



Published in final edited form as:

*Exp Aging Res.* 2013 January ; 39(1): 80–108. doi:10.1080/0361073X.2013.741996.

## The Framingham Heart Study Clock Drawing Performance: Normative Data from the Offspring Cohort

Justin A. Nyborn, MPH<sup>1,2</sup>, Jayandra J. Himali, MS<sup>2,3</sup>, Alexa S. Beiser, PhD<sup>1,2,3</sup>, Sherral A. Devine, PhD<sup>1,2</sup>, Yangchun Du, PhD<sup>2,3</sup>, Edith Kaplan, PhD ABCN<sup>1,4</sup>, Maureen K. O'Connor, PsyD ABCN<sup>1,5</sup>, William E. Rinn, PhD<sup>6</sup>, Helen S. Denison, PhD<sup>1,4</sup>, Sudha Seshadri, MD<sup>1,2</sup>, Philip A. Wolf, MD<sup>1,2</sup>, and Rhoda Au, PhD<sup>1,2</sup>

<sup>1</sup>Department of Neurology, Boston University School of Medicine, Boston, MA

<sup>2</sup>Framingham Heart Study, Boston MA

<sup>3</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA

<sup>4</sup>Department of Psychology, Suffolk University, Boston, MA

<sup>5</sup>Center for Translational Cognitive Neuroscience, Bedford VA Hospital, Bedford, MA

<sup>6</sup>Spaulding Rehabilitation Hospital, Boston, MA

### Abstract

**Background/Study Context**—While the Clock Drawing Test (CDT) is a popular tool used to assess cognitive function, limited normative data on CDT performance exists. The objective of the current study was to provide normative data on an expanded version of previous CDT scoring protocols from a large community-based sample of middle to older adults (aged 43 to 91) from the Framingham Heart Study.

**Methods**—The CDT was administered to 1476 Framingham Heart Study Offspring Cohort participants using a scoring protocol that assigned error scores to drawn features. Total error scores were computed, as well as for subscales pertaining to outline, numeral placement, time-setting, center, and “other.”

**Results**—Higher levels of education were significantly associated with fewer errors for time-setting (Command:  $p < .001$ ; Copy:  $p = .003$ ), numerals (Command:  $p < .001$ ) and “other” (Command:  $p < .001$ ) subscales. Older age was significantly associated with more errors for time-setting (Command:  $p < .001$ ; Copy:  $p = .003$ ), numeral (Command:  $p < .001$ ) and “other” (Command:  $p < .001$ ) subscales. Significant differences were also found between education groups on the Command condition for all but the oldest age group (75+).

**Conclusion**—Results provide normative data on CDT performance within a community-based cohort. Errors appear to be more prevalent in older compared with younger individuals, and may be less prevalent in individuals who completed at least some college compared with those who did not. Future studies are needed to determine whether this expanded scoring system allows detection of preclinical symptoms of future risk for dementia.

### Keywords

Clock Drawing Test; Normal aging; Scoring methods; Neuropsychological tests; Dementia; Cognitive screening

## INTRODUCTION

The Clock Drawing Test (CDT) is a multifaceted and multidimensional cognitive assessment tool widely used in research and clinical practice. It is particularly sensitive to detecting dysfunction among the elderly and those with neurological and psychiatric disorders (Brodaty & Moore, 1997; Freedman et al., 1994; Hermann et al., 1999; Hubbard et al., 2008). Previous studies have found the CDT to correlate well with a diagnosis of dementia (Heinik, Solomesh, Raikher, & Lin, 2002; Lessig, Scanlan, Nazemi, & Borson, 2008; Shulman, Pushkar, Cohen, & Zuccherro, 1993), including Alzheimer's disease (AD) (Brodaty & Moore, 1997; Sunderland et al., 1989; Wolf-Klein, Silverstone, Levy, & Brod, 1989). While many variations on CDT scoring systems have been reported, they largely share the same general protocol (Freedman et al., 1994; Henderson & Hotopf, 2007; Shulman, 2000). Shulman (2000) did a meta-analysis of CDT scoring systems and found high sensitivity and specificity, high inter-rater and test-retest reliability, and good predictive validity to detect cognitive change.

The utility of CDT as a test of cognitive status lies in the fact that it taps into multiple processes that involve different regions of the brain (Freedman et al., 1994). When asked to draw a clock and set the hands to a pre-specified time, individuals must possess adequate visuoconstructive and visuospatial skills, auditory comprehension, verbal and visual memory, motor programming, numerical knowledge, spatial attention, concentration, frustration tolerance, and executive functioning including organization, planning, abstract reasoning, and parallel processing (Freedman et al., 1994; Shulman, 2000; Libon, Malamut, Swenson, Sands, & Cloud, 1996; Spreen & Strauss, 1998; Ueda et al., 2002). In addition to measuring an array of cognitive skills, the CDT has been shown to offer other advantages to practitioners and researchers including: 1) brief and simple administration protocol (Bourke, Castleden, Stephen, & Dennis, 1995; Hubbard et al., 2008; Shulman, 2000; Shulman, Shedletsky, & Silver, 1986); 2) sensitivity to subtle impairments (Freedman et al., 1994; Libon et al., 1996); 3) low cost (Nishiwaki et al., 2004; Shulman et al., 1986); and 4) correlation with other cognitive tests such as the Blessed Dementia Rating Scale (Brodaty & Moore, 1997), Wechsler Memory Scale (Strauss, Sherman, & Spreen, 2006), and the Mini-Mental State Examination (MMSE) (Brodaty & Moore, 1997; Bourke et al., 1995; Shulman et al., 1986; Shulman, 2000). Ease of administration and acceptability to patients have fueled a recent resurgence of studies on the CDT among clinicians and researchers, each using modified clock-scoring systems (Hubbard et al., 2008; Shulman, 2000).

Studies report that CDT performance does show consistent decline with age (Crowe et al., 2010; Freedman et al., 1994; Spreen & Strauss, 1998; Hubbard et al., 2008; von Gunten, 2008), but no study to our knowledge has reported significant gender differences (Hubbard et al., 2008; Kim & Chey, 2010). While the CDT has been touted as an education unbiased (Lam et al., 1998; Shulman et al., 1986; Yamamoto, et al., 2004), and culture-fair (Marcopulos & McLain, 2003) test, the absence of educational or cultural bias is not clear.

Studies that used low education cut-offs (e.g., versus < than 7 years) found that those with higher education performed better on the CDT (Kim & Chey, 2010; Marcopulos & McLain, 2003). Another study that compared three CDT scoring systems used higher education cut-offs (e.g., versus < college graduate) and did not find education to impact CDT performance for two of the scoring systems (Hubbard et al., 2008). Further, the education effect disappeared for the third scoring system when controlling for educational attainment, as measured by a reading test. von Gunten et al. (2008) compared CDT performance across three educational levels and found poorer performance only among those in the lowest educational group that were over age 80. Ratliff et al. (2003) and Hubbard et al. (2008)

proposed there is a “ceiling effect” of education on CDT scores because healthy elderly obtain relatively high CDT scores no matter their education level, so education effects are difficult to observe.

Reported ethnic differences on the CDT (Manly et al, 2004; Crowe et al., 2008) find that African American participants do more poorly on the clock drawing test than people of Caucasian ethnicity, even after adjusting for other confounding factors such as age, education and level of urbanization (Crowe et al, 2008). But the use of years of education rather than more objective measures of educational attainment may be an underlying mediating factor (Crowe et al., 2010). Documented systematic social, regional and cultural biases have created educational inequities between the African American and Caucasian U.S. populations, whose impact is seen most significantly among the elderly. Thus, studies that use educational attainment measures, such as the Wide Range Achievement Test – Reading Subtest 3, Arrangement, in place of reported years of education, have found no difference based on ethnicity (Hubbard et al., 2008; Crowe et al., 2010). These studies suggest that in community-based studies with a diverse sample, reading ability may be a better predictor of cognitive performance than years of education.

Growing concern about early detection of AD, particularly at the pre-clinical stages (e.g., Mild Cognitive Impairment [MCI]), has propelled investigations as to whether the CDT can distinguish between normal and mildly impaired performance. While some studies find no differentiation (Lee et al., 2008; Powlishta et al., 2002; Seigerschmidt, Mosch, Siemen, Forstl, & Bickel, 2002), others do find CDT performance differences (De Jager, Hogervorst, Combrinck, & Budge, 2003; Yamamoto et al., 2004). The discrepancy in results may lie in the use of scoring systems with differential levels of sensitivity. Recent studies (Crowe et al. 2008; Crowe et al. 2010; Hubbard et al., 2008; von Gunten et al., 2008) have used scoring protocols that result in a wide range of scores (10–36 points). While a presumption might be made that more possible points reflects potential higher sensitivity in detecting cognitive change, in fact, it is the type of feature being scored that is a greater determinant of sensitivity rather than the number of points scored. With AD research focused on pre-clinical detection, a CDT scoring system that enhances sensitivity to subtle changes is warranted (Hubbard et al., 2008).

Despite widespread use of the CDT in both clinical and research settings, there is limited normative data on CDT performance (Hubbard et al., 2008). Freedman et al. (1994) developed one of the initial CDT scoring protocols to establish normative data on clock drawing performance. It incorporated qualitative features of the clock and consisted of 15 descriptors or “critical items” characteristic of a “good” clock, with one point assigned for presence of each critical item. The 15-point range was divided into 4 subgroups: clock outline (2 points), numerals (6 points), time setting (6 points), and center (1 point). They tested their CDT scoring system on 348 subjects ranging in age from 20 to 90 years in community centers in the Toronto area. To discriminate demented from non-demented participants, they also examined CDT performance in a subgroup of 18 non-demented participants, 20 with Parkinson’s disease but no dementia, 14 with Parkinson’s disease and dementia, and 13 with AD. They found that demented individuals scored below 12 out of a possible 15 points. Their cut-off score showed high sensitivity (.78), specificity (.82), and construct validity to performance on other neuropsychological tests including the Wechsler Adult Intelligence Scale-Revised (WAIS-R), the Wisconsin Card Sorting Test, and the Rey-Osterrieth Complex Figure (Freedman et al., 1994).

Hubbard et al. (2008) conducted a recent CDT normative study examining the reliability, validity, clinical utility, and effects of demographics of three previously published CDT scoring systems. They applied the scoring systems to clocks drawn by 207 cognitively intact

participants aged 55 to 98. They discovered that the range of CDT performance that would be considered normal is greater than previously reported in studies with smaller samples sizes. They suggested some variation in drawing ability was associated with age and education and should not be assumed to be abnormal. The authors proposed that what constituted “normal performance” based on these three systems was too narrow. They recommended further comparison of cognitively healthy participants to determine whether specific CDT errors are representative of the natural consequences of normal aging rather than necessarily indicators of the early stages of a progressive dementia. The objective of the current study is to provide normative data on an expanded version of previous CDT scoring protocols from a large community-based sample of middle to older adults (aged 43 to 91) from the Framingham Heart Study. Two primary reasons were the basis for expanding the scoring system. First, observations from Hubbard et al. (2008) that the range of performance that defines normal may be greater than has been documented suggests the need for a scoring system that can better detect a broader range of errors. Second, beginning with Peterson’s seminal paper on MCI (1999), the research on preclinical dementia particularly for Alzheimer’s disease, has become increasingly focused on earlier detection. Pre-MCI is emerging as a pre-symptomatic stage that neuropsychological tests, as currently scored, will not be able to detect (Sperling et al., 2011). Led by Dr. Edith Kaplan, who has developed earlier CDT scoring protocols, the Framingham Heart Study CDT scoring system sought to better capture subtle differences in clock drawing performance in people who are still years, possibly decades away from risk of Alzheimer’s disease.

## METHODS

### Participants

The FHS Offspring Cohort (n= 5124) was recruited in 1971 with the intent of establishing a prospective epidemiological study of young adults (mean age 37) to identify risk factors for cardiovascular and cerebrovascular diseases (Garrison, Kannel, Stokes, & Castelli, 1987). Biological Offspring of the Original FHS Cohort and their spouses were eligible to enroll. These members were invited to health examinations approximately every four years that included a detailed medical history, physical examinations, and laboratory tests (Kannel, Feinleib, McNamara, Garrison, & Castelli, 1979).

All Offspring participants who completed examination cycle 7 (n=3539) and also participated in a first round of neuropsychological testing (1999–2005) were recruited for a follow-up evaluation on brain imaging and cognition. A total of 1516 participants completed the Command and Copy CDT as part of a more comprehensive neuropsychological (NP) battery between 2005 and 2007. Participants with prevalent clinical stroke, dementia, or other neurological diseases (e.g., multiple sclerosis, severe head trauma, etc.) were excluded from the study (n=40). Thus, 1476 participants (680 men, 796 women) comprised the total sample for the normative study. Compared to the 2063 Offspring participants who completed examination cycle 7, but were not administered the Command and Copy CDT, the normative sample was younger (60.44 versus 63.05 years;  $p<.001$ ) and scored higher on the MMSE (29.03 versus 28.28;  $p<.001$ ). The normative sample was also healthier compared to the remaining participants from examination cycle 7 with less stage I hypertension (38% versus 52%;  $p<.001$ ), prevalent cardiovascular disease (8% versus 14%;  $p<.001$ ), diabetes (10% versus 16%;  $p<.001$ ), and smokers (10% versus 16%;  $p<.001$ ). Diagnosis of clinical stroke and dementia was done by consensus review panels and met standard diagnostic criteria (e.g., DSM IV and NINDS-ADRDA criteria for dementia/Alzheimer’s disease, respectively). Diagnostic procedures and criteria have been described in detail elsewhere (Seshadri et al., 2006). The proportion of excluded participants was relatively small because the majority of our sample was younger (only 6% over age 75 and

68% under age 65). The Institutional Review Board at Boston University Medical Center approved the study protocol and informed consent was obtained from all participants.

### **Clock Drawing Administration Procedure**

The CDT Command condition was administered using the standardized instructions, “Draw a clock, put in all the numbers, and set the hands to ten after eleven.” For the Copy condition, participants were shown an image of a clock and instructed to copy it. The Copy condition was administered approximately fifteen minutes after the Command Condition.

### **Clock Drawing Scoring Protocol**

A task group of five neuropsychologists (Devine et al., 2007) led by Edith Kaplan, one of the investigators who developed the Freedman et al. (1994) protocol, established the modified Framingham Heart Study Clock Drawing Test Scoring Protocol (FHS-CDT-SP) to enhance the sensitivity of the Freedman model. They assigned error scores to 38 qualitative features. Similar to Freedman’s model but calculating error scores instead of correct scores for critical features, the FHS-CDT-SP provides quantitative scores that include an overall summary error score, as well as subscale error scores related to outline, numeral placement, center, time-setting, and “other” (i.e. extraneous marks and self-corrected errors) characteristics of drawn clocks (Appendix A). Table 1 illustrates the FHS-CDT-SP Command and Copy condition breakdown of error points by subscale. Numerals (0–9 points) and time-setting (0–7 points) subscales comprise a majority of the 20.5 total possible error points and a majority of the 38 qualitative features. The FHS-CDT-SP algorithm was designed so that committing errors on some features (e.g., numerals are represented only by substitutes) excludes the possibility of making others (e.g., numerals are written out only). Additionally, participants can make multiple errors without receiving the maximum error score for each subscale (e.g., had only one 12 present [1 point], duplicated a numeral [1 point], and represented one numeral with a substitute [1 point] for a total of 3 out of 9 possible numeral placement error points).

Trained examiners scored the majority of Command and Copy condition clock variables based on observation, but some features required precise measurements. These measurements were calculated with two templates. Template 1 (Figure 1), a transparency sheet with concentric circles 10 millimeters apart, was placed directly over and pinned to the center of Template 2 (Figure 2), a transparency sheet with intersecting lines spaced 10 millimeters apart. Scorers used Template 1 to locate the center of the drawn circle and rotated Template 2 accordingly to calculate the longest diameter and its perpendicular bisecting diameter. Template 1 was also used to find the center of the numeral array, and that center was used with Template 2 to find the medial deviation, numeral displacement, and center of the hands (Appendix A).

Two quality control measures were taken to ensure consistency of scoring. First, two separate trained examiners independently score the same tests to ensure reliability of the scoring system. An inter-rater reliability score of 0.945 for the Command condition and 0.868 for the Copy condition was determined from a subset of 100 clocks. Second, a neuropsychologist (S.D.) performed a weekly review of a random set of tests for scoring accuracy. The average % of scoring errors across the entire 50 minute neuropsychological test battery, including the clock was less than 1%.

### **Statistical Analysis**

Measures of demographic characteristics and clock drawing error scores were summarized with means and standard deviations for the continuous measures and number and percentages for the categorical variables. Since the distributions of the command summary

error score (CmSES) and the copy summary error score (CoSES) were skewed, we performed our analyses using the natural logarithms of CmSES [LCmSES] and CoSES [LCoSES]. A general linear model was used to compare these error scores by age, gender, education and age $\times$ education; post-hoc pair-wise comparisons were made among the sub-groups using Tukey's procedure. Chi-square tests were used to assess relations between categorical measures of the components of CmSES and CoSES, and gender, age-groups, and education levels. All data were analyzed using SAS version 9.1.

## RESULTS

### Demographics

Table 2 provides demographic information on the total study sample. Mini-mental State Examination scores from the 7<sup>th</sup> cycle examination were high, with a mean of 29 out of 30 points, suggesting that that study sample was of normal cognition. The racial composition of the sample was Caucasian and fairly well-educated, with 65% having at least some college.

### Command Condition Error Score Totals by Subscale

The number and percent of participants making errors and no errors on each feature measured by the FHS-CDT-SP on the Command condition are displayed in Table 1. Commission of errors with low assigned point values was relatively common. For example, well over half (63.87%) of participants ( $n=928/1453$ ) drafted an outline that was insufficiently round (e.g., oval or distorted). Other prevalent errors included 626 of 1445 (43.22%) participants displaced a non-anchor numeral ( $n=626/1445$ , 43.22%) 1 or 1.5 half-hour positions from their measured correct location on the numeral array (e.g., placing a "2" in the "1" location) and 626 of 1436 participants (43.59%) drew their hands shifted their hands up or down from the true center of their numeral array.

Features with high assigned point values were less common. For instance, only 15 of 1467 participants (1.02%) did not construct a circular outline and 4 of 1468 participants (0.27%) omitted hands or included a clock with three ( $n=10/1468$ , 0.68%) or four or more hands ( $n=1/1468$ , 0.07%).

### Command Condition Error Scores [%]

Table 3a displays the mean and percentages for the CmSES stratified by gender, and age, and Table 3b provides the overall mean LCmSES, raw summary error scores, and percentage error scores by subscale for education categories. Log transformed scores were also provided because raw scores provide normative values of performance, while transformed scores allowed comparison across skewed CDT measures. Considering the distributions were skewed and most participant clocks had low error scores, we analyzed participant error percentages overall ( $>0$  points) and at higher error intervals ( $>0.25$  points,  $>0.5$  points) for each subscale to compare percentages of more critical errors. Error performance was found to be different across both age ( $p<0.001$ ) and education ( $p<0.001$ ), as well as for specific subscores (see Tables 3a & 3b). Numeral and time-setting subscale errors percentages overall and at higher error intervals were the most highly associated with age and education.

### Copy Condition Error Scores [%]

Similar to the Command Condition, Table 4a illustrates for the Copy Condition the comparison of LCoSES, raw error summary scores, and error percentage scores by gender and age and Table 4b presents the overall mean LCoSES, raw summary error scores, and error percentage scores by education. Again, log transformed scores were included in the table to provide normative distributions and comparative values across skewed measures.

Error performance was found to be significantly different across both age ( $p < 0.004$ ) and education ( $p < 0.001$ ). The time-setting subscale error percentage scores were most highly associated with age and education.

**Command and Copy Condition Errors Scores - Age  $\times$  Education**—The LCmSES and LCoSES and raw errors scores for the Command and Copy condition by age $\times$ education are shown in Table 5. Log transformed scores again allowed for across measure comparisons. The lowest education level (e.g., <high school) was not included because there were too few subjects. Significant differences were found between education groups on the Command condition for all but the oldest age group (75+).

## DISCUSSION

This study reported CDT findings on a large, community-based sample across both broader age and educational categories compared to earlier published normative data. The comprehensive FHS-CDT-SP scoring system, as applied to this sample, supports previous suggestions that the distribution of errors among participants with normal clock drawings is greater than has been published to date (Hubbard et al., 2008). The data derived from the FHS-CDT-SP suggest that age, education, and gender may need to be considered when evaluating whether specific types of errors made on the CDT are unusual relative and potentially clinically significant to a population of healthy aging individuals.

A major advantage of this study is the large cohort sample. The data show that presumed cognitively intact individuals make few errors of probable clinical significance; however, patterns emerged that suggest that older age and lower education are related to the commission of more errors. The overall mean transformed participant error score was 0.95 (raw error score = 1.93) on the clock Command condition and 0.77 on the Copy condition (raw score = 1.33) out of 3.02 possible points (raw score = 20.5 possible points). The overall mean error score is higher on the Command condition because presumably drawing a free-hand clock requires semantic memory and a greater dependence on abstract skills than copying a clock (Freedman et al., 1994).

Participants committed a multitude of minor errors on each of the major tasks involved in drawing or copying a clock. The most common errors included clocks with directional diameters (i.e., the outline was not round), clearly irregularly shaped outlines, displaced numerals from their correct half-hour location, hour and minute hands of the same length, and hands that were shifted from the center of the numeral array. Participants were not instructed to draw a precise clock and some were aware that they were being timed, so many of these minor errors may be a result of imprecision or rushing under pressure. Clocks rarely demonstrated errors on features with higher assigned point values such as an absent outline, absent numerals, duplication or omission of numerals, no hands or more than two hands, or hands that do not essentially meet in the center of the numeral array. These results demonstrate that although relatively infrequent, healthy aging participants do make errors.

### Gender Differences

Results show that gender differences on CDT performance are mixed. The overall Command and Copy error scores did not show a gender difference. Females made a significantly greater percentage of errors in centering their clock hands on the Command and Copy conditions. Men had significantly higher error percentage scores on time-setting on the Copy condition, but marginally significant lower error percentages scores on time-setting errors higher than 0.5 points on the Command condition. Men also had higher percentage scores on outline errors greater than 0.25 points on the Command condition and “other” subtype errors greater than 0.5 points on the Copy condition. Our results suggest that it is

not likely there are significant gender differences overall, but females may make more errors in centering their clock, while men may make more time-setting or outline errors. Further study is necessary to discern whether gender differences have clinical relevance.

## Education

Overall CDT error scores were lower in individuals who reported a college education on Copy and Command conditions. Error percentage scores were also slightly lower for individuals with higher education on the center subscale, and significantly lower on numerals, time-setting, and “other” subscales on the Command condition, and on numerals, time-setting, and “other” error subscales when the analysis for errors >0.5 points was performed. Error percentage scores among those with more education decreased significantly on the time-setting subscale overall, and on time-setting and center subscales for errors higher than 0.5 points on the Copy condition. These results are consistent with previous reports of better performance on the CDT was associated with the highest levels of reported education (Kim & Chey, 2010; von Gunten et al., 2008). In addition, significant differences occurred between the error percentages among the education groups on the center subscale of the Copy condition, but those with less than a high school education and college degree scored similarly, while those with a high school and some college had higher error percentages. This inconsistent finding may be due to the fact that the less than high school education group is small (n=42) compared to the other education groups (high school, n=473; some college, n=371; college degree, n=590). Numeral and time-setting subscale error percentages were the most highly associated with education on the Command condition and time-setting error percentages on the Copy condition. Time-setting errors percentages scores were more significant on the Command than Copy condition. These results may have clinical relevance, particularly with a higher percentage of time-setting errors among those with lower education on both the Command and Copy condition. Finally, the impact of education level on CDT error scores by age group still showed significant differences on the Command condition for all but the oldest age group (75+). However, significant differences between the education groups disappeared when analyzed by age on the Copy condition. Thus, these education differences should be observed longitudinally to see the relationship between education and CDT performance as individuals continue to age.

## Age

Previous literature reported a decline in CDT performance among cognitively normal adults as they increase in age (Freedman et al., 1994; Spreen & Strauss, 1998), which is corroborated by the current findings. Overall, the mean LCmSES were significantly different among those under 55, between 55 and 65, 65 to 75, and those over 75. The overall mean LCmSES for individuals over 75 was 1.21 compared to 0.84 for those under age 55. We also found significantly different error percentage scores among the age groups on the Command condition numeral, time-setting, and “other” subscales, and on outline, numerals, and time-setting subscales for errors assigned a value of greater than 0.5 points. The overall mean LCoSES increased across age groups on the Copy condition. Furthermore, time-setting and center subscale error percentages and time-setting and numeral subscale error percentages greater than 0.5 points were significantly more prevalent in the 75 and older age group on the Copy Condition. Freedman et al. (1994) similarly found that the greatest increase in error scores occurs above age 70. Similar to the results found for education, numeral and time-setting subscale error percentages scores were most highly associated with age on Command Condition and time-setting error percentages on the Copy Condition. Again, the associations were stronger on the Command condition than the Copy condition, and further research is necessary to determine if there is clinical relevance. Results suggest that as cognitively healthy individuals age, they may show subtle signs of difficulty in recall



and planning or sequencing associated with more frequent errors especially on numerals and time-setting, while minor errors in creating an outline may not change significantly.

## Limitations

This study has several limitations. The Offspring cohort is relatively well-educated, with 39.9% of the sample having at least a college degree and 97% of the sample finished high school. Analyses for the lowest education level (less than high school) was likely underpowered, masking any potential relationship between low education and poorer CDT performance. Also, this was a relatively young study sample, with fewer subjects in the highest age category (75+) compared to the younger age categories. Given that age and education were significant factors in predicting error scores, these data may be an underestimate of error scores in the general population. Further, the participants were Caucasian and thus results could not address the mixed findings of whether CDT performance is affected by race. Further, these results may not be generalizable to other ethnically diverse populations. This study also only examined cross-sectional clock drawing data. Longitudinal follow-up is necessary to determine whether these performance patterns are clinically meaningful.

## Conclusions

These normative data provide the distribution of quantified error scores on CDT performance in a cohort population. Results suggest that a variety of errors are seen on CDT Command and Copy conditions in a healthy aging population, and these errors are more prevalent in older compared with younger individuals, and may be less prevalent in individuals who completed at least some college compared with those who did not. Future studies are needed to determine whether this expanded scoring system allows detection of preclinical symptoms of future risk for dementia.

## Acknowledgments

The project described was supported by grants from the National Institute of Neurological Disorders and Stroke (grant number NS17950); the National Institute of Aging (grant numbers AG08122, AG16495); and the National Heart, Lung and Blood Institute's Framingham Heart Study (grant number NIH/NHLBI Contract # N01-HC-25195). Edith Kaplan, Ph.D. and Helen Denison, Ph.D, two core contributors to the development of the Framingham Heart Study Clock Scoring Protocol, are now deceased. Dr. Kaplan has influenced generations of neuropsychologists on test administration and interpretation, and is world renown for her expertise in analysis of the Clock Drawing Test. Dr. Denison's clinical wisdom has also guided many in the field and was critical in the creation of this scoring protocol. They are forever missed for what they have done and could still be doing.

## REFERENCES

- Bourke J, Castleden M, Stephen R, Dennis M. A comparison of clock and pentagon drawing in Alzheimer's disease. *International Journal of Geriatric Psychiatry*. 1995; 10:703–705.
- Brodsky H, Moore CM. The Clock Drawing Test for dementia of the Alzheimer's type: A comparison of three scoring methods in a memory disorders clinic. *International Journal of Geriatric Psychiatry*. 1997; 12:619–627. [PubMed: 9215942]
- Crowe M, Clay OJ, Sawyer P, Crowther MR, Allman RM. Education and reading ability in relation to differences in cognitive screening between African American and Caucasian older adults. *International Journal of Geriatric Psychiatry*. 2008; 23:222–223. [PubMed: 17907267]
- Crowe M, Allman RM, Triebel K, Sawyer P, Martin RC. Normative performance on an executive clock drawing task (CLOX) in a community-dwelling sample of older adults. *Archives of Clinical Neuropsychology*. 2010; 25:610–617. [PubMed: 20601672]
- De Jager CA, Hogervorst E, Combrinck M, Budge MM. Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment and Alzheimer's disease. *Psychological Medicine*. 2003; 33:1039–1050. [PubMed: 12946088]

- Devine, S.; Au, R.; Du, Y.; Beiser, A.; Denison, H.; Rinn, W.; O'Conner, M.; Seshadri, S.; Wolf, P.; Kaplan, E. Normative data for the clock drawing test: Results from the Framingham Heart Study; Poster presented to the Annual Meeting of the International Neuropsychological Society, Maui, HI; 2007 Aug.
- Freedman, M.; Leach, L.; Kaplan, E.; Winocur, G.; Shulman, KI.; Delis, D. Clock Drawing: A Neuropsychological Analysis. New York: Oxford University Press; 1994.
- Garrison RJ, Kannel WB, Stokes J III, Castelli WP. Incidence and precursors of hypertension in young adults: The Framingham Offspring Study. *Preventive Medicine*. 1987; 16:235–251. [PubMed: 3588564]
- Heinik J, Solomesh I, Raikher B, Lin R. 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*. 2002; 17:699–703. [PubMed: 12211117]
- Heinik J, Solomesh I, Berkman P. Correlation between the CAMCOG, the MMSE, and three clock drawing tests in a specialized outpatient psychogeriatric service. *Archives of Gerontology and Geriatrics*. 2004; 38:77–84. [PubMed: 14599707]
- Henderson M, Hotopf M. Use of the clock-drawing test in a hospice population. *Palliative Medicine*. 2007; 21:559–565. [PubMed: 17942493]
- Herrmann N, Kidron D, Shulman KI, Kaplan E, Binns M, Soni J, Leach L, Freedman M. The use of clock tests in schizophrenia. *General Hospital Psychiatry*. 1999; 21:70–73. [PubMed: 10068923]
- Hubbard EJ, Santini V, Blankevoort CG, Volkens KM, Barrup MS, Byerly L, Chaisson C, Jefferson AL, Kaplan E, Green RC, Stern RA. Clock drawing performance in cognitively normal elderly. *Archives of Clinical Neuropsychology*. 2008; 23:295–327. [PubMed: 18243644]
- Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *American Journal of Epidemiology*. 1979; 110:281–290. [PubMed: 474565]
- Kim H, Jeanyung C. Effects of education, literacy, and dementia in the Clock Drawing Test Performance. *Journal of the International Neuropsychological Society*. 2010; 16:1138–1146. [PubMed: 20961480]
- Lam LCW, Chui HFK, Ng KO, Chan C, Chan WF, Li SW, Wong M. Clock-face drawing, reading and setting tests in the screening of dementia in the Chinese Elderly Adults. *Journal of Gerontology: Psychological Sciences*. 1998; 53B(6):353–357.
- Lee KS, Kim EA, Hong CH, Lee DW, Oh BH, Cheong HK. Clock drawing test in mild cognitive impairment: Quantitative analysis of four scoring methods and qualitative analysis. *Dementia and Geriatric Cognitive Disorders*. 2008; 26:483–489. [PubMed: 18987468]
- Lessig MC, Scanlan JM, Nazemi H, Borson S. Time that tells: Critical clock-drawing errors for dementia screening. *International Psychogeriatrics*. 2008; 20:459–470. [PubMed: 17908348]
- Libon DJ, Malamut BL, Swenson R, Sands LP, Cloud BS. Further analyses of clock drawings among demented and nondemented older subjects. *Archives of Clinical Neuropsychology*. 1996; 11:193–205. [PubMed: 14588923]
- Manly JJ, Byrd DA, Touradji P, Stern Y. Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*. 2004; 11(1):37–46. [PubMed: 15471745]
- Marcopulos BA, McLain CA. Are our norms “normal”? A 4-year follow up study of a biracial sample of rural elders with low education. *Clinical Neuropsychologist*. 2003; 17(1):19–33. [PubMed: 12854008]
- Nishiwaki Y, Breeze E, Smeeth L, Bulpitt CJ, Peters R, Fletcher AE. Validity of the Clock-Drawing Test as a screening tool for cognitive impairment in the elderly. *American Journal of Epidemiology*. 2004; 160:797–807. [PubMed: 15466502]
- Peterson RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokman E. Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*. 1999; 56(3):303–308. [PubMed: 10190820]
- Powlishta KK, Von Dras DD, Stanford A, Carr DB, Tsering C, Miller JP, Morris JC. The clock drawing test is a poor screen for very mild dementia. *Neurology*. 2002; 59:898–903. [PubMed: 12297574]

- Ratcliff G, Dodge H, Birzescu M, Ganguli M. Tracking cognitive function over time: Ten-year longitudinal data from a community-based study. *Applied Neuropsychology*. 2003; 10(2):76–88. [PubMed: 12788682]
- 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*. 2002; 17:1048–1054. [PubMed: 12404654]
- Seshadri S, Beiser A, Kelly-Hayes M, Kase CS, Au R, Kannel W, Wolf PA. The lifetime risk of stroke: Estimates from the Framingham Heart Study. *Stroke*. 2006; 37:345–350. [PubMed: 16397184]
- Shulman KI. Clock-drawing: Is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*. 2000; 15:548–561. [PubMed: 10861923]
- Shulman KI, Pushkar Gold D, Cohen CA, Zuccherro CA. Clock-drawing and dementia in the community: A longitudinal study. *International Journal of Geriatric Psychiatry*. 1993; 8:487–496.
- Shulman K, Shedletsky R, Silver I. The challenge of time: Clock drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*. 1986; 1:135–140.
- Spreen, O.; Strauss, E. *A Compendium of Neuropsychological tests: Administration, norms and commentary*. 2nd Ed. New York: Oxford University Press; 1998.
- Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011; 7:280–292.
- Strauss, E.; Sherman, S.; Spreen, O. *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. 3 Ed.. New York: Oxford University Press; 2006.
- Sunderland T, Hill JL, Mellow AM, Lawlor BA, Gundersheimer J, Newhouse PA, Grafman JH. Clock drawing in Alzheimer's disease: A novel measure of dementia severity. *Journal of the American Geriatrics Society*. 1989; 37:725–729. [PubMed: 2754157]
- Ueda H, Kitabayashi Y, Narumoto J, Nakamura K, Kita H, Kishikawa Y, Kenji F. 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 Neurosciences*. 2002; 56:25–29. [PubMed: 11929568]
- Wolf-Klein GP, Silverstone FA, Levy AP, Brod MS. Screening for Alzheimer's disease by clock drawing. *Journal of the American Geriatrics Society*. 1989; 37:730–734. [PubMed: 2754158]
- von Gunten A, Ostos-Wiechetek M, Brull J, Vaudaux-Pisquem I, Cattin S, Duc R. Clock-drawing test performance in the normal elderly and its dependence on age and education. *European Neurology*. 2008; 60:73–78. [PubMed: 18480611]
- Yamamoto S, Mogi N, Umegaki H, Suzuki Y, Ando F, Shimokata H, Akihisa I. The clock drawing test as a valid screening method for mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*. 2004; 18:172–179. [PubMed: 15211073]

## Appendix A

Clock Drawing Subscale Error Explanations [Please note: Although examiners captured data on all the following variables, we were still in the process of accurately defining some of the variables, so not all of the following were included in our analysis]

## Outline

The outline variables relate to the outer form, typically a circle, drawn by the participant.

1. Clock outline is present: This variable codes whether an outline is present, and, if so, whether it is a circle or some other shape (e.g., a square).

2. Outline formed by a continuous curve: The examiner watches the participant draw the outline and codes whether a continuous curve is used or the participant lifts his or her pen.
3. Direction of motion in creating outline: Again, the examiner observes, and records, whether the participant draws the outline in a clockwise, counterclockwise, or mixed direction.
4. Length of longest diameter: Template 2 (Figure 2) is rotated until the x-axis lies along the longest diameter. Scorers record the two radii measurements of this diameter in millimeters.
5. Length of perpendicular bisecting line of the longest diameter: Once the longest diameter is identified and measured, its perpendicular bisector is measured, again in millimeters.
6. Direction of the longest diameter: Recorded is whether the longest diameter is horizontal, vertical, or oblique. If the circle is nearly round, direction of the longest diameter is coded as “not applicable”.
7. Clock outline is *clearly* irregular, angular, lumpy, or has indentations: To be coded as irregular, the irregularity has to be obvious to the examiner looking at the outline; subtle irregularity is not considered.
8. Clock outline perseveratively overdrawn: This is coded as present if the participant draws the circle and then continues to draw along the circle, going more than two times around.
9. Perseveration of clock outline: The outline is considered perseverated if participants draw multiple outlines, without indication that they are trying to correct, or improve the drawing.

## Numerals

1. The presence of numerals: Numerals and numeral substitutes may be absent altogether, numeral substitutes may be present without actual numerals, or numerals can be written out (e.g., “six”), Arabic, Roman, or some combination of these last three.
2. Rotated paper while placing numerals/numeral substitutes: This is coded as present if the participant rotates the paper 180 degrees while placing the numerals, resulting in the numbers in the bottom quadrant being upside-down or angled.
3. Anchor numerals or substitutes placed before any others: Anchor numerals are 12, 3, 6, and 9. Placing these numerals first often assists in spatial placement of the other numerals. All anchors placed before the first non-anchor are recorded (e.g., if a participant started by writing “12,” “6,” “1,” then 12 and 6 are coded).
4. Numerals or substitutes are placed *on or outside* the outline: This can be coded as none, at least one but not all, or all.
5. Measurement of the most medially deviated numeral: Medial deviation refers to how far the numerals are from the drawn outline. Template 1 (Figure 1) is used to find the center of the numeral array. Template 2 (Figure 2) is used to find the numeral that is furthest from the drawn outline (toward the center) by measuring the line that is projected through the center of that numeral to find the location on the outline on which to measure. Using the outside edge of the number, scorers

measure the distance to the inside edge of the drawn circle, measured in millimeters.

6. Displacement of anchor (12, 3, 6, 9): Template 1 is used to find the center of the numeral array. Template 2 is aligned with the numbers on the drawn clock such that it is the “best fit” for the anchor numeral. The anchor that is most displaced from the correct half-hour position is measured. Starting at the numerals true hour line, scorers count the number of ½ hour lines to the location of the displaced number.
7. Displacement of nonanchor (1, 2, 4, 5, 7, 8, 10, 11): Without moving the template from the position determined by the anchor numerals, the nonanchor numeral that is furthest from where it should be is measured in the same way as displacement of anchor.
8. Dots, words, symbols, or other marks are substituted for one or more numerals.
9. Two 12s are present
10. Duplication of numerals (other than 12)
11. Omission of numerals
12. Numerals (or substitutes) beyond 12 are present.
13. Sequencing errors: Sequencing errors refer to placement of numerals that do not follow the correct order (e.g., 1, 2, 4, 3). This does *not* include omitted numerals, duplicated numerals, or numerals placed in a counter-clockwise order (e.g., the 1 in the 11 position, 2 in the 10 position).

## Time Setting

As mentioned above, participants are asked to place the hands to “ten after eleven.”

1. Number of drawn hands: This is coded as “none,” “one,” “two,” “three,” or “four or more.”
2. One hand is correctly pointing to the 11 AND the 11 is in the correct location.
  - a. If not, then the 11 OR 11 position are indicated in some manner (e.g., a circle is drawn around the 11, but no hand is pointing to it or a hand points to the “11” location but there is no “11” in that location).
3. One hand is correctly pointing to the 2 AND the 2 is in the correct location.
  - a. If not, then the 2 OR 2 position is indicated in some manner.
4. One hand incorrectly points to the 10.
5. Length of the hour hand versus the minute hand: This is measured as minute hand longer, hour hand longer, or hands are equal.

## Center

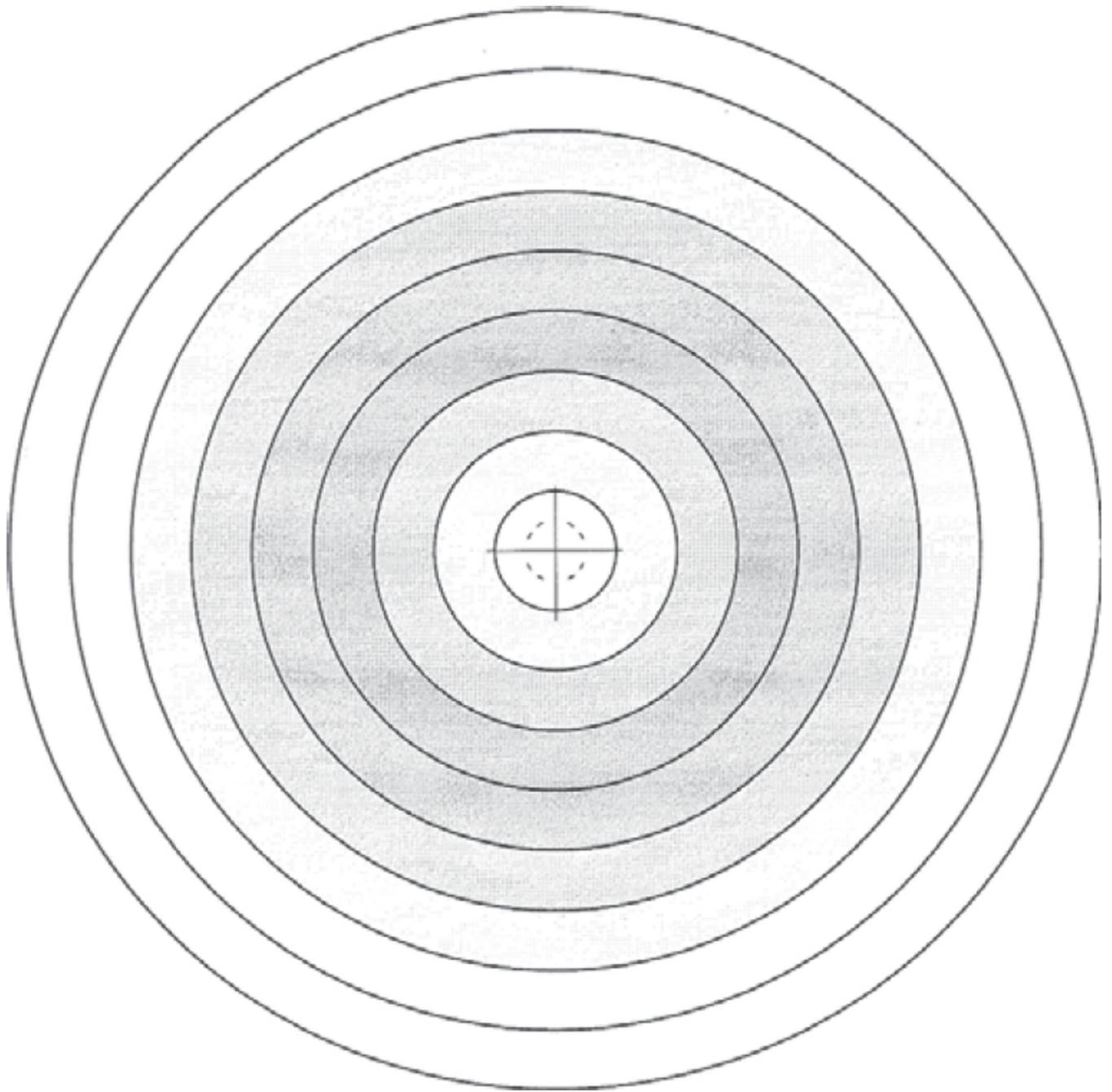
The center is determined by the intersection point of two or more hands. Two hands must be present, and they must *essentially* meet. That is, if the ends of the hands come close enough that it is clear where they would meet if their end points did connect, then that determined intersection is measured. Template 1 and 2 are used to determine the location of the center relative to the numeral array.

1. Two hands are present AND they essentially meet.

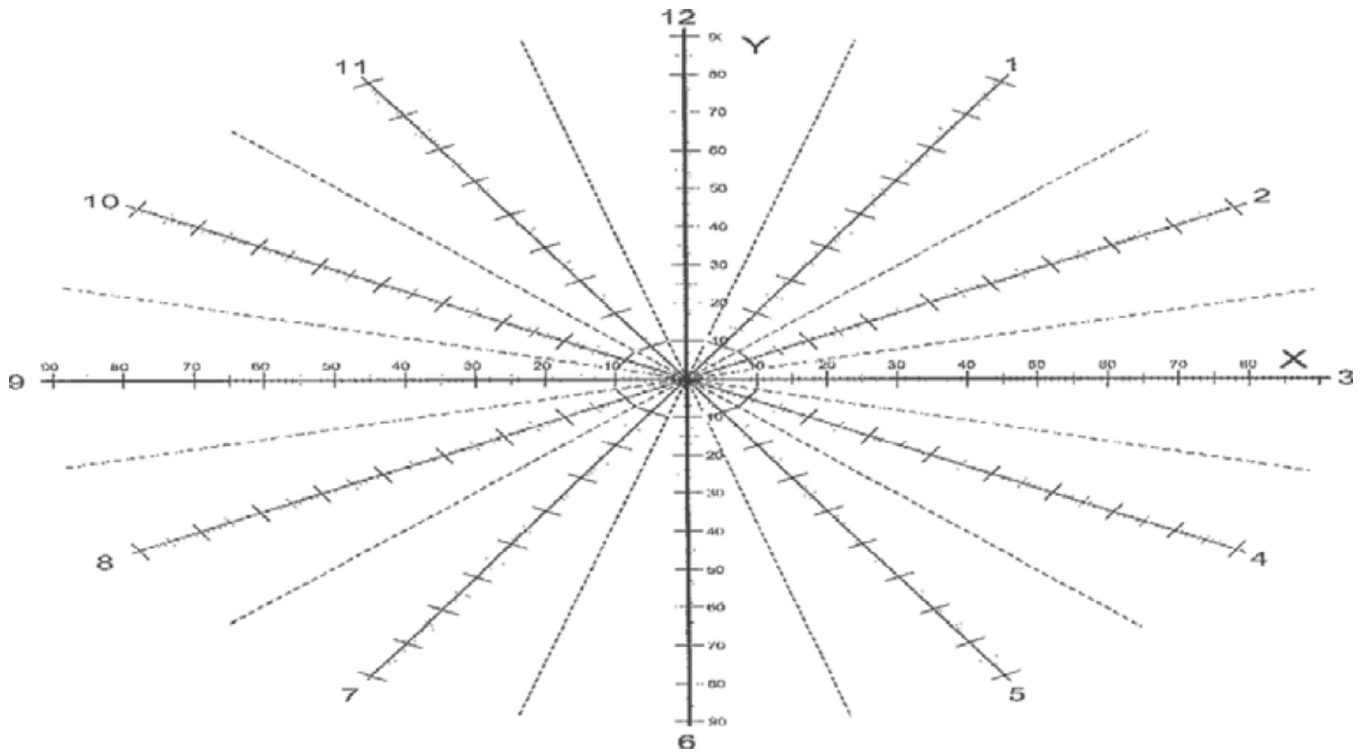
2. On the horizontal axis, the point of intersection of the hands is left, center, or right.
3. On the vertical axis, the point of intersection of the hands is down, center, or up.

### Other Variables

1. Participant asks for a reminder of the time for time-setting.
2. Attempt to self-correct any error: If no errors are present, this variable is coded as “not applicable.” If at least one error is present, this variable is scored as either “no” (there was no attempt to correct an error), “Yes, correct” (the correction resulted in an error-free clock), or “Yes, NOT correct” (a correction was attempted, but it did not lead to an error-free clock).
3. Time to completion.
4. Extraneous marks in clock: This refers to extra marks such as bisecting lines, criss-crossing lines, or radiating lines. Tick marks used as numeral place markers are *not* considered extraneous marks.
5. Other observations: Any unusual characteristics of the drawn clock are indicated here, with a text description.
6. Tester’s clinical assessment of clock drawing: The scorer makes a judgment about whether the overall clock is normal, mildly impaired, moderately impaired, or severely impaired.



**Figure 1.**  
CDT Template 1: 9 cm × 9 cm.



**Figure 2.**  
CDT Template 2: 9 cm × 9 cm.



Table 1

## FHS-CDT-SP Error Score Rating System By Subtype

Qualitative Features	Possible Points	#/total N	%
Clock Outline (range 0–1.5)			
<b>Outline is:</b>			
Present	0	1452/1467	98.98
Absent	1.5	0	0
Present but not a circle	1	15/1467	1.02
A continuous curve	0	1400/1468	95.37
Not a continuous curve	0.25	68/1468	4.63
Not irregular, lumpy, or has indentation	0	1203/1453	82.79
Clearly irregular, angular, lumpy, or has indentation	0.5	250/1453	17.21
<b>Longest diameter is:</b>			
≥ 25 mm	0	1445/1451	99.59
< 25 mm	0.5	6/1451	0.41
Not directional (i.e. all diameters are essentially equal)	0	525/1453	36.13
Directional (i.e. vertical, horizontal, or oblique)	0.25	928/1453	63.87
<b>Perpendicular diameter is:</b>			
≥ 25 mm	0	1437/1450	99.1
< 25 mm	0.5	13/1450	0.9
Numerals (range 0–9)			
<b>Numerals or numeral substitutes are:</b>			
Present	0	1474/1476	99.86
Absent	5.5	2/1476	0.14
<b>Numerals are represented by:</b>			
Only substitutes (dots, dashes, symbols, words)	4	10/1466	0.68
Written out only	0.25	0	0
Only Arabic or Roman numerals	0	1456/1466	98.64
>1 form (Arabic, Roman, Written out)	0.5	0	0
No substitutes (dots, dashes, symbols, words)	0	1449/1466	98.84
At least one, but not all substitutes	1	17/1466	1.16
<b>While placing numerals/substitutes, paper is:</b>			
Not rotated	0	1459/1467	99.46
Rotated	0.5	8/1467	0.54
<b>Number of 12's present:</b>			
One	0	1460/1462	99.86
Two or more	1	2/1462	0.14
<b>Duplication of numerals (1–11 only):</b>			
Absent	0	1446/1462	98.91
Present	1	16/1462	1.09

Qualitative Features	Possible Points	#/total N	%
<b>Omission of numeral(s):</b>			
Absent (i.e. all numerals are drawn)	0	1414/1464	96.58
Present (i.e. at least one numeral is omitted)	1	50/1464	3.42
<b>Numerals (or substitutes) beyond 12:</b>			
Absent	0	1462/1464	99.86
Present	1	2/1464	0.14
<b>Anchor is displaced (correct half-hour positions [value of 1 correct position= 15 degrees]):</b>			
0 or 0.5	0	1051/1439	73.03
1 or 1.5	0.5	337/1439	23.42
2–4	0.75	49/1439	3.41
>4	1	2/1439	0.14
<b>Non-anchor is displaced (correct half-hour positions):</b>			
0 or 0.5	0	757/1445	52.39
1 or 1.5	0.5	626/1445	43.22
2–4	0.75	61/1445	4.22
>4	1	1/1445	0.07
<b>Medial deviation is:</b>			
< the perpendicular diameter divided by 5	0	1386/1464	94.67
the perpendicular diameter divided by 5	1	78/1464	5.33
<b>Time-Setting (range 0–7)</b>			
<hr/>			
One hand incorrectly points to the “10”	1	68/1464	4.64
There is NOT a hand incorrectly pointing to the “10”	0	1396/1464	95.36
<b>The “11” location and numeral are correctly targeted in some manner:</b>			
Both	0	1424/1466	97.14
Either (i.e. the location <u>or</u> the “11” is targeted, but not both)	1	25/1466	1.71
Neither	2	17/1466	1.16
<b>The “2” location and numeral are correctly targeted in some manner:</b>			
Both	0	1357/1466	92.56
Either (i.e. the location <u>or</u> the “2” is targeted, but not both)	1	38/1466	2.59
Neither	2	71/1466	4.84
<b>Two hands are present and:</b>			
Minute hand is longer than the hour hand	0	1052/1430	73.57
Minute hand is shorter than the hour hand	1	150/1430	10.49
Hands are equal length	0.5	228/1430	15.94
<b>Number of hands present (if other than 2):</b>			
Four	0.75	1/1468	0.07
Three	0.5	10/1468	0.68
One	1	15/1468	1.02
None	3	4/1468	0.27
<hr/>			
Center (range 0–1)			

Qualitative Features	Possible Points	#/total N	%
<b>At least two hands are present and:</b>			
Meet to form a "center"	0	1437/1450	99.10
Do not essentially meet	1	13/1450	0.90
<b>"Center" is inside the outline and shifted left or right:</b>			
Absent (i.e. the hands meet in the true horizontal center of the numeral array)	0	1024/1436	71.31
Present (i.e. the hands meet in a location shifted horizontally from the true center of the numeral array)	0.5	412/1436	28.69
<b>"Center" is inside the outline and shifted up or down:</b>			
Absent (i.e. the hands meet in the true vertical center of the numeral array)	0	810/1436	56.41
Present (i.e. the hands are shifted vertically from the true center of the numeral array)	0.5	626/1436	43.59
Other (range 0–2)			
<b>Extraneous marks (bisecting, criss-crossing, or radiating line(s); tick marks not used as number place marks, etc.) are:</b>			
Absent	0	1425/1468	97.07
Present	1	43/1468	2.93
<b>Self-corrected errors:</b>			
No errors requiring correction	0	691/1467	47.10
Successful attempt	0.25	322/1467	21.95
Unsuccessful attempt	0.5	27/1467	1.84
No attempt	1	427/1467	29.11

**Table 2**

## Offspring Cohort Descriptive Statistics

Characteristics	Overall	Men	Women
	<b>N = 1476</b>	<b>N = 680</b>	<b>N = 796</b>
Age at NP L2	67.46±8.93	67.61±8.82	67.33±9.02
[Range]	[43, 91]	[44, 87]	[43, 91]
MMSE Score	29.03±1.26	28.82±1.40	29.20±1.09
[Range]	[20, 30]	[20, 30]	[22, 30]
Education(%)			
Below High School	2.9	3.8	2.0
High School	32.1	26.2	37.1
Some College	25.1	21.9	27.9
College Degree	39.9	48.1	33.0

*Note.* Results are Mean ± SD unless stated otherwise. L2= Time of follow-up NP testing. MMSE= Mini-Mental State Examination, a brief 30-point questionnaire used to screen for cognitive impairment.

Table 3

a. CDT Command Condition Error Scores by Gender & Age										
	Total Sample	Gender		Age Group					p-value	
		Male	Female	p-value	<55	55-65	65-75	75+		
N	1476	680	796		448	554	392	82		
LCmSES	0.95±0.49	0.94±0.49	0.96±0.50	0.498	0.84±0.47	0.92±0.47	1.06±0.50	1.21±0.53		<0.001
RCmSES Range	1.93±1.57 [0, 11.75]	1.90±1.57 [0, 10]	1.95±1.58 [0, 11.75]		1.59±1.31 [0, 7.25]	1.82±1.42 [0, 9]	2.29±1.85 [0, 11.75]	2.81±1.90 [0, 8.25]		
Command Score[%]										
Outline (>0)	72	73	71	0.630	69	72	73	77		0.093
Numerals (>0)	52	51	53	0.424	43	53	59	72		<0.001
Time Setting (>0)	33	31	34	0.232	21	33	44	51		<0.001
Center (>0)	53	50	56	0.026	52	51	55	64		0.087
Other (>0)	53	51	55	0.121	47	52	60	63		<0.001
Command Score[%] *										
Outline (>0.25)	22	25	19	0.014	19	21	24	35		0.005
Numerals (>0.5)	29	28	29	0.697	20	29	34	50		<0.001
Time Setting (>0.5)	18	16	20	0.068	13	16	26	32		<0.001
Center (>0.5)	21	19	22	0.247	20	20	22	26		0.170
Other (>0.5)	31	32	29	0.137	30	29	32	35		0.366

b. CDT Command Condition Error Scores by Education						
	Total Sample	Education Group			p-value	
		< High School	High School	Some College		
N	1476	42	473	371	590	
LCmSES	0.95±0.49	1.19±0.49	1.05±0.48	0.97±0.48	0.84±0.49	<0.001
RCmSES Range	1.93±1.57 [0, 11.75]	2.71±1.87 [0.25, 8.25]	2.21±1.62 [0, 10]	1.97±1.54 [0, 11.75]	1.62±1.47 [0, 9.25]	
Command Score [%]						
Outline (>0)	72	63	73	74	70	0.603
Numerals (>0)	52	76	59	54	45	<0.001

**b. CDT Command Condition Error Scores by Education**

	Total Sample	Education Group					p-value
		< High School	High School	Some College	College		
Time Setting (>0)	33	54	42	35	23	<0.001	
Center (>0)	53	54	57	56	49	0.013	
Other (>0)	53	66	58	55	47	<0.001	
<b>Command Score [%] *</b>							
Outline (>0.25)	22	24	23	24	20	0.197	
Numerals (>0.5)	29	44	34	31	23	<0.001	
Time Setting (>0.5)	18	37	24	19	12	<0.001	
Center (>0.5)	21	24	21	22	19	0.198	
Other (>0.5)	31	39	34	32	27	0.006	

*Note.*

\* Error score percentages computed at higher assigned point values  
 LCmSES= Natural logarithms of command summary error score  
 RCmSES= Raw command summary error score

Table 4

a. CDT Copy Conditions Error Scores by Gender & Age									
	Total Sample	Gender		p-value	Age Group				p-value
		Male	Female		<55	55-65	65-75	75+	
N	1476	680	796		448	554	392	82	
LCoSES	0.77±0.40	0.76±0.40	0.77±0.39	0.673	0.74±0.41	0.77±0.38	0.77±0.39	0.91±0.43	0.005
RCoSES Range	1.33±0.92 [0, 6.75]	1.32±0.92 [0, 4.75]	1.34±0.92 [0, 6.75]		1.27±0.91 [0, 5.25]	1.32±0.88 [0, 5]	1.33±0.94 [0, 6.75]	1.72±1.14 [0, 4.75]	
Copy Score [%]									
Outline (>0)	68	69	68	0.797	68	68	66	81	0.441
Numerals (>0)	53	51	54	0.167	48	55	54	53	0.111
Time Setting (>0)	15	18	12	<0.001	11	16	15	26	0.003
Center (>0)	45	41	48	0.003	43	45	44	59	0.070
Other (>0)	34	34	34	0.735	35	33	33	38	0.836
Copy Score[%]*									
Outline (>0.25)	21	22	20	0.555	21	19	22	31	0.215
Numerals (>0.5)	18	18	19	0.467	17	17	22	23	0.044
Time Setting (>0.5)	4	4	4	0.627	3	3	5	14	<0.001
Center (>0.5)	17	15	19	0.021	17	17	16	23	0.713
Other (>0.5)	23	26	20	0.012	25	23	19	27	0.191
b. CDT Copy Conditions Error Scores by Education									
	Total Sample	Education Group				p-value			
		< High School	High School	Some College	College				
N	1476	42	473	371	590				
LCoSES	0.77±0.40	0.85±0.44	0.82±0.39	0.76±0.41	0.73±0.38	0.004			
RCoSES Range	1.33±0.92 [0, 6.75]	1.59±1.29 [0, 6.75]	1.44±0.92 [0, 4.75]	1.31±0.93 [0, 5.50]	1.24±0.88 [0, 5.25]				
Copy Score [%]									
Outline (>0)	68	62	70	65	69	0.918			

**b. CDT Copy Conditions Error Scores by Education**

	Total Sample	Education Group				p-value
		< High School	High School	Some College	College	
Numerals (>0)	53	62	56	51	51	0.094
Time Setting (>0)	15	31	16	17	12	0.003
Center (>0)	45	38	50	45	41	0.011
Other (>0)	34	46	35	33	33	0.211
<b>Copy Score [%] *</b>						
Outline (>0.25)	21	21	24	18	21	0.381
Numerals (>0.5)	18	13	21	20	16	0.097
Time Setting (>0.5)	4	10	5	4	3	0.047
Center (>0.5)	17	23	20	17	15	0.019
Other (>0.5)	23	33	24	22	21	0.098

Note:

\* Error score percentages computed at higher assigned point values  
 LCoSES= Natural logarithms of copy summary error score  
 RCoSES= Raw copy summary error score



**Table 5**

CDT Command and Copy Condition Error Scores By Age × Education

Age Group	Education Group			p-value
	High School	Some College	College	
<55 years	n	100	108	237
	LCmSES	0.97±0.47	0.90±0.47	0.75±0.46
	RCmSES	1.94±1.43	1.75±1.37	1.36±1.18
	Range	[0, 6.75]	[0, 7.25]	[0, 7.25]
55–65 years	LCoSES	0.79±0.40	0.73±0.41	0.72±0.41
	RCoSES	1.38±0.94	1.25±0.84	1.23±0.93
	Range	[0, 4.25]	[0, 3.75]	[0, 5.25]
65–75 years	n	179	156	208
	LCmSES	0.98±0.47	0.94±0.46	0.85±0.48
	RCmSES	1.97±1.35	1.84±1.37	1.64±1.49
	Range	[0, 6.25]	[0, 9.00]	[0, 8.50]
75+ years	LCoSES	0.82±0.37	0.74±0.40	0.74±0.37
	RCoSES	1.43±0.84	1.27±0.93	1.25±0.85
	Range	[0, 4.50]	[0, 4.75]	[0, 5.00]
65–75 years	n	157	86	122
	LCmSES	1.12±0.49	1.08±0.48	0.94±0.50
	RCmSES	2.46±1.83	2.32±1.85	1.94±1.75
	Range	[0, 10.00]	[0, 11.75]	[0, 9.25]
65–75 years	LCoSES	0.78±0.39	0.79±0.43	0.73±0.35
	RCoSES	1.36±0.90	1.42±1.03	1.21±0.77
	Range	[0, 4.50]	[0, 5.50]	[0, 3.50]
75+ years	n	37	21	23
	LCmSES	1.29±0.50	1.17±0.56	1.11±0.57
	RCmSES	3.06±1.98	2.69±1.84	2.51±1.86
	Range	[0.25, 8.25]	[0, 6.75]	[0, 6.50]
LCoSES	1.00±0.46	0.86±0.40	0.83±0.40	0.2786

Age Group	Education Group	Some College	College	p-value
RCoSES	High School	1.54±0.92	1.48±1.12	
Range	[0, 4.75]	[0, 3.00]	[0.25, 4.25]	

Note.  
 LCmSES= Natural logarithms of command summary error score  
 LCoSES= Natural logarithms of copy summary error score  
 RCmSES= Raw command summary error score  
 RCoSES= Raw copy summary error score