

GUEST EDITORIAL

The clock-drawing test: a critique

Let me begin with a declaration of disinterest. The esteemed editor of *International Psychogeriatrics* commissioned this article after reading my letter to Another Journal (Philpot, 2003) in which I called for a moratorium on all clock-drawing test (CDT) research. It seemed to me that an awful lot of effort was being put into a fairly trivial issue and that researchers' time might be better spent. That much effort is being expended cannot be denied. Using the search words 'clock drawing', MEDLINE currently lists 177 publications, 91 since the turn of the century. The first reference to this specific form of drawing test comes in 1983 (Goodglass and Kaplan), and the first paper on validation is that of Shulman *et al.* (1986).

The CDT has emerged as the front runner from a range of drawing tests that have been used to investigate cognition over the last 40 years. When I trained at the Maudsley Hospital in the early 1980s we used the 'draw-a-cat' test and used what is now known as a 'naïve', informal scoring method to judge 'normality' or 'abnormality'. Cube-drawing used to be popular but is strongly influenced by educational and cultural background (Rosselli and Ardila, 2003). The problem of three-dimensionality and perspective also applies to the house-drawing test (Moore and Wyke, 1984), but is the specific focus of the 'draw-a-tree-lined-avenue' test (Rennert, 1971). The House-Person-Tree test can be used to assess intellectual function, as well as having a projective element (Fukunishi *et al.*, 1997).

The properties of any screening test should not only include ease of administration and acceptability but ease of scoring and, if used sensibly as a first stage in disease identification, there should be a high sensitivity and positive predictive value. Specificity and negative predictive value are less important in this respect, unless the screening method is itself lengthy, aversive or expensive, none of which apply to any of the forms of the CDT. It seems to this reviewer that, while the CDT is often proposed as a means of identifying patients with early stages of dementia who might benefit from specific treatment, it has largely been validated in groups of patients with established diagnoses. Or worse still, validated in patients whose 'dementia' is only established by scores on other screening tests.

Ken Shulman (2000) has written an excellent and balanced review of the CDT in which he emphasized that it should be used in conjunction with other broader tests of intellectual function, such as the Mini-mental State Examination (MMSE), the verbal fluency test and informant reports. My quibble here would

be that, if the test can't be used in isolation, then it is not really a screening test in its own right and is merely part of a battery, as can be said of the digit span test for example.

What have we learnt since Shulman's review? The following represents a selective review of recent publications.

Drawing and scoring methods

Shulman (2000) emphasized the virtue of simple scoring methods. He identified twelve different procedures for eliciting and scoring clock drawings that had been published up to 1998. These differ a good deal in complexity. At least three new methods have appeared. Manos (1999) has a 10-point scale, Royall and co-workers (1999) have developed a version (the CLOX) that examines more rigorously the executive function components and Lin and colleagues (2003) employ drawing and copying tasks with both complex and simple scoring methods. Many papers have compared the validity and reliability of scoring methods. Almost without exception, the sensitivity and specificity values are derived from comparisons of patients with well-established Alzheimer's disease (AD) and healthy controls matched for age and education. However, several studies have demonstrated that the CDT, however scored, is only modestly successful in identifying cases of mild dementia (Storey *et al.*, 2001; Powlishta *et al.*, 2002; Seigerschmidt *et al.*, 2002). Similarly, some doubts have been expressed about its usefulness in the community (Kirby *et al.*, 2001).

Why are there so many different scoring systems? Perhaps it is now the time to concentrate on two or three. The Mendez system (Mendez *et al.*, 1992) was identified as the most accurate discriminator by Storey *et al.* (2001) and Scanlan *et al.* (2002) among others. 'Naïve' judgment (abnormal or not) has been shown to be quite adequate for quick clinical purposes (Scanlan *et al.*, 2002). The simple method of Shulman and colleagues (1986) provides a happy medium.

Several authors have explored the added value, in terms of diagnostic discriminating power, that might be achieved by combining the CDT with other tasks. The Mini-Cog is a composite of a three-item recall test and the CDT (Borson *et al.*, 2000), and is reported to have high reliability and validity in identifying cases of dementia (Scanlan and Borson, 2001). Other combinations have been suggested with the Syndrom Kurz Test (Schramm *et al.*, 2002), the Short Cognitive Evaluation Battery (Robert *et al.*, 2003) and the MMSE (Heinik *et al.*, 2003). The CDT was originally part of the Boston Aphasia battery (Goodglass and Kaplan, 1983). There is a certain irony in advocating its return to use in combination with other tests, after all that has been written about it in the last twenty years.

Transcultural acceptability

Advocates of the CDT often cite the ease with which the test can be administered to patients from different cultures and languages. Successful validation studies have been carried out in Italy (Ravaglia *et al.*, 2003), non-English US citizens (Borson *et al.*, 1999), Thailand (Jitapunkul *et al.*, 2000), Brazil (Fuzikawa *et al.*, 2003), Mexico (Royall *et al.*, 2003) and Hong Kong (Lam *et al.*, 1998). However, Storey *et al.* (2002) found only modest sensitivity and low specificity in a culturally diverse group in Australia.

Discrimination between AD and other dementias

Sensitivity and specificity values are more modest in the studies comparing groups of patients with different dementias. These include Cahn-Weiner *et al.* (2003), comparing AD with Parkinson's disease dementia, and dementia with Lewy bodies, and Heinik *et al.* (2002) comparing AD with vascular dementia. Rascovsky *et al.* (2002) and Moretti *et al.* (2002) both found that patients with fronto-temporal dementia (FTD) performed better on the CDT than AD patients matched for severity, whereas the FTD patients were worse at verbal fluency tests. Here my concern is whether the CDT score adds anything to the usual diagnostic process, clinical history and supporting investigations. What are presented in these studies are comparisons of group means rather than the predictive value of a particular score.

Neurological localization

Some researchers have argued, largely based on correlational data with other test methods, that the CDT is chiefly a test of executive function (Royall *et al.*, 1999). However, brain imaging studies do not support this view. Using a variety of techniques CDT scores have been associated with ventricular enlargement (Heinik *et al.*, 2000), grey matter volume in the non-dominant temporal lobe (Cahn-Wiener *et al.*, 1999) and increased cerebral blood flow in the dominant posterior temporal regions (Ueda *et al.*, 2002). Ino and colleagues (2003) found that actual performance of the CDT was associated with increased activity on functional magnetic resonance imaging in the posterior parietal cortices.

Calling time

While many authors do sensibly couch their enthusiasm for the CDT in qualifying statements about its limitations and its proper use in conjunction with other tests, many do not. Clearly it is essential for us to have adequate screening methods for dementia that are quick and easy to administer, and that

examine the range of disabilities with which dementia might present. My original call for a moratorium was motivated in part by an altruistic urge to encourage colleagues to focus their minds on something more robust than the CDT (or ‘get a life’ as modern parlance has it). With the exception of the interesting but inconsistent brain imaging studies, I am not impressed that we have learnt much more since Shulman’s review (2000). Perhaps now is the moment to ‘Stop all the clocks, cut off the mobile phone . . .’ (to misquote W. H. Auden . . . again).

MICHAEL PHILPOT
 Consultant Psychiatrist
 Maudsley Hospital
 London
 U.K.
 Email: mike.philpot@slam.nhs.uk

References

- Borson, S., Brush, M., Gil, E., Scanlan, J., Vitaliano, P. et al.** (1999). The Clock-Drawing Test: utility for dementia detection in multiethnic elders. *Journal of Gerontology (A): Biological Science and Medical Science*, 54, M534–M540.
- Borson, S., Scanlan, J., Brush, M., Vitaliano, P. and Dokmak, A.** (2000). The Mini-Cog: a cognitive ‘vital signs’ measure for dementia screening in multilingual elderly. *International Journal of Geriatric Psychiatry*, 15, 1021–1027.
- Cahn-Weiner, D. A., Sullivan, E. V., Shear, P. K., Fama, R., Lim, K. O. et al.** (1999). Brain structural and cognitive correlates of clock-drawing performance in Alzheimer’s disease. *Journal of the International Neuropsychology Society*, 5, 502–509.
- Cahn-Weiner, D. A., Williams, K., Grace, J., Tremont, G., Westervelt, H. and Stern, R. A.** (2003). Discrimination of dementia with Lewy bodies from Alzheimer disease and Parkinson disease using the Clock-Drawing Test. *Cognition and Behavioural Neurology*, 16, 85–92.
- Fukunishi, I., Mikami, N. and Kikuchi, M.** (1997). Alexithymic characteristics in responses to the Synthetic House-Tree-Person (HTP) Drawing Test. *Perceptual and Motor Skills*, 85, 939–942.
- Fuzikawa, C., Lima-Costa, M. F., Uchoa, E., Barreto, S. M. and Shulman, K.** (2003). A population-based study on the intra- and inter-rater reliability of the Clock-Drawing Test in Brazil: the Bambui Health and Ageing Study. *International Journal of Geriatric Psychiatry*, 18, 450–456.
- Goodglass, H. and Kaplan, E.** (1983). *The Assessment of Aphasia and Related Disorders*. Lea and Feiber: Philadelphia.
- Heinik, J., Reider-Groswasser, I. I., Solomesh, I., Segev, Y. and Bleich, A.** (2000). Clock-drawing Test: correlation with linear measurements of CT studies in demented patients. *International Journal of Geriatric Psychiatry*, 15, 1130–1137.
- Heinik, J., Solomesh, I., Raikher, B. and Lin, R.** (2002). Can Clock-drawing Test help to differentiate between dementia of the Alzheimer’s type and vascular dementia? A preliminary study. *International Journal of Geriatric Psychiatry*, 17, 699–703.
- Heinik, J., Solomesh, I., Bleich, A. and Berkman, P.** (2003). Are the Clock-drawing Test and the MMSE combined interchangeable with CAMCOG as a dementia evaluation instrument in a specialized outpatient setting? *Journal of Geriatric Psychiatry and Neurology*, 16, 74–79.

- Ino, T., Asada, T., Ito, J., Kimura, T. and Fukuyama, H.** (2003). Parieto-frontal networks for clock drawing revealed with fMRI. *Neuroscience Research*, 45, 71–77.
- Jitapunkul, S., Worakul, P. and Kiatprakoth, J.** (2000). Validity of clinical use of the Clock-drawing Test in Thai elderly patients with memory problems. *Journal of the Medical Association of Thailand*, 83, 342–347.
- Kirby, M., Denihan, A., Bruce, I., Coakley, D. and Lawlor, B. A.** (2001). The Clock-drawing Test in primary care: sensitivity in dementia detection and specificity against normal and depressed elderly. *International Journal of Geriatric Psychiatry*, 16, 935–940.
- Lam, L. C., Chiu, H. F., Ng, K. O., Chan, C., Chan, W. F. et al.** (1998). Clock-face drawing, reading and setting tests in the screening of dementia in Chinese elderly adults. *Journal of Gerontology (B). Psychological Science and Social Science*, 53, P353–P357.
- Lin, K. N., Wang, P. N., Chen, C., Chiu, Y. H., Kuo, C. C. et al.** (2003). The three-item Clock-drawing test: a simplified screening test for Alzheimer's disease. *European Neurology*, 49, 53–58.
- Manos, P. J.** (1999). Ten-point clock test sensitivity for Alzheimer's disease in patients with MMSE scores greater than 23. *International Journal of Geriatric Psychiatry*, 14, 454–458.
- Mendez, M. F., Ala, T. and Underwood, K.** (1992). Development of scoring criteria for the clock drawing task in Alzheimer's disease. *Journal of the American Geriatrics Society*, 40, 1095–1099.
- Moretti, R., Torre, P., Antonello, R. M., Cazzato, G. and Bava, A.** (2002). Ten-point clock test: a correlation analysis with other neuropsychological tests in dementia. *International Journal of Geriatric Psychiatry*, 17, 347–353.
- Moore, V. and Wyke, M. A.** (1984). Drawing disability in patients with senile dementia. *Psychological Medicine*, 14, 97–105.
- Philpot, M.** (2003). Stop all the clocks. *International Journal of Geriatric Psychiatry*, 18, 89–90.
- Powlishta, K. K., Von Dras, D. D., Stanford, A., Carr, D. B., Tsering, C. et al.** (2002). The clock drawing test is a poor screen for very mild dementia. *Neurology*, 59, 898–903.
- Rascovsky, K., Salmon, D. P., Ho, G. J., Galasko, D., Peavy, G. M. et al.** (2002). Cognitive profiles differ in autopsy-confirmed fronto-temporal dementia and AD. *Neurology*, 58, 1801–1808.
- Ravaglia, G., Forti, P., Maioli, F., Arnone, G., Pantieri, G. et al.** (2003). The clock-drawing test in elderly Italian community dwellers: associations with socio-demographic status and risk factors for vascular cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 16, 287–295.
- Rennert, H.** (1971) [The tree-lined avenue drawing test. Preliminary report on a method to diagnose space drawing abnormalities in mental patients] *Psychiatrie, Neurologie und Medizinische Psychologie (Leipzig)*, 23, 601–609.
- Robert, P. H., Schuck, S., Dubois, B., Olie, J. P., Lepine, J. P. et al.** (2003). Screening for Alzheimer's disease with the short cognitive evaluation battery. *Dementia and Geriatric Cognitive Disorders*, 15, 92–98.
- Roselli, M. and Ardila, A.** (2003). The impact of culture and education on non-verbal neuropsychological measurements: a critical review. *Brain and Cognition*, 52, 326–333.
- Royall, D. R., Mulroy, A. R., Chiodo, L. K. and Polk, M. J.** (1999). Clock drawing is sensitive to executive control: a comparison of six methods. *Journal of Gerontology (B) Psychological Science and Social Science*, 54, P328–P333.
- Royall, D. R., Espino, D. V., Polk, M. J., Verdeja, R., Vale, S. et al.** (2003). Validation of a Spanish translation of the CLOX for use in Hispanic samples: the Hispanic EPESE study. *International Journal of Geriatric Psychiatry*, 18, 135–141.
- Scanlan, J. and Borson, S.** (2001). The Mini-Cog: receiver operating characteristics with expert and naïve raters. *International Journal of Geriatric Psychiatry*, 16, 216–222.

- Scanlan, J. M., Brush, M., Quijano, C. and Borson, S.** (2002). Comparing clock tests for dementia screening: naïve judgments vs. formal systems—what is optimal? *International Journal of Geriatric Psychiatry*, 17, 14–21.
- Schramm, U., Berger, G., Muller, R., Kratzsch, T., Peters, J. et al.** (2002). Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population. *International Journal of Geriatric Psychiatry*, 17, 254–260.
- Seigerschmidt, E., Mosch, E., Siemen, M., Förstl, H. and Bickel, H.** (2002). The clock drawing test and questionable dementia: reliability and validity. *International Journal of Geriatric Psychiatry*, 17, 1048–1054.
- Shulman, K. I.** (2000). Clock-drawing: is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*, 15, 548–561.
- Shulman, K. I., Shedletsky, R. and Silver, I. L.** (1986). The challenge of time: clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, 1, 135–140.
- Storey, J. E., Rowland, J. T., Basic, D. and Conforti, D. A.** (2001). A comparison of five clock scoring methods using ROC (receiver operating characteristic) curve analysis. *International Journal of Geriatric Psychiatry*, 16, 394–399.
- Storey, J. E., Rowland, J. T., Basic, D. and Conforti, D. A.** (2002). Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients. *International Psychogeriatrics*, 14, 259–271.
- Ueda, H., Kitabayashi, Y., Narumoto, J., Nakamura, K., Kita, H. et al.** (2002). Relationship between clock-drawing test performance and regional cerebral blood flow in Alzheimer's disease: a single photon emission computed tomography study. *Psychiatry and Clinical Neuroscience*, 56, 25–29.