DIFFERENTIAL SCORING PATTERNS ON THE CLOCK DRAWING TEST: A COMPARISON OF VASCULAR DEMENTIA AND ALZHEIMER'S DEMENTIA

Alaina Everitt, B.A.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2006

APPROVED:

Charles Guarnaccia, Major Professor
Jerry McGill, Committee Member
Andrew Houtz, Committee Member
Linda Marshall, Chair of the Department of Psychology
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

Everitt, Alaina*, Differential scoring patterns on the clock drawing test: A comparison of vascular dementia and Alzheimer's dementia.* Master of Science (Psychology), May 2006, 19 pp., 1 table, 27 titles.

This study examined differences in scoring patterns among those diagnosed with Alzheimer's dementia and vascular dementia on the clock-drawing test. Archival clock drawing data was retrieved on 279 patients presenting at a county hospital-based memory clinic. Analysis of drawings was based on frequency of qualitative errors, as well as an overall quantitative score. Mean comparisons found those patients with Alzheimer's dementia to perform worse on both quantitative and qualitative scoring measures. However, Pearson's chi-squared test revealed a significantly higher rate of spacing errors among subjects with vascular dementia. Such lends support to my hypothesis that impaired executive functioning in vascular dementia patients would lead to poor qualitative performance. Logistic regression found significant predictive ability for the qualitative criteria in diagnosis ($\chi^2 = 25.49$, $p < .001$), particularly the rate of omission ($z = 8.96$, $p = .003$) and addition errors ($z = 7.58$, $p = .006$). Such findings hold important implications for the use of qualitative criteria in cognitive screening assessments.

TABLE OF CONTENTS

INTRODUCTION

As the number of elderly Americans continues to grow (Federal Interagency Forum on Aging-Related Statistics, 2004), clinicians are seeing accelerated rates of dementias and cognitive disorders in this aging population (National Institute on Aging, 2001). Despite the influx of new cases, there is, as of yet, no definitive ante mortem diagnostic technique for Alzheimer's dementia (AD). Physicians continue to rely on biological correlates of the disease such as post-mortem plaque counts, and levels of cholineacetyltransferase in the brain to provide them with a direction for treatment (Sunderland et al., 1989). Despite the recent availability of magnetic resonance imagining (MRI) technology able to image amyloidal plaques previously too small to be viewed in living subjects (Mayo Foundation, 2005); functional and clinical assessment tools remain invaluable supportive evidence of cognitive decline. In particular, cognitive screening of geriatric populations has become a popular means for early identification of dementia. The use of clock drawing is one such screening device.

Clock drawing, though typically incorporated into neurocognitive assessment batteries, demonstrates one of the most useful likelihood ratios (likelihood of a disease given specific test results) in diagnosing dementia. In Shulman's review of the literature (2000), he cites research by Siu (1991) that reports a likelihood ratio of 24 for abnormal clock drawings and .2 for normal representations. Clock drawing tests also have well-established correlations with other cognitive tests such as the Mini Mental Status Exam (MMSE) (Brodaty & Moore, 1997; Juby, 1999; Mendez et al, 1992) the Blessed Dementia Scale (Brodaty & Moore, 1997), the Global Deterioration Scale (Mendez et al., 1992), and the Hierarchical Dementia Scale (Shulman, 2000). Schramm et al. (2000) noted that the MMSE, one of the most popular screening tools, has only a 20% sensitivity rate in detecting milder forms of dementia. Furthermore, it serves as a measure of verbal skills with no assessment of visuoconstructive ability, which is one of the earliest areas of dysfunction in patients later diagnosed with dementia (Ihl et al., 1992; Schramm et al., 2000).

The heavy reliance on language demonstrated by many cognitive screening tests makes them subject to influences of education and background (Silverstone et al., 1993). In addition, scoring for clock drawing tests does not depend upon speed of performance, as do other tests such as the Short Performance Test (Erzigkeit, 1989), and is thus less confounded by the presence of depression (Schramm et al., 2002). Despite the existence of several scoring methodologies, clock drawing tests have consistently shown general sensitivity rates of 81% to 95%, an above-average percentage for cognitive screening tools (Schramm et al., 2002).

The variability in diagnostic sensitivity rate depends upon the samples studied (clinical vs. general populations) as well as the type of scoring criteria utilized. Heinik et al. (2002) has identified over fifteen different original scoring systems for clock drawing, all of which exhibit exceptional inter-rater reliabilities of .90 to .95 (Tuokko, 1992). However, the true differences in scoring methods lie in their ability to differentiate milder forms of dementia from normal aging (Manos, 1999). Early detection of dementia holds serious implications for treatment, as it is now the case that new classes of memory-enhancing drugs can greatly improve the quality of life for those diagnosed and their caregivers (Simard & van Reekum, 1999).

Qualitative scoring methods for the clock drawing tests were some of the earliest to evolve. Shulman (1986) devised a five-point system to rank clocks according to severity level with one point assigned to those with minor errors. Even prior to Shulman's scoring protocol, Goodglass and Kaplan (1983) were evaluating clock drawings on a similar qualitative principle. However, Goodglass and Kaplan used criteria with more objectively defined definitions.

A ten-point scoring system outlined by Manos and Wu (1994) shows promise in its ability to detect milder dementia. As previously noted, the sensitivity of the MMSE in diagnosing mild dementia is approximately twenty percent while Manos and Wu's Clock Drawing Test (1994) demonstrates 77% sensitivity. Manos and Wu's scoring methodology divides a clock face into eighths and assigns a point for the positioning of numbers 1,2,4,5,7,8,10 and 11 in the correct segments with two additional points for correct minute hand placement at ten minutes after eleven (Manos, 1999). These researchers found that by using a cut-off score of eight points, a clinician would be able to correctly diagnose delirium (92% sensitivity), dementia (90%) and adjustment disorders (5%). Moretti et al. (2002) also found high sensitivity for the ten-point clock test, as it correctly identified 71% of Alzheimer's patients and 77% of those patients diagnosed with milder dementia. These findings are important, particularly when coupled with the brevity, ease of scoring, practicality and acceptability of the task, for detection and differentiation of vascular dementia, often seen in younger patients, from Alzheimer's dementia.

The current study examined differences in scoring patterns on clock drawing between patients diagnosed with either Alzheimer's or vascular dementias. Previous studies have found subjects with vascular dementia (VaD) to have better-preserved insight and naming ability than those with Alzheimer's dementia. However, executive functioning was considerably more deteriorated (Inzitari & Pantoni, 1998). Given clock drawing's heavy reliance on executive function, particularly goal-directed behavior with ambiguous instructions (Royall, 1996), one might expect patients with vascular dementia to have poorer performance on this screening tool. This however, has not been the case in previous research. Heinik et al. (2002) found significantly poorer performance among those with vascular dementia, only with regard to clock-hand placement. Similarly, Moretti et al. (2002) found overall scores for vascular dementia patients to be significantly better than for Alzheimer's dementia counterparts (AD, M = 2.96, SD = 1.86; VaD, M = 4.23, SD = 2.01; difference between group means $p < .05$).

However, neither the Heinik et al. (2002) study nor the Moretti et al. (2002) study examined qualitative differences in clock drawing that may likely be affected by executive dysfunction. The present study hoped to clarify the performance differences proposed by these authors by incorporating a comparison of qualitative scoring as outlined by Goodglass and Kaplan (1983), and Shulman (2000). It was hypothesized that though quantitatively, vascular dementia patients would perform better than those with Alzheimer's dementia, the qualitative scores would suffer as a result of their impaired executive functioning. Such findings would offer additional support clock drawing in its use as a cognitive screening tool for all types of dementia.

METHOD

Participants

Archival data was retrieved on 279 patients referred for neuropsychological screening at the John Peter Smith County Hospital, Memory Clinic in Fort Worth, TX. Patients were selected based on age and availability of demographic information. No subjects over the age of 92 were used, and the sample ranged in age from 45 to 92 with a mean of 70.66 (*SD* = 7.92) years. Approximately 79% of the subjects were female (*n* = 220) and 21% male (*n* = 59). African American individuals made up the largest ethnic/racial category at 46% (*n* = 128) followed by Caucasian (38%) (*n* = 126), and then Hispanic (13%) (*n* = 36) individuals. Seven percent of participants (*n* = 19) spoke a language other than English. Though clinically trained translators were available during assessment, research has demonstrated little effects of ethnicity on clock drawing given its reliance on universal symbols of time (Borson et al., 1999). Diagnoses were based on criteria developed by the International Neuropsychological Society (INS) (Roman et al., 1993) and patients were diagnosed by a medical professional prior to data retrieval and analysis.

Procedure

Prior to retrieval of data, the present research was approved by the Institutional Review Boards of the University of North Texas and John Peter Smith Hospital. In addition, consent forms were signed by all participants during initial intake as customary procedure for John Peter Smith Hospital. After the brief intake interview including primarily demographic information and medical/psychiatric history, patients were given a blank sheet of paper and asked to draw the face of a clock set at ten past eleven. Such instructions allowed for an evaluation of circle quality as based on a collaboration of criteria proposed by Shulman (2000), and by Goodglass and Kaplan (1983). Circles were rated as either poor quality or acceptable quality using these criteria and assigned into appropriate categorical groups for analysis. In addition, Manos and Wu (1994) scoring criteria were applied according to the authors' original standards, resulting in a score of one to ten, with ten representing a perfect clock drawing. Additional qualitative criteria were added based on work done by Tuokko et al. (1990) that identified various objective error types including additions and omissions as well as more subjectively defined micro/ macrographic errors and spacing errors not detected through Manos and Wu (1994) criteria. Analyzed scores for such errors represented frequency counts of the individual error categories.

RESULTS

Two clinicians, blind to diagnosis, scored 52 clock drawings independently. Inter-rater reliability for quantitative scoring among this data subset was excellent at .9 similar to analyses in previous studies (Tuokko, 1992). Reliabilities for qualitative scoring were more modest and ranged from .25 (micro/macrographia) to .81 (omissions).

Independent sample *t*-test comparison of means ($n = 117$) yielded some significant differences in various scoring components between subjects diagnosed with vascular and Alzheimer's dementias (see Table 1 for mean scores by diagnosis, and the mean difference tests). Version 11.5 of SPSS was used to analyze data. The results of the Levene's test revealed significant differences for spacing ($F = 6.25$, $p < .05$) and micro/macrographia variables ($F = 6.48$, $p < .05$), indicating possible heterogeneity of variance. Therefore, the non-pooled values were examined for each of the variables in violation of this assumption.

The results of the independent samples t-test found significant differences for the frequency of omission errors ($t = 3.54$, $df = 115$, $p = .001$ two-tailed), incorrect time ($t = 2.37$, $df = 115$, $p = .02$ two-tailed) and the Manos quantitative score ($t = 3.13$, $df = 115$, $p = .002$ two-tailed). Because incorrect time is an integral component of the Manos score, this variable was interpreted only with regard to quantitative differences in diagnoses. Those patients with vascular dementia appeared to perform better than those with Alzheimer's dementia with regard to overall quantitative scoring (Manos score). Similar performance was noted for the majority of qualitative variables.

However, spacing errors were significantly more frequent among vascular dementia subjects as evidenced by Pearson's chi-square test ($\chi^2 = 24.31$, $df = 6$, $p < .001$).

Though the hypothesis suggesting greater frequencies of qualitative errors among the vascular dementia sample was not entirely supported, it is interesting to note the problematic spacing issues in these subjects given the previously mentioned link between vascular dementia and executive functioning. These spacing errors likely represent difficulties with planning and goal-directed behavior; hallmarks of impaired executive function.

A supplemental analysis was performed to determine the ability of patients' qualitative and quantitative scores to predict diagnostic category. A direct logistic regression was performed between the dependent variable, diagnostic category, and quantitative and qualitative scores as independent variables or predictors. Continuous data (Manos score) appeared normal ($z = .8$) therefore, transformations were unnecessary. Correlations between variables were low ranging from -.018 to .326 and there were no evident problems with multicollinearity, as tolerance values all fell below 1.0. A test of the full model demonstrated significant predictive ability as compared to the null or constant–only model as χ^2 (8, $n = 55$) = 26.12, $p = .001$, indicating that both quantitative and qualitative scoring, together, reliably differentiated diagnostic categories. The variance accounted for in diagnosis based on this model was substantial as Nagelkerke's $R^2 = .518$, an overall effect of 51.8%. In terms of predictive ability the model performed exceptionally well, correctly classifying 88.6 % of those subjects diagnosed with vascular dementia and 70% of those with Alzheimer's dementia, resulting in an overall success rate of 81.8%.

In a second regression analysis, diagnostic category was analyzed as the dependent variable as predicted by only the qualitative scores (additions, omissions, spacing, circle quality, and micro/macrographia). Results indicated that the model retained its significant predictive ability, χ^2 (6, *n* = 55) = 25.49, *p* < .001, regardless of having left out quantitative scores. Furthermore, this combination of variables accounted for 50.8% of the variance in diagnosis; only 1% less than the full model. Predictive accuracy remained high as well, with successful classification of subjects identical to that of qualitative and quantitative scores (81.8%). Examination of individual variable contributions according to the Wald criterion, found only frequency of omission errors, (*z* = 8.96, *p* = .003) and addition errors (*z* = 7.58, *p* = .006) to reliably differentiate diagnostic categories. This finding is partially supported by the comparison of means yielding significant differences between diagnoses in the frequency of omissions.

DISCUSSION

This research attempted to identify differential scoring patterns on the clock drawing test among those patients diagnosed with vascular and Alzheimer's dementia. Previous research has found certain performance differences among these dementia subtypes (Heinik et al., 2002) yet no study has examined both quantitative scoring systems as well as specific qualitative criteria. A combination of scoring methodologies taken from Tuokko et al. (1990), Shulman (1986), Manos and Wu (1994), and Goodglass and Kaplan (1983) allowed us to look more specifically at differences in functioning. Frequencies of omission and addition errors were calculated, as well as spacing problems, circle quality and micro/macrographia. It was hypothesized that the presence of these error types could be attributed to frontal dysfunction, particularly executive functions, one of the earliest areas to be affected by vascular dementia (Inzitari & Pantoni, 1998). Though Alzheimer's dementia is considered a more advanced and debilitating form of dementia, we expected to see poorer performance overall, in terms of quantitative scoring, but better performance on the specific qualitative criteria. The results of this research supported this hypothesis while also providing details on the particular strengths of certain qualitative variables in distinguishing patients with vascular dementia from those with Alzheimer's dementia.

Significant differences were found in both qualitative and quantitative scoring criteria as vascular dementia patients demonstrated significant difficulties with the spacing of numbers on the clock face while outperforming Alzheimer's dementia patients on other qualitative measures and the overall Manos score. Thus, it is likely that the more severe executive dysfunction in the vascular subjects played a role in their clock drawing performance. Due to the relatively small sample size and the categorical, frequency-related variables, a logistic regression was performed on the data to supplement findings from the mean comparisons. Analysis took into account the unequal sample sizes in each category and based probabilities off specific group numbers. Not only were differences in scoring patterns found between these two populations, but it was also discovered that such differences are strong enough to allow for prediction and classification of protocols. Using all scoring elements (qualitative and quantitative), prediction accuracy was excellent at 81.8%. However a closer look at individual contributions of variable combinations saw no decrease in accuracy when using only qualitative variables. Though other researchers might explain this finding in terms of the strong relationship between components of the Manos score and the individual qualitative criteria, we found no issues with multicollinearity. Furthermore, the specific error types, omissions and additions, proved to be the strongest predictors in both models, providing strong support for their use in assessment and diagnosis.

Despite our important findings, several methodological weaknesses must be mentioned. First, in classifying subjects into diagnostic groups, several diagnoses were included in the database but not analyzed. A significant number of subjects (n = 89) were diagnosed with vascular symptoms but failed to meet INS criteria for vascular dementia. Had these subjects been grouped in the vascular dementia category, effects may have been more pronounced. Other diagnoses included Cognitive Disorder Not Otherwise Specified (n = 50) and mixed vascular and Alzheimer's dementias (n= 6).

A second important weakness of the present study involved demographics. Though research has found the clock-drawing test to be less influenced by education than other cognitive screening tools (Silverstone et al., 1993), no effort was made to match subjects on educational level. In addition, no tests were run to determine if our sample was skewed toward any particular ethnic group. In the data collection procedure, minimal information was requested regarding possible comorbid disorders that might affect cognitive performance. Subjects with depression, obsessive-compulsive disorder and diabetes were identified but no analyses were performed to highlight any possible effects on clock drawing performance.

Finally, the research hypotheses were based primarily on the assumption that the utilized qualitative measures were a good representation of executive functioning. It might have been useful to test these criteria alongside proven executive function assessments before forming our hypotheses. Royall et al. (1998) has found that the CLOX scoring method is the strongest correlate of executive functioning. Future studies might look to incorporate this system into their design so as to assure adequate measurement of true executive function.

Despite the previously mentioned limitations, this study found support for the research hypotheses. Implications for future research are vast, as similar results may offer important evidence for the use of qualitative criteria in making accurate diagnoses. Furthermore, if, as in this study, only a few scoring criteria are found to account for significant variance, clinicians may be able to greatly abbreviate length of assessments, further appeasing managed care companies already enthralled with the use of screening tools.

Table 1

Mean Error Scores by Diagnosis

	Diagnostic Group					
	AD (*n*= 48)		VaD (*n*= 69)		Group diff. test	
Error types	*M*	*SD*	*M*	*SD*	*t*-test	*p* diff.
Incorrect time 1 = present 2 = absent	1.38	.49	1.59	.49	2.37	.02
Spacing errors 1 = present 2 = absent	1.35	.48	1.49	.50	1.49	.139
Errors of omission 1 = present 2 = absent	1.40	.49	1.71	.46	3.54	.001
Addition errors 1 = present 2 = absent	1.42	.50	1.58	.50	1.74	.084
Micro/macrographic errors 1= micro 2= macro 3= none	2.50	.85	2.26	.90	-1.19	.235
Circle quality 1= poor 2= good	1.56	.50	1.65	.48	.976	.331
Manos score (0= worst 10= best)	4.60	3.26	6.41	2.92	3.13	.002

REFERENCES

Borson, S., Brush, M., Gil, E., Scanlan, J., Vitaliano, P., Chen, J., Cashman, J., Sta Maria, M.M., Barnhart, R., & Roques, J. (1999). The Clock Drawing Test: Utility for dementia detection in multiethnic elders. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 54*, 534-540.

Erzigkeit, H. (1989). *SKT: Ein Kurtest zur Erfassung von Gedachtnis und Aufmerksamkeitssorungen* [Short Performance Test Manual]. Weinheim, 4. Auflage: Beltz. (Original work published 1989).

Federal Interagency Forum on Aging-Related Statistics (n.d). *Older Americans 2004: Key Indicators of Well-Being.* Retrieved March 17, 2006 from http://www.agingstats.gov/chartbook2004/default.htm

Goodglass, H., Kaplan, E. (1983). *The Assessment of Aphasia and Related Disorders*. Philadelphia: Lea and Febiger.

Heinik, J., Solomesh, I., Raikher, B., & Lin, R. (2002). Can clock drawing test help to differentiate between dementia of the Alzheimer's type and vascular dementia? A preliminary study. *Journal of International Geriatric Psychiatry, 17,* 699-703.

Heinik, J., Solomesh, I., Shein, V., & Becker, D. (2002). Clock Drawing Test in mild and moderate dementia of the Alzheimer's type: A comparative and correlation study. *International Journal of Geriatric Psychiatry, 17,* 480-485.

Ihl, R., Frolich, L., Dierks, T., Martin, E., & Maurer, K. (1992). Disease in patients with MMSE scores greater than 23. *Psychiatry Research, 44,* 93-106.

Inzitari, D., & Pantoni, L. (1998). Subcortical vascular dementia: Therapeutical aspects. *Archives of Gerontology and Geriatrics,* 263-268.

Manos, P. J. (1997). The utility of the Ten-Point Clock Test as a screen for cognitive impairment in general hospital patients. *General Hospital Psychiatry, 19,* 439-444.

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.*

Mayo Foundation for Medical Education and Research (2005). Visualizing Alzheimer’s. *Discovery’s Edge*. Retrieved March 5, 2006 from http://discoverysedge.mayo.edu/discoverys.cfm

Mendez, M. F., Ala, T., & Underwood, T. (1992). Development of scoring criteria for the clock drawing task in Alzheimer's Disease. *Journal of American Geriatric Society, 40,* 1095-1098.

Moretti, R., Torre, P., Antonello, R. M., Cazzato, G., & Bava, A. (2002). Ten-point Clock Test: A correlation analysis with other neuropsychological tests in dementia. *International Journal of Geriatric Psychiatry, 17,* 347-353.

National Institute on Aging (Nov. 12, 2001). *Cognitive Impairment High Among Older People, Study Suggests*. Retrieved March 17, 2006 from http://www.nia.nih.gov/Alzheimers/ResearchInformation/NewsReleases/Archives/PR20011112cognitiveimpairment.htm

Roman G.C., Tatemichi T.K., Erkinjuntti T., Cummings J.L., Masdeu J.C., Garcia J.H., Amaducci L., Orgogozo J.M., Brun A., Hofman A., et al. (1993). Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology, 43*, 250-260.

Royall, D.R. (1996). Comments on the executive control of clock-drawing. *Journal of the American Geriatric Society, 44*, 218-219

Royall, D.R., Cordes, J. A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology and Neurosurgical Psychiatry, 64*, 588-594.

Schramm, U., Berger, G., Muller, R., Kratzsch, T., Peters, J., & Frolich, L. (2002). Psychometric properties of the 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., Forstl, H., & Bickel, H. 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., Shedletsky, R., & Silver, I.L. (1986). The challenge of time: Clock-drawing and cognitive functioning in the elderly. *International Journal of Geriatric Psychiatry, 1,* 135-140.

Silverstone, F. A., Duke, W. M., & Wolf-Klein, G. P. (1993). Clock drawing helps when communication fails [Letter to the editor]. *Journal of American Geriatric Society, 41,* 1155.

Simard, M., & van Reekum, R. (1999). Memory assessment in studies of cognition enhancing drugs for Alzheimer's Disease. *Drugs and Aging, 14*, 197-230.

Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., et al. (1989). Clock drawing and Alzheimer's Disease: A novel measure of dementia severity. *Journal of the American Geriatrics Society, 37,* 725-729.

Tuokko, H., Hadjistavropoulos, T., Miller, J. A., & Beattie, B. L. (1992). The Clock Test: A sensitivity measure to differentiate normal elderly from those with Alzheimer's Disease. *Journal of the American Geriatrics Society, 40,* 579-584.

Tuokko, H., Horton, A., Hadjistavropoulos, T. et al. (1990). *The Clock Test: Manual for Administration and Scoring.* Vancouver, BC: University Hospital-UBC Site.

Wolf-Klein, G. P., Silverstone, F. A., Levy, A. P., Brod, M. S., Breuer, J. (Authors), & Breuer, J. (Statistical Consultant). (1989). Screening for Alzheimer's Disease by clock drawing. *Journal of the American Geriatrics Society, 37,* 730-734.