile translation of the 3-min Mini-Cog test (Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000) has been developed, but no comparative data are yet available.

One study highlighted concerns about the use of mobile-platform screening for dementia or MCI, particularly when used to translate traditional screening assessments to computerized formats (Ruggeri, Maguire, Andrews, Martin, & Menon, 2016). Significant differences in overall scores between traditional assessments and their mobile adaptations were recorded, even after adjusting analyses for design flaws in the mobile screening application developed (Ruggeri et al., 2016). This indicates that mobile translations of existing assessments may require the development of new normative standards and greater validation.

Building on previous work, in the current study, we addressed design issues of a mobile adaptation of a multidomain cognitive screening tool and incorporated a specific assessment of executive function, a core characteristic of vascular dementia, which is the second most common form of dementia (Román & Royall, 1999). Dysfunction of memory, the hallmark of Alzheimer's disease, is often the predominant focus of many current screening tests and other indicators of dementia, such as executive function, language, and praxis, are often overlooked (Cullen, O'Neill, Evans, Coen, & Lawlor, 2007). Thus a mobile measure of executive function would be well-placed in an era of CNADS, should equivalence be determined.

This between-groups study examines the equivalence of traditional and mobile-platform cognitive function assessments for MCI and dementia. The Self-Administered Gerocognitive Exam (SAGE; Scharre et al., 2010) was used to determine if there were any preexisting differences in cognitive function between participants. In this study, we sought to determine the following.

(a) Is the Cambridge University Pen to Digital Equivalence assessment (CUPDE; Ruggeri et al., 2016), a multidomain mobile-platform translation of an MCI screening tool, equivalent to its original pen-and-paper counterpart, the Saint Louis University Mental Status examination (SLUMS; Feliciano et al., 2013), following changes made to the initial CUPDE application?

(b) Is the electronic Color Trails Test (eCTT), an electronic tablet assessment of executive function, equivalent to its original pen-and-paper counterpart, the Color Trails Test (CTT; D'Elia, Satz, Uchiyama, & White, 1996; Maj et al., 1993), upon an as-direct-as-possible translation?

(c) Is the eCTT equivalent to a specifically modified pen-and-paper version of the CTT (pCTT)?

Method

Participants

A total of 42 volunteers aged between 55 and 71 years (M = 60.6, SD = 4.24) completed this study. Participants were required to be native English speakers, living independently and have normal or corrected vision and hearing. Those with a history or presence of memory complaints, psychiatric disorder, or neurological disease were excluded.

Color Trails Test

The CTT (D'Elia et al., 1996; Maj et al., 1993) was selected as an assessment of executive function because of its freedom from language and cultural biases. This two-part task requires participants to connect numbers first in ascending order from 1 to 25 (Part 1), then by alternating between numbers and two colors (i.e., 1-pink, 2-yellow, 3-pink, etc.) from 1 to 25 (Part 2) as quickly as possible using a pencil. The CTT possesses test–retest reliabilities between .64 to .79 (D'Elia et al., 1996). Although Part 1 of the test principally involves visual scanning and processing speed, Part 2 of the task requires additional attention and mental flexibility (Donoghue et al., 2012). Both parts of the test are timed separately by the test administrator. An index of executive function (set shifting) was determined by subtracting time to complete Part 1 of the test from time to complete Part 2 of the test (i.e., time to connect alternating color and number dots, minus time to connect only numbered dots; Bowie & Harvey, 2006; D'Elia et al., 1996). Lower scores indicate better performance.

A number of studies have determined the CTT to be valid for a variety of nationalities and age groups (e.g., Dugbartey, Townes, & Mahurin, 2000; Messinis, Malegiannaki, Christodoulou, Panagiotopoulos, & Papathanasopoulos, 2011; Rabelo et al., 2010), having found that it is a reliable and valid measure for assessment and comparison within as well as between populations. Specifically, Messinis et al. (2011) found that older adults were significantly slower at completing Part 2 while concluding that, even in a small sample, the CTT showed good convergent and criterion validity. This study examined the effects of mobile translation on this index.

Modified Colour Trails Test

In the original CTT (D'Elia et al., 1996), the participant is required to join up the numbers by drawing a continuous line in pencil. The motion closest to this on a mobile device requires the user to continuously swipe between the numbers to draw a line. However, because of programming challenges, swipe motions can be difficult to maintain on tablets for long periods of time, as needed for this test.

In response to this design challenge, the newly developed mobile eCTT requires that participants tap the numbers in order, then the program connects the numbers with lines. The eCTT records time to complete each task automatically. To test whether the changes—from joining the numbers to selecting them in order—fundamentally altered the test, a modified pCTT was adapted. In pCTT, the numbers and colors are selected in order by crossing through them, rather than joining with a continuous line. The difference in time taken to complete Part 1 and Part 2 was used as an index of executive function (set shifting) for both eCTT and pCTT. In this study, we compared these index scores against each other and scores obtained on the original CTT (D'Elia et al., 1996; Maj et al., 1993).

Saint Louis University Mental Status Examination

SLUMS (Feliciano et al., 2013) is an 11-item, reliable (Tariq, Tumosa, Chibnall, Perry, & Morley, 2006), multidomain screening tool for MCI that has a UK-specific version and has been used in previous research on this topic (Ruggeri et al., 2016). The SLUMS has been argued to be a better test than the most common cognitive battery used, the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975), based on better ability to detect mild neurocognitive disorder (Tariq et al., 2006), fewer ceiling effects (Feliciano et al., 2013), and generally better psychometric properties, particularly for older populations in multiple countries (Cruz-Oliver, Malmstrom, Allen, Tumosa, & Morley, 2012; Kaya et al., 2016).

Cambridge University Pen to Digital Equivalence Assessment

CUPDE (Ruggeri et al., 2016), the digital format of the SLUMS, was originally developed using JavaScript, HTML, and Cascading Style Sheets (CSS) code. Initial reliability and validity assessments of CUPDE showed significant differences from SLUMS, which catalyzed the need for improvements and further evaluation. For the present study, the application was recoded into the native iOS language, Swift, to remove limitations that resulted from coding in HTML. Those limitations included, among others, problems with Internet connection and accessing iPad features such as the microphone. Dictation technology was integrated, and user-interface design was adapted to generate an application as equivalent as possible to the traditional assessment.

Self-Administered Gerocognitive Exam

Paper-based SAGE (Scharre et al., 2010) was used to compare construct validity and potential differences in sensitivity between CUPDE (Ruggeri et al., 2016) and SLUMS (Scharre et al., 2010). SAGE is a reliable and sensitive measure for the identification of MCI from any cause, as well as early-onset dementia, demonstrating 95% specificity and 79% sensitivity for clinical diagnosis in a population over the age of 59 (Scharre et al., 2010). It is particularly useful in older populations and presents solid psychometric properties in its standard format and various adaptations (Scharre, Chang, Nagaraja, Vrettos, & Bornstein, 2017). This test was used to compare baseline cognitive function between the testing groups. Upon determining overall group comparisons between the assessment formats, participants were paired by age and scores on SAGE to draw more direct comparisons between the traditional and mobile-platform cognitive assessments.

Procedure

Participants were randomly assigned using a coin toss to complete Condition A, pCTT and the traditional format SLUMS (Feliciano et al., 2013; $n = 21$) administered by a researcher, or Condition B, eCTT and CUPDE (Ruggeri et al., 2016; $n = 21$), presented via iPad application. For validation and comparative purposes, both conditions completed the SAGE and the standard CTT (D'Elia et al., 1996; Maj et al., 1993). Testing took approximately 40 min for each condition.

The Engineering Design Centre and the Department of Psychology at the University of Cambridge provided ethical review, due to multiple department involvement. All participants provided informed consent at the beginning of assessment.

Analysis

Descriptive and inferential statistical testing was conducted using a level of significance at $p < .05$, two-tailed. Given the number of test items and the nature of the matched *t* test, our desired power was 60 per group for an overall of 120 participants. Having attempted these steps before (Ruggeri et al., 2016), it was considered unlikely that the ideal sample size be achieved and seeing as the actual sample size ($n = 42$) was primarily appropriate for nonparametric tests, no power calculation was done. Due to the small sample size, the Shapiro-Wilk test was conducted to test whether the data was normally distributed (Razali & Wah, 2011). Inferential statistical tests were used to test if there were any differences or relationships between measures. For data meeting the assumptions of normality, the Pearson's correlation coefficient and independent and paired *t* tests were used. For non-normally distributed data, Spearman's rank correlation and Mann–Whitney *U* and Wilcoxon signed-ranks test were used. All participants were utilized in these analyses. Equivalence between CUPDE (Ruggeri et al., 2016) and SLUMS (Feliciano et al., 2013) and eCTT and pCTT was examined by matching participants according to cognitive function (SAGE scores; Scharre et al., 2010) and age. A paired *t* test was then used to test for equivalence.

Results

Shapiro–Wilk normality test indicated that SAGE ($W = .901$, $p < .01$) and eCTT ($W = .903$, $p < .05$) scores were not normally distributed. As such, nonparametric analyses were used to test relationships and differences using SAGE and eCTT scores. Independent difference tests were used on all data to compare overall results between conditions (see Table 1).

Comparison of Baseline Cognitive Function on SAGE and CTT Between Conditions A and B

Mann–Whitney tests were conducted to determine whether baseline differences in cognitive function existed between participants in Condition A and Condition B. Results of the test indi-

Table 1
Descriptive Statistics and Comparisons for Measures

Measure	Statistics
	Mean (*SD*)
SAGE: Condition A	19.05 (2.89)
SAGE: Condition B	18.95 (2.54)
CTT: Condition A	1.75 (.35)
CTT: Condition B	1.98 (.62)
SLUMS: Condition A	25.29 (3.33)
CUPDE: Condition B	21.29 (3.94)
pCTT: Condition A	1.65 (.50)
eCTT: Condition B	1.71 (.62)
Correlations between measures	ρ
SAGE and SLUMS[a]	.608***
SAGE and CUPDE[b]	.486*
CTT and eCTT[b]	−.144
	r
CTT and pCTT[a]	.475*
Differences between measures	*t* (*df*)
SLUMS[a] and CUPDE[b]	–3.55 (40)***
CTT[a] and CTT[b]	1.50 (40)
	Mann–Whitney *U*
SAGE[a] and SAGE[b]	207.50
pCTT[a] and eCTT[b]	210.00

Note. SLUMS = Saint Louis University Mental Status Examination; CUPDE = Cambridge University Pen to Digital Equivalence assessment; SAGE = Self-Administered Gerocognitive Exam; CTT = Color Trails Test; pCTT = modified, pen-and-paper Color Trails Test; eCTT = electronic Color Trails Test.
[a] Condition A (paper measures): SLUMS and pCTT, $n = 21$. [b] Condition B (electronic measures): CUPDE and eCTT, $n = 21$.
* significant at .05 level. ** significant at .01 level. *** significant at .005 level.

cated that there were no significant differences between participants' SAGE scores (Scharre et al., 2010) in Condition A ($Mdn = 20$, $n = 21$) and Condition B ($Mdn = 19$, $n = 21$), $U = 207.5$, $z = -.331$, $p = .741$. A Mann–Whitney test also indicated no significant difference in standard CTT scores (D'Elia et al., 1996; Maj et al., 1993) between the two conditions, $U = 164$, $z = -1.42$, $p = .155$.

Comparison of CUPDE With SLUMS

A Spearman's rank correlation indicated that there was a moderate, positive association between CUPDE (Ruggeri et al., 2016) and SAGE scores (Scharre et al., 2010), $\rho = .486$, $n = 21$, $p < .005$, and a strong, positive association between SLUMS (Feliciano et al., 2013) and SAGE scores, $\rho = .608$, $n = 21$, $p < .05$.

Independent *t* tests demonstrated significant differences in total scores between SLUMS ($M = 25.29$, $SD = 3.33$) and CUPDE ($M = 21.29$, $SD = 3.94$), $t_{(40)} = -3.55$, $p < .001$. To control for potential group differences, paired *t* tests were conducted on participants from the two conditions after being matched for SAGE scores and age. These data met normality assumptions. When participants were matched on cognitive function via SAGE scores and age, CUPDE scores ($M = 22.29$, $SD = 3.81$) were significantly lower than SLUMS ($M = 25.21$, $SD = 3.62$) scores, $t_{(13)} = 2.55$, $p < .05$, $d = .680$.

Comparison of eCTT With CTT

A negative relationship between eCTT and standard CTT (D'Elia et al., 1996; Maj et al., 1993) was observed using Spearman's rank correlation, but this was not significant in Condition B, $\rho = -.144$, $n = 21$, $p = .533$. Wilcoxon signed ranks indicated significant differences between eCTT and standard CTT in Condition B, $z = -1.964$, $p < .05$.

Comparison of Modified pCTT With CTT

Pearson's correlation coefficient signified a moderate positive relationship between pCTT and standard CTT in Condition A, $r = .475$, $n = 21$, $p < .05$. A paired-samples *t* test demonstrated that there was no difference in scores between pCTT and standard CTT in Condition A, $t_{(20)} = -.926$, $p = .366$.

Comparison of eCTT With Modified pCTT

A Mann–Whitney test indicated no significant differences between eCTT ($Mdn = 1.61$, $n = 21$) and pCTT ($Mdn = 1.42$, $n = 21$) scores, $U = 210$, $z = -.264$, $p = .792$.

To control for potential group differences and determine the relationship between eCTT and pCTT, paired *t* tests were run on participants from the two conditions after being matched for SAGE scores (Scharre et al., 2010) and age. These data met normality assumptions. Though no difference was identified between eCTT and pCTT scores, $t_{(13)} = .092$, $p = .928$, there was also no relationship between eCTT and pCTT, as assessed with a Pearson correlation, $r = .139$, $n = 14$, $p = .635$.

Discussion

Despite addressing issues of design, the use of mobile-based dementia-screening tools appears to remain controversial, particularly regarding translation of traditional measures to mobile platforms. In this small-scale study, two multidomain assessments, CUPDE (Ruggeri et al., 2016) and SLUMS (Feliciano et al., 2013), could not be considered suitably equivalent, even with multiple design iterations, because of significant differences in total scores between the two measures. Despite between-groups analyses indicating no difference between the testing groups on overall cognitive function as measured by the SAGE (Scharre et al., 2010), significant differences were observed between scores on CUPDE and SLUMS. Although correlations with the SAGE had improved compared with previous research (Ruggeri et al., 2016), indicating moderate construct validity, CUPDE continued to demonstrate greater variability in results than SLUMS. This suggests that CUPDE, a mobile translation of SLUMS, is less sensitive as a screening tool to identify MCI than its standardized, original pen-and-paper counterpart. Results emphasize the need for new scoring approaches and standardizations for the mobile translations of multidomain assessments before comparisons or conclusions between studies and protocols can be established.

Complex challenges also exist for importing measures of individual cognitive function such as the CTT (D'Elia et al., 1996; Maj et al., 1993) onto mobile platforms. This study specifically focused on the translation of a measure of executive function (CTT). Results of between-groups analyses indicated that there was no true difference in executive function scores between the pen-and-paper (Condition A) and mobile-platform groups (Condition B), as measured by the traditional CTT. There was no relationship between eCTT, the mobile translation of CTT, and the original CTT, and there were significant differences in scores, with eCTT performance poorer than CTT performance. Due to programming restrictions, namely recording and presenting lines drawn in real time on the tablet, a modified pCTT was administered to assess whether differences in scores between the eCTT and the original CTT could be accounted for by the program restrictions of the eCTT. Though no difference in scores between the eCTT and pCTT was identified, there was also no relationship between these measures. Presented data suggest that, whereas differences in outcomes between the original CTT and a modified pen-and-paper format appear minimal, there was no relationship between the mobile version (eCTT) and either pen-and-paper versions (CTT or pCTT). This tentatively suggests that the modified version may not effectively measure the same cognitive functions, and determining differences in scores between the mobile and traditional measures may be redundant. However, closer translation between the CTT and the eCTT by joining the numbers using a stylus or swiping motion on the eCTT may help to minimize the discrepancies between the testing formats.

Researchers should consider extending future studies across a broader participant base. Presently, technology-based skills should not be assumed to be equivalent between elderly populations and digital natives, those who have grown-up in an age of technology (Prensky, 2001). In addition, gender should be considered, as this study's sample consisted of predominantly male participants. Although the intention within this study was to test individuals with no prior diagnosis of cognitive decline, it is pertinent to note that studying participants with a history of cognitive decline may have an even greater effect on the results. Such factors may increase the differences present in mobile cognitive function tests, or may even skew the results.

One key limitation of this study is the small sample size, which was a result of the limited availability of suitable participants. Individual differences such as familiarity with mobile technology, educational achievement, and premorbid IQ may exist between the two test conditions and act as confounding variables, limiting the reliability of the results. Thus, though the findings may indicate altered translation of screening measures from a traditional format to mobile platforms, larger, more adequately powered studies are required to draw robust conclusions. In addition, comparison between testing platforms using a single measure of executive function does not definitively determine the fallibility of mobile translations of paper-based, single assessments. Rather, the results highlight a need for careful measurement selection and translation, given the absence of relationships between the eCTT and the CTT (D'Elia et al., 1996; Maj et al., 1993), as well as between the eCTT and the modified version of the CTT, the pCTT. Alternative measures may prove to be more equivalent, and other domains of cognitive function affected by dementia should be considered.

"Mobile health" (mHealth) continues to gain favor, and mobile screening at a population level could help detect dementia and related disorders early on, allowing for timely intervention. Still, the results of this small study suggest that such screening has to be approached carefully, as both single assessments and multidomain screening tools of cognitive function may be altered in their translation. There can only be limited certainty that the translated measures used in this study reliably test the same cognitive functions as those determined by the traditional measures, particularly with regard to the CTT, considering that its mobile translation, the eCTT, held no relationship to the paper-based measures. Instrument specificity must be high in screening programs to avoid false negatives as much as possible and the sensitivity and specificity of each new CNAD, translated or purposefully developed, should be tested against existing, validated measures. Additional measures could also be taken into account for comprehensive assessment, even as mHealth screening will be better developed in the future. In this study, the CUPDE (Ruggeri et al., 2016) demonstrated improved correlation with another screening tool, the SAGE (Scharre et al., 2010), following redesign, compared with earlier research (Ruggeri et al., 2016). With continued design iterations and comprehensive neurocognitive assessments combined with neuroimaging methods, the true equivalence of mobile adaptations and traditional dementia screening tools may be reliably established. Should evidence eventually confirm that both approaches equally capture the same function and process, there will then be a need for the standardization and renorming of scores. Such work has been carried out with the SAGE over the period of this study (Scharre et al., 2017) and there will be tremendous benefit from further such attempts.

At present, scores derived from mobile adaptations in this study appeared distinct from the established norms of traditional assessments, which emphasizes the need for further investigation into the listed aspects before mobile-based applications can be considered for reliable population-level screening of MCI and ultimately, dementia.

References

Alzheimer's Society. (2014). *Dementia 2014: Opportunity for change*. Retrieved from https://www.alzheimers.org.uk/download/downloads/id/2317/dementia_2014_opportunity_for_change.pdf

Ashford, J. W., Borson, S., O'Hara, R., Dash, P., Frank, L., Robert, P., . . . Fillit, H. (2007). Should older adults be screened for dementia? It is important to screen for evidence of dementia! *Alzheimer's & Dementia, 3,* 75–80. http://dx.doi.org/10.1016/j.jalz.2007.03.005

Aslam, R. W., Bates, V., Dundar, Y., Hounsome, J., Richardson, M., Krishan, A., . . . Sikdar, S. (2018). A systematic review of the diagnostic accuracy of automated tests for cognitive impairment. *International Journal of Geriatric Psychiatry, 33,* 561–575. http://dx.doi.org/10.1002/gps.4852

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology, 27,* 362–373. http://dx.doi.org/10.1093/arclin/acs027

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15,* 1021–1027. http://dx.doi.org/10.1002/1099-1166(200011)15:11<1021::AID-GPS234>3.0.CO;2-6

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1,* 2277–2281. http://dx.doi.org/10.1038/nprot.2006.390

Brooker, D., La Fontaine, J., Evans, S., Bray, J., & Saad, K. (2014). Public health guidance to facilitate timely diagnosis of dementia: ALzheimer's COoperative Valuation in Europe recommendations. *International Journal of Geriatric Psychiatry, 29,* 682–693. http://dx.doi.org/10.1002/gps.4066

Cambridge Cognition. (2017). CANTAB [Cognitive assessment software]. Retrieved from www.cantab.com

Cruz-Oliver, D. M., Malmstrom, T. K., Allen, C. M., Tumosa, N., & Morley, J. E. (2012). The Veterans Affairs Saint Louis University Mental Status Exam (SLUMS Exam) and the Mini-Mental State Exam as predictors of mortality and institutionalization. *The Journal of Nutrition, Health & Aging, 16,* 636–641. http://dx.doi.org/10.1007/s12603-012-0098-9

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry, 78,* 790–799. http://dx.doi.org/10.1136/jnnp.2006.095414

D'Elia, L. F., Satz, P., Uchiyama, C. L., & White, T. (1996). *Color Trails Test: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Department of Health. (2009). *Living well with dementia: A national dementia strategy*. London, UK: Author.

Di Rosa, E., Hannigan, C., Brennan, S., Reilly, R., Rapčan, V., & Robertson, I. H. (2014). Reliability and validity of the Automatic Cognitive Assessment Delivery (ACAD). *Frontiers in Aging Neuroscience, 6,* Article 34. http://dx.doi.org/10.3389/fnagi.2014.00034

Donoghue, O. A., Horgan, N. F., Savva, G. M., Cronin, H., O'Regan, C., & Kenny, R. A. (2012). Association between timed up-and-go and memory, executive function, and processing speed. *Journal of the American Geriatrics Society, 60,* 1681–1686. http://dx.doi.org/10.1111/j.1532-5415.2012.04120.x

Dugbartey, A. T., Townes, B. D., & Mahurin, R. K. (2000). Equivalence of the Color Trails Test and Trail Making Test in nonnative English-speakers. *Archives of Clinical Neuropsychology, 15,* 425–431. http://dx.doi.org/10.1093/arclin/15.5.425

Erlanger, D. M., Kaushik, T., Broshek, D., Freeman, J., Feldman, D., & Festa, J. (2002). Development and validation of a web-based screening tool for monitoring cognitive status. *The Journal of Head Trauma Rehabilitation, 17,* 458–476. http://dx.doi.org/10.1097/00001199-200210000-00007

Feliciano, L., Horning, S. M., Klebe, K. J., Anderson, S. L., Cornwell, R. E., & Davis, H. P. (2013). Utility of the SLUMS as a cognitive screening tool among a nonveteran sample of older adults. *The American Journal of Geriatric Psychiatry, 21,* 623–630. http://dx.doi.org/10.1016/j.jagp.2013.01.024

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State:" A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Fox, C., Lafortune, L., Boustani, M., Dening, T., Rait, G., & Brayne, C. (2013). Screening for dementia: Is it a no brainer? *International Journal of Clinical Practice, 67,* 1076–1080. http://dx.doi.org/10.1111/ijcp.12239

Kaya, D., Isik, A. T., Usarel, C., Soysal, P., Ellidokuz, H., & Grossberg, G. T. (2016). The Saint Louis University Mental Status Examination is better than the Mini-Mental State Examination to determine the cognitive impairment in Turkish elderly people. *Journal of the American Medical Directors Association, 17,* 370.e11–370.e15. http://dx.doi.org/10.1016/j.jamda.2015.12.093

Maj, M., D'Elia, L., Satz, P., Janssen, R., Zaudig, M., Uchiyama, C., . . . Chervinsky, A. (1993). Evaluation of two new neuropsychological tests designed to minimize cultural bias in the assessment of HIV-1 seropositive persons: A WHO study. *Archives of Clinical Neuropsychology, 8,* 123–135. http://dx.doi.org/10.1093/arclin/8.2.123

Messinis, L., Malegiannaki, A. C., Christodoulou, T., Panagiotopoulos, V., & Papathanasopoulos, P. (2011). Color Trails Test: Normative data and criterion validity for the Greek adult population. *Archives of Clinical Neuropsychology, 26,* 322–330. http://dx.doi.org/10.1093/arclin/acr027

Nirjon, S., Emi, I. A., Mondol, M. A. S., Salekin, A., & Stankovic, J. A. (2014). *MOBI-COG: A mobile application for instant screening of dementia using the Mini-Cog Test* (pp. 1–7). New York, NY: ACM Press. http://dx.doi.org/10.1145/2668883.2668889

Prensky, M. (2001). Digital natives, digital immigrants: Pt. 1. *On the Horizon, 9,* 1–6. http://dx.doi.org/10.1108/10748120110424816

Prince, M., Wimo, A., Guerchet, M., Ali, G., Wu, Y., & Prina, M. (2015). *World Alzheimer report 2015: The global impact of dementia*. London, UK: Alzheimer's Disease International.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics, 2,* 21–33.

Robinson, L., Tang, E., & Taylor, J.-P. (2015). Dementia: Timely diagnosis and early intervention. *British Medical Journal, 350,* h3029. http://dx.doi.org/10.1136/bmj.h3029

Román, G. C., & Royall, D. R. (1999). Executive control function: A rational basis for the diagnosis of vascular dementia. *Alzheimer Disease and Associated Disorders, 13* (Suppl. 3), S69–S80. http://dx.doi.org/10.1097/00002093-199912003-00012

Ruggeri, K., Maguire, Á., Andrews, J. L., Martin, E., & Menon, S. (2016). Are we there yet? Exploring the impact of translating cognitive tests for dementia using mobile technology in an aging population. *Frontiers in Aging Neuroscience, 8,* Article 21. http://dx.doi.org/10.3389/fnagi.2016.00021

Sant'Ana Rabelo, I., Pacanaro, S. V., Rossetti, M. d. O., Leme, I. F. A. D. S., Castro, N. R. D., Güntert, C. M., . . . Lucia, M. C. S. D. (2010). Color Trails Test: A Brazilian normative sample. *Psychology & Neuroscience, 3,* 93–99. http://dx.doi.org/10.3922/j.psns.2010.1.012

Scharre, D. W., Chang, S. I., Murden, R. A., Lamb, J., Beversdorf, D. Q., Kataki, M., . . . Bornstein, R. A. (2010). Self-administered Gerocognitive Examination (SAGE): A brief cognitive assessment instrument for mild cognitive impairment (MCI) and early dementia. *Alzheimer Disease and Associated Disorders, 24,* 64–71. http://dx.doi.org/10.1097/WAD.0b013e3181b03277

Scharre, D. W., Chang, S. I., Nagaraja, H. N., Vrettos, N. E., & Bornstein, R. A. (2017). Digitally translated Self-Administered Gerocognitive Examination (eSAGE): Relationship with its validated paper version, neuropsychological evaluations, and clinical assessments. *Alzheimer's Research & Therapy, 9,* 44. http://dx.doi.org/10.1186/s13195-017-0269-3

Tariq, S. H., Tumosa, N., Chibnall, J. T., Perry, H. M., III, & Morley, J. E. (2006). Comparison of the Saint Louis University Mental Status Examination and the Mini-Mental State Examination for detecting dementia and mild neurocognitive disorder: A pilot study. *The American Journal of Geriatric Psychiatry, 14,* 900–910. http://dx.doi.org/10.1097/01.JGP.0000221510.33817.86

Uflacker, A., & Doraiswamy, P. M. (2017). Alzheimer's disease: An overview of recent developments and a look to the future. *Focus, 15,* 13–17.

Wild, K., Howieson, D., Webbe, F., Seelye, A., & Kaye, J. (2008). Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's & Dementia, 4,* 428–437. http://dx.doi.org/10.1016/j.jalz.2008.07.003

Zygouris, S., & Tsolaki, M. (2015). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias, 30,* 13–28. http://dx.doi.org/10.1177/1533317514522852

Received June 29, 2017
Revision received July 19, 2018
Accepted July 22, 2018 ■