Archives of Clinical Neuropsychology 29 (2014) 1–6

Clock Drawing as a Screen for Impaired Driving in Aging and Dementia: Is It Worth the Time?

Kevin J. Manning[1,2], Jennifer D. Davis[1,*], George D. Papandonatos[3], Brian R. Ott[4]

[1]*Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, RI, USA*
[2]*Department of Psychiatry, New York Presbyterian Hospital, NY, USA*
[3]*Department of Biostatistics, Brown University, RI, USA*
[4]*Department of Neurology, Alpert Medical School of Brown University, RI, USA*

*Corresponding author at: Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, RI, USA.
Tel.: +1-401-444-4500; Fax: +1-401-444-6643.
E-mail address: jennifer_davis@brown.edu (J.D. Davis).

Accepted 5 November 2013

Abstract

Clock drawing is recommended by medical and transportation authorities as a screening test for unsafe drivers. The objective of the present study was to assess the usefulness of different clock drawing systems as screening measures of driving performance in 122 healthy and cognitively impaired older drivers. Clock drawing was measured using four different scoring systems. Driving outcomes included global ratings of safety and the error rate on a standardized on-road test. Findings revealed that clock drawing was significantly correlated with the driving score on the road test for each of the scoring systems. However, receiver operator curve analyses showed limited clinical utility for clock drawing as a screening instrument for impaired on-road driving performance with the area under the curve ranging from 0.53 to 0.61. Results from this study indicate that clock drawing has limited utility as a solitary screening measure of on-road driving, even when considering a variety of scoring approaches.

Keywords: Clock drawing; Driving; Dementia; Alzheimer's disease

Introduction

Cognitive impairment places older drivers at risk for accidents and unsafe driving (Ott, Heindel, et al., 2008). Specifically, deficits in processing speed, executive functioning, and spatial abilities are associated with poor driving performance in older adults with and without neurological disease (Dawson, Anderson, Uc, Dastrup, & Rizzo, 2009; Dawson, Uc, Anderson, Johnson, & Rizzo, 2010). Considerable efforts have been made in developing screening instruments that measure these cognitive functions to identify potentially unsafe older drivers who are appropriate for more comprehensive assessment (Molnar, Patel, Marshall, Man-Son-Hing, & Wilson, 2006). The American Medical Association (AMA) and the National Highway Traffic Safety Administration recommend clock drawing as a screening test of driving abilities in older adults as a component of the Assessment of Driving-Related Skills (ADReS). Clock drawing is a logical candidate for a screening test of driving performance; it is brief, easy to administer, sensitive to cognitive impairment and is associated with other measures of executive functioning and spatial abilities (Cosentino, Jefferson, Chute, Kaplan, & Libon, 2004).

Screening tests should correctly identify poor drivers as unsafe to ensure individual and public safety (sensitivity) and correctly identify strong drivers as safe (specificity) to prevent premature driving cessation and maintain independence (Oxley & Whelan, 2008). Recommendations are that ideal screening measures should effectively classify 80%–90% of older drivers to minimize unnecessary referrals for on-road testing (Bedard, Weaver, Darzins, & Porter, 2008). Evidence for clock drawing as an effective screening test in older drivers is inconclusive. Using the command to draw a clock for "ten after eleven," Freund, Gravenstein, Ferris, Burke, and Shaheen (2005) found a strong association between a novel scoring method and total errors on a virtual

doi:10.1093/arclin/act088 Advance Access publication on 1 December 2013

reality simulator ($r = .68$) in a sample of 109 mildly cognitively impaired older drivers. However, in discriminating clinician-rated safe and unsafe driving simulator performance, Freund and colleagues (2005) found high specificity (97.7%) but low sensitivity (64.2%) for a clock drawing score of 4/7. Other cutoff scores (e.g., 6/7) were more likely to identify unsafe drivers (sensitivity 96.2%), but at the cost of increasing false positives (specificity 58.1%). No cutoff score resulted in both high sensitivity and high specificity (i.e., 80%) in classifying simulated driving. Elsewhere, using a different clock drawing test from Freund and colleagues (2005), Oswanski and colleagues (2007) reported both low sensitivity (70%) and low specificity (65%) for a clock drawing score of 3/4 in discriminating safe and unsafe on-road driving (as rated by a driving specialist) in a general sample of 232 older adults. However, Oswaski and colleagues (2007) did not provide information on overall cognitive status or the presence of neurological disorders that could have influenced results. Furthermore, sample bias may have influenced the findings, as all participants were referred for a clinical driving evaluation due to concern over driving capabilities.

Despite limited empirical evidence, AMA (2010) recommends older drivers who earn less than a perfect score on the Freund 7-point clock drawing test be referred to a driving specialist and be considered for on-road testing. To date, no studies have reported cutoff scores for this scoring system in older drivers who have completed performance-based on-road testing. Moreover, single cutoff scores from screening instruments are unlikely to confidently predict driving performance as the distributions of good and bad drivers likely overlap (Molnar et al., 2006). In an effort to increase the clinical utility of screening instruments, Molnar and colleagues (2006) recommend using both upper and lower cutoff scores to create intervals for driving risk (e.g., clearly unsafe, indeterminate, clearly safe); however, upper and lower cutoff scores have not been reported in clock drawing. Finally, the Freund 7-point clock drawing method is only one of several clock drawing systems available. No studies have examined potential differences in clock drawing methods and road-test performance in older drivers.

The present study aimed to address the issues by comparing the value of multiple clock scoring systems in predicting unsafe on-road driving in older drivers without cognitive impairment and those with very mild or mild dementia.

Methods

Participants

Study participants were 122 active drivers ranging in age from 60 to 89, with a valid driver's license. Healthy older adults ($n = 47$) had no history of dementia and a Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) score greater than 26. Cognitively impaired participants ($n = 75$) were recruited from a hospital-based memory disorders clinic. Patients met diagnostic criteria for possible or probable Alzheimer's disease (AD) based on the NINCDS-ADRDA criteria (McKhann et al., 1984). Cognitive assessment was included in the diagnosis when available. Dementia severity was measured with the Clinical Dementia Rating scale which was completed according to standardized procedures by a memory disorder specialist (Morris, 1993); only patients with CDR scores of 0.5 or 1.0 (questionable or mild dementia) were included. Exclusion criteria for both groups included reversible causes of dementia, physical or ophthalmologic disorders that might impair driving abilities, severe psychiatric illness, or alcohol/substance abuse within the previous year. Because many authorities feel that drivers with dementia who have an at-fault crash should not drive, we excluded drivers with history of at-fault crash in the past year. Patients were on a stable dose of a cholinesterase inhibitor for 6 weeks, if prescribed. Anxiolytic and antipsychotic medications were permitted. The Rhode Island Hospital Institutional Review Board approved all procedures.

Clock Drawing

The clock drawing test was administered during an office visit prior to the road test. Using the same scoring system as Freund and colleagues (2005), participants were given a blank piece of paper and instructed to "Draw a clock. Put in all the numbers, and set the hands for ten after eleven." The instructions were repeated as necessary but no further help was given in completing the task. Four scoring systems were evaluated. These included the 7-point scoring system (Freund et al., 2005) as well as an earlier version of this scoring system (scored out of 8 points) previously recommended as a component of the ADReS (Wang, Kosinski, Schwartzberg, & Shanklin, 2003). Despite the present use of the 7-point system, no studies have compared the performance of the 7- and 8-point scoring systems in predicting on-road driving. Other clock drawing methods evaluated included a 10-point scoring system (Rouleau, Salmon, Butters, Kennedy, & McGuire, 1992), and a modified version of a scoring system based on the ability to conceptualize the correct clock drawing time (Cahn-Weiner et al., 2003). This last method was selected as a way of identifying older drivers who have lost the semantic knowledge necessary for identifying time; these drivers may be at considerable risk for unsafe driving as a result of this level of cognitive impairment. This scoring method assigned 1 point for the correct conceptualization of time and 0 point if time was not correctly conceptualized. Scoring guidelines for the remaining 7-, 8-, and 10-point methods are published in detail elsewhere (Freund et al., 2005; Rouleau et al., 1992; Wang et al., 2003).

Standardized Road Test

The Rhode Island Road Test (RIRT) is a standardized on-road driving assessment with established reliability and validity with older adults. See Davis and Colleagues (2012) for a detailed description of the RIRT. Briefly, the RIRT is administered by a professional driving instructor on a fixed course that is comprised of 28 routine driving maneuvers (e.g., response to signage, right and left turns, lane keeping, parking). The instructor was blind to diagnosis and rated the driving maneuvers/behaviors on a 3-point scale (0 = unimpaired, 1 = mildly impaired, 2 = moderately to severely impaired) and made a global rating of overall driving ability (Pass, Marginal, or Fail). Pass implied that the participant's driving performance would unlikely result in crashes or violations. Marginal indicated an increased crash risk requiring remediation with driving lessons or restriction (e.g., obstructing traffic flow). A Fail rating indicated that the driver exhibited behavior that had a high probability of leading to accidents, but could not be easily remediated (e.g., bypassing a traffic signal). Given the prime objective of this study was to determine the value of clock drawing as a screening test to identify individuals who would require follow-up assessments, Marginal and Fail ratings were grouped together and considered "Unsafe" driving, whereas Pass was considered "Safe" driving.

Results

Sample characteristics, clock drawing performances, and driving outcomes are presented in Table 1. Subjects demonstrated mild difficulties with on-road driving. Considering all participants, the road test error rate was 0.07 ($SD = 0.59$). The majority of participants were rated as safe drivers (60%) whereas 40% were rated as unsafe (23% Marginal, 16% Fail). As expected, AD drivers had higher error scores and had poorer global ratings of driving than the HC drivers, as well as poorer performance on clock drawing as measured by all four scoring methods.

Clock Drawing and Driving

We first examined the association among the different clock drawing systems using Spearman's rank or point biserial correlation coefficients and found all four were strongly correlated (7-point/8-point $\rho = .77$, $p < .001$; 7-point/10-point $\rho = .76$, $p < .001$; 7-point/conceptualization $r_{pb} = .78$, $p < .001$; 8-point/10-point $\rho = .85$, $p < .001$; 8-point/conceptualization $r_{pb} = .68$, $p < .001$; 10-point/conceptualization $r_{pb} = .76$, $p < .001$). Next, we computed correlation analyses between the clock drawing scoring systems and the road test error rate. Clock drawing using each of the four scoring systems was significantly correlated with road test performance. The 7-point ($\rho = -.32$, $p < .001$) scoring system was moderately correlated with road-test performance, whereas the remaining scoring systems were modestly related to road test performance (8-point $\rho = -.27$, $p < .05$; 10-point $\rho = -.28$, $p < .05$; conceptualization $r_{\text{pb}} = -.25$, $p < .05$).

We next used receiver operating characteristic (ROC) curves to evaluate the clinical utility of each of the clock drawing systems in predicting unsafe driving. Overall discriminatory ability, as indicated by the area under the curve (AUC), was low, ranging from 0.53 to 0.61 for different scoring methods (Table 2). Based on the recommendation of Molnar and colleagues (2006), we examined

Table 1. Participant characteristics, clock drawing, and driving outcomes

	Healthy controls ($n = 47$)	Alzheimer's disease ($n = 75$)
Age	71.87 (7.77)	76.65 (6.24)*
Gender (%, M/F)	61.7/38.3	53.3/46.7
Education	16.30 (3.78)	13.53 (3.28)*
MMSE	29.48 (0.65)	25.05 (2.83)*
CDR sum of boxes	0.00 (0.00)	2.76 (1.60)*
Visual acuity (≥20/20)	45%	15%*
7-point system	5.61 (1.07)	4.16 (1.67)**
8-point system	7.43 (1.06)	6.05 (2.08)**
10-point system	9.17 (.94)	7.52 (2.20)**
Conceptualization error	10.60%	43.80%**
Driving experience (years)	53.11 (6.80)	56.28 (9.23)*
Miles driven/week	195.53 (114.87)	87.64 (88.20)**
Road test error rate	0.04 (0.03)	0.09 (0.07)**
Unsafe driver ratings	14.90%	54.70%**

Notes: Values are means (*SD*) unless otherwise noted; MMSE = Mini-Mental State Examination; CDR = Clinical Dementia Rating Scale.
*$p < .05$.
**$p < .001$.

Table 2. Sensitivity and specificity of clock drawing in predicting unsafe driving ($n = 122$)

Scoring system	Sensitivity	Specificity
7-point (AUC = 0.61, 95% CI = 0.51–0.71)		
0	1.00	0.00
1.5	0.99	0.08
2.5	0.91	0.15
3.5	0.78	0.30
4.5	0.69	0.41
5.5	0.43	0.72
6.5	0.18	0.98
8-point (AUC = 0.60, 95% CI = 0.50–0.71)		
0	1.00	0.00
1.5	1.00	0.06
2.5	1.00	0.15
3.5	0.95	0.17
4.5	0.89	0.19
5.5	0.86	0.29
6.5	0.72	0.44
7.5	0.50	0.62
10-point (AUC = 0.58, 95% CI = 0.48–0.69)		
0	1.00	0.00
1.5	1.00	0.02
2.5	1.00	0.03
3.5	1.00	0.04
4.5	0.97	0.15
5.5	0.90	0.17
6.5	0.82	0.25
7.5	0.77	0.31
8.5	0.62	0.54
9.5	0.34	0.75
Conceptualization (AUC = 0.53, 95% CI = 0.43–0.64)		
Yes/No	0.71	0.35

Notes: AUC = area under the curve; CI = confidence interval.

the ROC curves to determine if there were confident upper and lower cutoff scores as defined by sensitivity or specificity of at least 0.80. The 7-point scoring method appeared to be a mild improvement compared with other scoring systems as it was the only one to provide both a confident upper (>6) and lower (<3) cut point (Table 2). However, the application of the 7-point upper and lower cutoff scores to the present sample was inaccurate in classifying the majority of drivers. Only 24% of the sample scored within the upper (>6, $n = 14$) and lower (<3, $n = 15$) range, leaving 76% ($n = 93$) of participants with an indeterminate rating of road-test performance. Furthermore, although the upper cut score was highly specific (positive predictive value = 0.43, negative predictive value = 0.93), reliably classifying safe drivers 93% of the time (13/14 drivers who scored >6 passed the road test), the use of a lower cut score was insensitive, only classifying unsafe drivers accurately 53% of the time (8/15 drivers who scored <3 failed the road test; positive predictive value = 0.37, negative predictive value = 0.46).

Discussion

The principle finding of the current study is clock drawing showed limited clinical utility as a sole screening instrument for impaired on-road driving performance in healthy older and cognitively impaired adults. Overall, results varied little based on different clock drawing scoring methods, although the 7-point scoring method appeared to be somewhat better than the alternative methods, since it was the only one to provide upper and lower cut scores with fair sensitivity and specificity. However, the application of the 7-point upper and lower cut scores to the present sample was inaccurate in classifying driving performance for the majority (76%) of participants. The major implication of these findings is that clock drawing alone is not effective as a screening test for unsafe driving in older healthy adults and adults with mild dementia.

In a prior study, a score of 4/7 on the 7-point clock drawing system demonstrated high specificity (97.7%) in classifying poor older driver performance on a driving simulator (Freund et al., 2005). In contrast, a score of 4/7 on the 7-point system demonstrated both poor sensitivity and specificity in classifying unsafe on-road driving in the present sample. Consistent with the AMA's recommendations, the best cutoff score was a 6/7, resulting in a specificity of 98%. However, only 11% (14/122) of the present sample earned a perfect 7/7 score. Therefore, while a perfect score on the 7-point system reliably indicated a safe driver, too few drivers scored within this range to make this cutoff score clinically useful. A lower cutoff score (3/7) resulted in poor sensitivity, and using this cutoff in conjunction with the higher cut score 6/7 was ineffective in classifying driving risk.

We have previously reported that the application of the complete ADReS test battery in the present sample revealed limited overall utility as a screening test (Ott et al., 2013), whereas similar office-based measures in a more impaired sample of older drivers produced higher classification accuracy (Carr, Barco, Wallendorf, Snellgrove, & Ott, 2011). Additional work is needed to determine the additive value of other cognitive measures besides clock drawing that could enhance predictive ability in the office setting. However, compared with individual screening tests, extensive neuropsychological assessments may prove more useful in driving prediction of older healthy adults and those with mild cognitive impairment or early dementia. Composite measures of executive functioning, visuospatial, and sensory functioning tests are moderately to strongly associated with on-road driving performance in healthy older adults and adults with probable mild AD (Dawson et al. 2009, 2010; Ott, Festa, et al. 2008), but neuropsychological tests explain only modest amounts of variance in on-road driving performance (Dawson et al., 2010). The cognitive underpinnings of challenging driving scenarios (e.g., self-guided navigation) are not well understood, and alternative approaches to driving assessment may strengthen our understanding of the relationship between cognition and driving skills.

The composition of our sample limits the generalizability of these findings. The classification accuracy in other samples may differ dependent on their particular clinical characteristics. Although a strength of our sample composition is that it provided for a range of cognitive and driving abilities with which to study the predictive power of clock drawing, the mix of subjects with cognitive impairment in our sample is likely higher than that seen in most general clinical practices. Therefore, the percentage of indeterminate drivers would likely be considerably lower in a physician's general office practice where most drivers do not exhibit significant cognitive deficits that would be expected to impair driving performance. Future investigators may wish to compare the cutoff scores from the present study to those obtained from different settings to assess for potential differences in positive and negative predictive values.

Despite these limitations, the present findings add to the existing literature and inform recommendations for screening tests by demonstrating that clock drawing is not useful as a solitary screening test for older healthy adults and mildly cognitively impaired drivers.

Funding

This work was supported by the National Institute on Aging at the National Institutes of Health (R01 AG016335 to BRO).

Conflict of Interest

None declared.

Acknowledgements

The authors wish to thank Mr. Scott Hewitt with his help in data collection.

References

American Medical Association. Physician Guide to Assessing and Counseling Older Drivers. Retrieved November 25, 2012, from http://www.ama-assn.org.2010.

Bedard, M., Weaver, B., Darzins, P., & Porter, M. M. (2008). Predicting driving performance in older adults: We are not there yet! *Traffic Injury Prevention*, *9*, 336–341.

Cahn-Weiner, D. A., Williams, K., Grace, J., Tremont, G., Westervelt, H., & Stern, R. A. (2003). Discrimination of dementia with Lewy bodies from Alzheimer disease and Parkinson disease using the clock drawing test. *Cognitive and Behavioral Neurology*, *16*, 85–92.

Carr, D. B., Barco, P. P., Wallendorf, M. J., Snellgrove, C. A., & Ott, B. R. (2011). Predicting road test performance in drivers with dementia. *Journal of the American Geriatric Society*, *59*, 2112–2117.

Cosentino, S., Jefferson, A., Chute, D. L., Kaplan, E., & Libon, D. J. (2004). Clock drawing errors in dementia: Neuropsychological and neuroanatomical considerations. *Cognitive and Behavioral Neurology*, *17*, 74–84.

Davis, J. D., Papandonatos, G. D., Miller, L. A., Hewitt, S. D., Festa, E. K., Heindel, W. C., et al. (2012). Road test and naturalistic driving performance in healthy and cognitively impaired older adults: Does environment matter? *Journal of the American Geriatric Society*, *60*, 2056–2062.

Dawson, J. D., Anderson, S. W., Uc, E. Y., Dastrup, E., & Rizzo, M. (2009). Predictors of driving safety in early Alzheimer disease. *Neurology*, *72*, 521–527.

Dawson, J. D., Uc, E. Y., Anderson, S. W., Johnson, A. M., & Rizzo, M. (2010). Neuropsychological predictors of driving errors in older adults. *Journal of the American Geriatric Society*, *58*, 1090–1096.

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 Psychiatry Research*, *12*, 189–198.

Freund, B., Gravenstein, S., Ferris, R., Burke, B. L., & Shaheen, E. (2005). Drawing clocks and driving cars. *Journal of General Internal Medicine*, *20*, 240–244.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. *Neurology*, *34*, 939–944.

Molnar, F. J., Patel, A., Marshall, S. C., Man-Son-Hing, M., & Wilson, K. G. (2006). Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: A systematic review. *Journal of the American Geriatric Society*, *54*, 1809–1824.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*, 2412–2414.

Oswanski, M. F., Sharma, O. P., Raj, S. S., Vassar, L. A., Woods, K. L., Sargent, W. M., et al. (2007). Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine and Rehabilitation*, *86*, 190–199.

Ott, B. R., Davis, J. D., Papandonatos, G. D., Hewitt, S. D., Festa, E. K., Heindel, W. C., et al. (2013). Assessment of driving related skills (ADRes) prediction of unsafe driving in older adults in the office setting. *Journal of the American Geriatric Society*, *61*, 1164–1169.

Ott, B. R., Festa, E. K., Amick, M. M., Grace, J., Davis, J. D., & Heindel, W. C. (2008). Computerized maze navigation and on-road performance by drivers with dementia. *Journal of Geriatric Psychiatry and Neurology*, *21 (1)*, 18–25.

Ott, B. R., Heindel, W. C., Papandonatos, G. D., Festa, E. K., Davis, J. D., Daiello, L. A., et al. (2008). A longitudinal study of drivers with Alzheimer’s disease. *Neurology*, *70*, 1171–1178.

Oxley, J., & Whelan, M. (2008). It cannot be all about safety: The benefit of prolonged mobility. *Traffic Injury Prevention*, *9*, 367–378.

Rouleau, I., Salmon, D. P., Butters, N., Kennedy, C., & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer’s and Huntington’s disease. *Brain and Cognition*, *18*, 70–87.

Wang, C. C., Kosinski, C. J., Schwartzberg, J. G., & Shanklin, A. V. (2003). *Physician’s guide for assessing and counseling older drivers*. Washington, DC: National Highway Traffic Safety Administration.