A Reduced Scoring System for the Clock-Drawing Test Using a Population-Based Sample

by

Alexandra Jouk
B.A., Scripps College, 2007

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

A Reduced Scoring System for the Clock-Drawing Test

Using a Population-Based Sample

by

Alexandra Jouk
B.A., Scripps College, 2007

Supervisory Committee:

Dr. Holly A. Tuokko, (Department of Psychology)
Supervisor

Dr. Mauricio Garcia-Barrera, (Department of Psychology)
Departmental Member

Dr. Stuart MacDonald, (Department of Psychology)
Departmental Member

Abstract

Supervisory Committee:

Dr. Holly A. Tuokko, (Department of Psychology)
Supervisor

Dr. Mauricio Garcia-Barrera, (Department of Psychology)
Departmental Member

Dr. Stuart MacDonald, (Department of Psychology)
Departmental Member

This thesis project examined the generalizabilty of the simplified scoring system for the Clock Drawing Test developed by Lessig and her colleagues in 2008 using a population-based sample. Clock-drawings from 356 participants (80 with dementia, 276 healthy controls) from the Canadian Study on Health and Aging were analyzed using logistic regression and Receiver Operating Characteristic curves. The new scoring system reduced the Lessig system down even further to include five critical errors: missing numbers, repeated numbers, number orientation, extra marks, and number distance, and produced a sensitivity of 81% and a specificity of 68%. The results from this study improve our current state of knowledge concerning the Clock Drawing Task by validating the simplified scoring system proposed by Lessig and her colleagues among a more representative sample and provides further evidence in support of a simple and reliable dementia-screening tool.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor, Dr. Holly Tuokko, for providing me the guidance and encouragement needed to complete this thesis. Her dedication and excellence as a researcher continually motivates me to strive for my best. I would also like to thank my committee members, Dr. Stuart MacDonald and Dr. Mauricio Garcia-Barrera, for their valuable comments and continual support. In addition, I thank Dr. John Walsh for acting as my external department member on my committee. This thesis was funded by the BC Network for Aging Research (BCNAR).

Chapter 1:

Introduction

Neurologists and neuropsychologists have used clock-drawing tasks for decades to examine the cognitive processes underlying many neurological disorders, with particular focus on the assessment of constructional apraxia (Critchley, 1966). In the past 30 years, variations of the Clock-Drawing Test (CDT) have risen to the forefront as a dementia screening tool (Lessig, Scanlan, Nazemi, & Borson, 2008; Manos, 1999; Mendez, Ala, & Underwood, 1992; Royall, Cordes, & Polk, 1998; Rouleau, Salmon, Butters, Kennedy, & McGuire, 1992; Shulman, Gold, Cohen, & Zucchero, 1993; Sunderland, Hill, Mellow, Lawlor, Gundersheimer, Newhouse, & Grafman, 1989; Tuokko, Hadjistavropoulos, Miller, & Beattie, 1992; Watson, Arfken, & Birge, 1993; Wolf-Klein, Silverstone, Levy, Brod, & Breuer, 1989). However, among the many proposed scoring systems used to evaluate the CDT, researchers and clinicians have yet to reach a consensus as to which system best predicts dementia.

Experts in the field agree that if clock-drawing tasks are to be used as a screening measure, they must be brief, generalizable to diverse populations, highly predictive of dementia, quick and easy to score, yet at the same time retain the qualities of a comprehensive examination (Ainslie & Murden, 1993; Borson et al., 1999; Shulman, 2000). This thesis aims to examine whether or not a procedure that was used to identify six critical errors needed for dementia detection (Lessig et al., 2008) in a clinic-based sample would produce similar results in a Canadian population-based subsample.

I. The History and Present Uses of the Clock Test

In the field of neurology and neuropsychology, performance on assessment tests provides doctors with the gateway for understanding how the brain affects behavior. Results from these tests can not only reveal brain abnormalities, but they can also provide insight into a patient's functional capacities and prognosis for treatment. Clock drawing hails as one of the oldest and most widely used neuropsychological assessment tools (Head, 1926; Luria, 1966; Shulman, Herrmann, Brodaty, Chiu, Lawlor, Ritchie, & Scanlan, 2006). For decades, clinicians and later researchers used clock-drawing tasks to assess a multitude of neurological and psychiatric disorders (Freedman, Leach, Kaplan, Winocur, Shulman, & Delis, 1994). The clock-drawing task has withstood the test of time and to this day remains a staple tool among the neuropsychological community, with its frequency of administration trailing only the Mini-Mental Status Exam (MMSE, Folstein, Folstein, & McHugh, 1975; Shulman et al., 2006).

Clock-drawing was first used to assess for constructional apraxia (Critchley, 1966; Goodglass & Kaplan, 1983; Head, 1926; Luria, 1966; Mayer-Gross, 1935; Van der Horst, 1934). The familiarity of a standard clock face made clock-drawing an ideal tool to screen for neurological impairments with drawing or copying. Visuo-spatial deficits commonly found in patients with constructional apraxia could be easily detected using clock-drawing and include frequent errors in hand setting (Critchley, 1966) and comprehending spatial relationships through the interchange of symmetrically opposite points (e.g., the 3 and the 9 are swapped; Luria, 1966).

The clock-drawing task's reliance on visuo-spatial processing led researchers to apply this assessment methodology to other disorders affecting visuoconstructional abilities, including hemi-spatial neglect (Battersby, Bender, Pollack, & Kahn, 1956; Di Pellegrino, 1995). A person with unilateral neglect fails to process information on one side of space, the visual field contra lateral to the lesion (e.g., damage to the left cerebral hemisphere will result in right visual field neglect and damage to the right cerebral hemisphere will result in left visual field neglect; Gazzaniga, Ivry, & Mangun, 2002). A common clock drawing error in a person with hemispatial neglect is the clustering of numbers on one side of the clock face (Figure 1.; Smith, Gilchrist, Butler, Muir, Bone, Reeves, & Harvey, 2007). Due to the clock-drawing task's sensitivity in detecting hemispatial neglect, it can now be found as part of the Behavioural Inattention Test, a visual neglect battery consisting of six paper and pencil tests (Wilson, Cockburn, & Halligan, 1987).

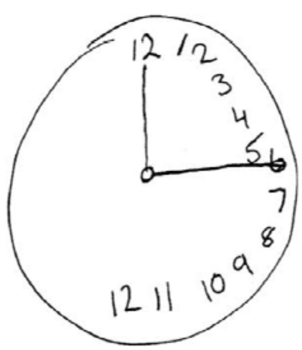


Figure 1. A clock drawing by a patient with hemispatial neglect. Reproduced from Smith et al., 2007.

Although historically the clock-drawing task's first use centered on the assessment of visuo-spatial disorders, more recently its application has shifted to the detection of dementia (Kim, Lee, Choi, Sohn, & Lee, 2009; Lessig et al., 2008; Libon, Swenson, Baronski, & Sands, 1993; Manos, 1999; Mendez et al., 1992; Pinto & Peters, 2009; Royall et al., 1998; Rouleau, Salmon, & Butters, 1996; Rouleau et al., 1992; Shulman et al., 1993; Sunderland et al., 1989; Tuokko et al., 1992, 1995; Watson et al., 1993; Wolf-Klein et al., 1989). Visual and spatial deficits are commonly associated with the early stages of dementia (Freedman et al., 1994), thus making the CDT a valid assessment measure of dementia. Many experts in the field of neurodegenerative disorders, however, argue that CDTs rely more on semantic memory and executive functions than visual-constructive processes (Libon, Malamut, Swenson, Sands, & Cloud, 1996; Rouleau et al., 1992; Tuokko et al., 1992). According to Tuokko and her colleagues (1992), errors in clock drawing in persons with Alzheimer's disease (AD) reflect a deterioration in abstract thinking and reasoning abilities, revealing a fundamental impairment in the understanding of time representation. Deficits in semantic memory and executive functioning (e.g., abstract thinking and reasoning) exemplify two of the core features of AD as defined by the Diagnostic and Statistical Manual of Mental Disorders, 4^{th} ed., text revision (DSM-IV-TR, American Psychological Association, 2000). This makes the CDT a particularly sensitive measure of dementia, specifically of the Alzheimer type. As a result, the co-chair of the national panel spearheading the craftsmanship of dementia-screening guidelines recommended the incorporation of a form of the clock-drawing task into a professional screening battery (Williams, 1992). Since this proposal, clock drawing has been added to a multitude of screening batteries, including the Mini-Cog (Borson, Scanlan, Brush, Vitaliano, & Dokmok, 2000), the 7-Minute Neurocognitive Screening Battery (Solomon, Hirschoff, Kelly, Relin, Brush, DeVeaux, & Pendlebury, 1998) and the Microcog (Powell, Kaplan, Whitla, Weintraub, Catlin, & Funkenstein, 1993). A vast majority of clinicians use some form of clock drawing either alone or as part of a battery not just because of its sensitivity as a dementia screening measure, but also because it is quick and easy to administer, relatively simple to score, requires less language ability than other commonly used screening measures, is fairly resistant to the influences of depression, and is well-tolerated by most patients (Peters & Pinto, 2008; Shulman et al., 2006).

II. Neuroanatomical Correlates of the Clock Drawing Test

The CDT is a seemingly straightforward assessment measure. However, when broken down into its individual components, it proves to be a complex task that requires the synthesis of various neurological functions (e.g., visuoperceptual, visuospatial, visuomotor, executive functioning, auditory) and the brain regions that subserve those functions (e.g., parietal, occipital, frontal, temporal lobes). When one or many of these regions becomes subject to neural damage, the resulting clock drawing can change drastically depending on the placement of the lesion (Freedman et al., 1994). Additionally, a patient's final clock production represents an interaction between his/her deficits and spared functions. Taken together, the type and location of neural pathology coupled with compensating functions can render one person's clock drawing markedly different from another person's.

As mentioned in the previous section, clock drawing was historically first assumed to assess for disruptions in visuospatial processes, including visual recognition and visual memory, in patients with constructional apraxia (Critchley, 1966; Goodglass & Kaplan, 1983; Luria, 1966). In 1874, English neurologist, John Hughlings Jackson, broadly proposed the localization of these processes to the posterior region of the right hemisphere of the brain. At the turn of the 20^{th} century, several neurologists expanded on Jackson's theory and developed the idea of a 'disconnection syndrome' to account for visuospatial disorders (Kliest, 1912; Poppelreuter, 1917). A 'disconnection syndrome' is a blanket term used to describe any neurological condition resulting from an interruption in association pathways localized to one hemisphere or linking the two hemispheres together (Geschwind, 1965). Although the founding fathers of apraxia all supported this disconnection syndrome, they could not agree on its localization. For example, Poppelreuter (1917) believed either or both occipital lobes governed the visuomotor abilities responsible for design copying, whereas Kleist (1912) felt that impairments in figure copying resulted from lesions to the parieto-occipital region of the dominant hemisphere (see Freedman et al., 1994). Recent advances in neuropsychological assessments, imaging techniques, and electrical stimulation have clarified some of the confusion associated with the neural correlates of visuospatial processing. The parietal lobes appear to play a role in many of the cognitive functions needed to complete the CDT (e.g., visual association, long-term memory, spatial relations, movement; Critchley, 1966; Gazzaniga et al., 2002). Split-brain surgical research (cutting the fibrous white matter tracts connecting the left and right hemispheres) has also allowed researchers to better understand hemispheric specialization associated with figure drawing. In 1995, Kingstone and Gazzaniga found that both the left and right hemisphere contribute to drawing a figure, like a clock face, with a slight bias in favor of the left hemisphere.

As greater knowledge in the field of neuroscience amassed over the years, it became clear that many other brain regions in addition to the parietal lobes were responsible for generating a complete rendition of a clock face. When a patient is first given the verbal command to "draw a clock", auditory processes mainly controlled by the temporal lobes (Gazzaniga et al., 2002) are used to decode the verbal information into semantic knowledge. Hippocampal regions are also activated when long-term visual memory and the associated retrieval mechanisms are needed to remember the appropriate clock face layout (Bauer, Grande, & Valenstein, 2003). If verbal instructions are also given to set the clock to a certain time, as in certain versions of the clock task (Goodglass & Kaplan, 1983; Spreen & Strauss, 1998; Tuokko et al., 1992), working memory used to retain the instructions until the patient is ready to perform that portion of the test. Working memory is also primarily associated with activation in the hippocampus, as well as areas in the prefrontal cortex (Baddeley & Della Sala, 1998; Bauer, Grande, & Valenstein, 2003). Accurate visual perception is needed to guide the recreation of a clock's spatial arrangement and has been found to be predominantly associated with parietal lobe function, as noted above (Critchley, 1966; Gazzaniga et al., 2002). Executive functions, associated with the frontal lobes, monitor and correct for any errors, while also assisting in the planning, organization, and coordination of all the multiple steps involved in clock drawing. Finally, motoric processes, controlled primarily by the motor cortex, execute the drawing of the clock (Gazzaniga et al., 2002), whereupon receptive language skills are needed to decode verbal commands into graphomotor representations of the clock numbers, whether they are Arabic numbers or Roman numerals (Freedman et al., 1994). When broken down into its individual components, it becomes clear that the CDT relies on the integration of a wide variety of cortical and subcortical regions.

Recent neuroimaging studies have also contributed to the findings listed above (Cahn-Weiner et al., 1999; Kim, Lee, Choi, Sohn, & Lee, 2009; Nagahama, Okina, Suzuki, Nabatame, & Matsuda, 2005). In 1999, Cahn-Weiner and her colleagues examined clock-drawing performance in 29 patients with probable AD, as indexed by the National Institute of Neurological and Communicative Disease and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria. It was found with magnetic resonance imaging (MRI) that clock-drawing performance in AD patients was significantly correlated with right hemisphere regional gray matter, specifically with right anterior and posterior-superior temporal lobe volumes, compared to normal elderly controls. Nagahama et al. (2005) investigated regional cerebral blood flow (rCBF) in 100 patients with mild and severe forms of AD using single photon emission computed tomography (SPECT) imaging. Impairments in clock-drawings by patients with probable AD related to reduced cerebral blood flow in the post lateral region of the left temporal lobe. These correlative imaging results underscore the general finding that a multitude of diverse brain regions are involved in the CDT.

III. Clock Drawings in Healthy Older Adults and Patients with Dementia

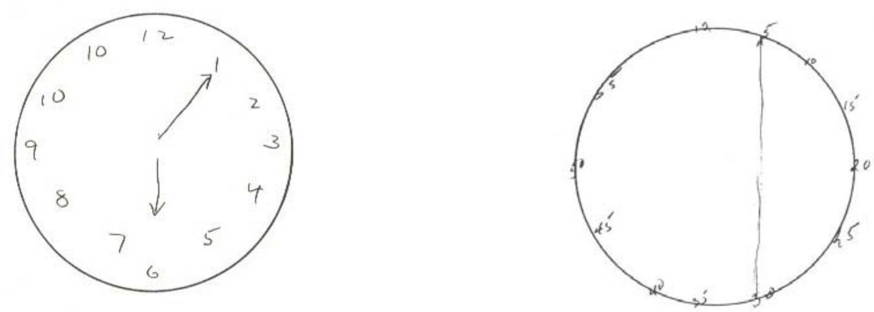
The widespread neurodegeneration that can occur in individuals with dementia, as evidenced above, can impact clock drawing in many different ways. Before a clock drawing can be deemed indicative of dementia, a normative baseline must be established to provide a reference point for comparison. Freedman et al.'s (1994) study compared clock drawings from healthy older adults with the clock drawings produced by individuals with dementia. They found that cognitively intact older adults produced wide-ranging variations in clock drawing. It should be noted that the examples and figures presented below are only representations of a small sub-set of clock drawings, since there are as many clock variations as there are individuals.

Most of the free-drawn clocks were circular or oval in shape (Figure 2A and B) and contained all Arabic numbers “1” through “12”.

A **B**

Figure 2. Acceptable clock contours. (A) 70-year-old male. (B) 52-year-old male. Reproduced from Freedman et al., (1994).

In approximately 10% of the healthy controls, however, errors of number omission, addition, and substitution occurred. Figure 3A presents an example of a pre-drawn clock with the numbers filled in by a healthy 65-year old male. In this clock, the number “11” is omitted with perseveration of the number “10”. Figure 3B illustrates the worst-case scenario of omissions, additions and substitutions drawn by a healthy individual. In this pre-drawn clock, filled in by a healthy 77-year old woman, the anchor-number “12” is properly written, but then the numbers “5 – 55” are placed around the clock in 5-minute increments as a substitute for the numbers “1” through “11”. Hand placement is also poorly represented, with the hour hand and minute hand depicted by the same line length.



A **B**

Figure 3. Examples of number omissions and additions. (A) 65-year-old healthy male; omission of number “10” with perseveration of number “11”. (B) 77-year-old healthy female; substitution of numbers 1 through 12 with “5 – 55” at 5-minute intervals. Hand placement also poorly represented. Reproduced from Freedman et al., (1994).

According to Freedman et al. (1994), hand placement is the most critical component of the CDT. In their normative study, they found that over 92% of the healthy controls properly indicated the hour hand, with more variation occurring with minute hand placement. Figure 4 shows the clock from a healthy 77-year-old woman illustrating an over-representation of the minute hand in conjunction with off-center hand origin. The results from Freedman et al.’s (1994) study indicate that errors on the CDT are a normative aspect of healthy aging, with the older age groups (70-79 and 80+) making more errors on the test compared to their younger counterparts.

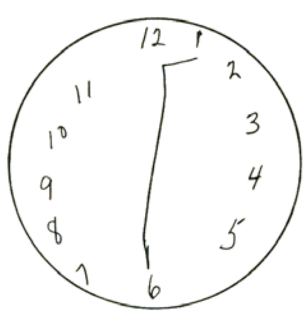
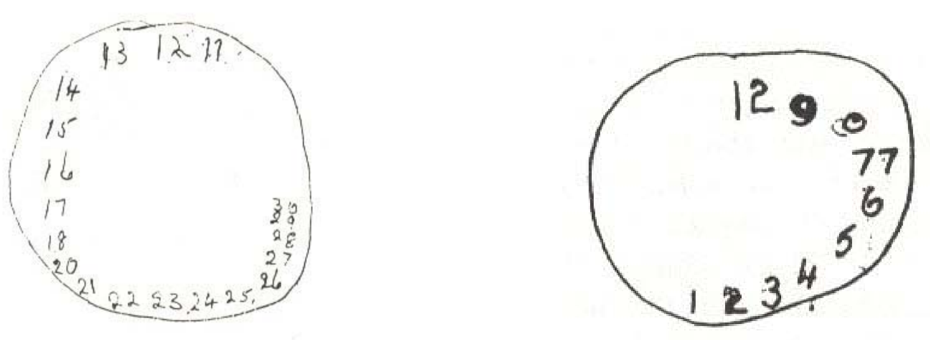


Figure 4. 77-year-old healthy female. Overrepresentation of the minute hand with off-center hand origin. Reproduced from Freedman et al., (1994).

Compared to cognitively healthy older adults, individuals with dementia have been shown to produce significantly more errors on the CDT (Freedman et al., 1994; Tuokko et al., 1992, 1995). Number omission, addition, and substitution were found to be very common errors produced by individuals with dementia (Freedman et al., 1994). Figure 5A shows the clock drawing of a 73-year-old woman with severe Alzheimer's disease. In this clock drawing, the elderly woman omitted the number "1" through "10", added the numbers 13 through 30, and perseverated on the sequence of numbers in the counter-clockwise direction. The clock drawn in figure 5B was created by a 63-year-old woman with dementia who also displays number omission.



A **B**

Figure 5. Examples of number omissions, additions, and substitutions in dementia patients. (A) 73-year-old female. (B) 63-year-old female. Reproduced from Freedman et al., (1994).

In addition to number omissions, additions, and substitutions, the majority of dementia patients exhibited trouble with hand placement. Freedman et al. (1994) found that more severe forms of dementia correlated with the placement of only one hand or no hands on the clock face. Another common error found among individuals with dementia is the inability to recode time-setting instructions. Figure 6 shows the clock drawing of a 70-year-old female who was asked to set her clock to a "quarter to 7". This woman was unable to recode a "quarter" into the "9" and as a result wrote "¼" next to the four on the right side of the clock.

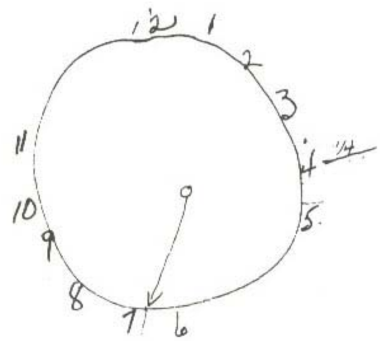


Figure 6. 70-year-old female with Alzheimer's disease. Reproduced from Freedman et al., (1994).

The examples presented in this section only provide a rough summary of some of the common clock-drawing errors found in healthy older adults and those suffering from dementia. The wide-ranging variations in clock drawing, as evidenced above, demonstrate the need for a valid scoring system.

IV. Scoring the Clock Drawing Test: A Review of the Currently Used Scoring Systems

As illustrated above, clock drawings in patients with dementia can vary drastically. Many different administration methods and scoring systems have been developed over the years to best capture the variety of errors found among clock drawings in the dementia population (Freedman et al., 1994; Lessig et al., 2008; Lin et al., 2003; Manos & Wu, 1994; Mendez et al., 1992; Rouleau et al., 1992; Royall et al., 1998; Shulman, Sheldetsky, & Silver, 1986; Tuokko et al., 1992; Sunderland et al., 1989; Watson et al., 1993; Wolf-Klein et al., 1989), however, no single administration procedure and corresponding scoring system has been universally accepted. This may be due, in part, to the fact that the current CDT administration procedures and their corresponding scoring methods have been shown to produce varying degrees of validity (Pinto & Peters, 2008; Tuokko & O'Connell, 2006), making it difficult to choose one particular scoring system over another. It is possible for both false positives and false negatives to occur when using clock drawing as part of a dementia screening battery. In other words, the CDT may occasionally fail to detect all cases of dementia (false negatives) and may conversely misidentify healthy older adults as having dementia (false positives). Due to the inherent imperfections associated with the available clock drawing scoring systems, this test is best used in conjunction with other neuropsychological tests and interpreted by a trained clinician to safeguard against false positives and false negatives during a screening examination.

Several different administration procedures for clock drawing currently exist and can range from slight changes in the wording of the instructions to considerable differences in hand-setting placement. All the CDTs currently in existence fall under two main administration categories: (a) a clock drawn freely (free-hand); and (b) a clock drawn in a pre-outlined circle (pre-drawn). It has been argued that although these two variations in administration practices ultimately use similar cognitive processes, they emphasize different functions (Tuokko & O'Connell, 2006). Pre-drawn clocks place demands on perceptual functioning, whereas free-drawn clocks rely more on memory, language, and executive function abilities (Freedman et al., 1994). Pre-drawn circles can vary in size, but most fall in the range of 3.4 to 4.5 inches (8.6 to 11.4 cm) in diameter (Manos & Wu, 1994; Shulman et al., 1986; Tuokko et al.1995, Watson et al., 1993; Wolf-Klein et al., 1989). Some researchers favor the pre-drawn condition due to its standardized procedure (Shulman et al., 1986). At the same time, many CDT developers utilize the free-drawn administration method (Mendez et al., 1992; Royall et al., 1998; Shua-Haim, Koppuzha, Shua-Haim, & Gross, 1997; Spreen & Strauss, 1998). This technique, however, has been criticized since a poorly drawn circle (e.g., too small, drastically misshapen) can compromise the rest of the task and limit interpretation (Tuokko & O'Connell, 2006). Freedman et al., (1994) suggested a compromise between these two debated procedures by recommending that patients first freely draw the clock face, and only if distortions and/or size limitations arise, should a pre-drawn clock be provided.

Administration procedures can also vary in terms of time settings. Some tests require patients to set clocks to one time setting (e.g., Mendez et al., 1992; Royall et al., 1998; Sunderland et al., 1989; Tuokko et al., 1995) and some do not require any time setting at all (e.g., Watson et al, 1993). The most commonly used time settings in descending order of frequency are "10 past 11", "20 past 8", and "3 o'clock" (Freedman et al., 1994). The clock settings of "10 past 11" and "20 past 8" require the patient to draw a hand in both the left and right visual field, thereby concurrently assessing bilateral attention. These two settings also rely more on semantic processing and abstract thinking since both the "10" and "20" must be mentally recoded in order to be placed in their proper "2" and "4" respective positions on the clock. The key difference between these two settings is that the "10 past 11" setting places the two hands in the upper right and left quadrants – the temporal fields which have been suggested to employ executive functions – whereas the "20 past 8" setting places both hands in the lower right and left quadrants – the parietal fields which have been suggested to mediate visuospatial processes (Freedman et al., 1994). As a dementia-screening test, the "10 past 11" setting has been found to be most sensitive to neurocognitive dysfunction (Freedman et al., 1994).

Over the years, many research groups developed scoring systems to capture the impairments associated with dementia. The collective goals of these varying scoring systems included: (a) quick and easy scoring, done in a matter of minutes; (b) high inter-rater reliability (consensus in scoring outcomes remains among different raters); and (c) high sensitivity and specificity, that is the scoring system will as accurately as possible detect persons with the disease of interest (disease-positive) – in this case dementia – while correctly identifying individuals without the disease (disease-negative). Sensitivity and specificity are typically reported in percentages, with a higher percentage reflecting a more discriminating diagnostic tool (Bewick, Cheek, & Ball, 2004). Detailed information about determining and interpreting sensitivity and specificity is further explored in the Methods section.

According to O'Rourke, Tuokko, Hayden, and Beattie (1997), all clock-drawing test scoring takes approximately 3 – 5 minutes to administer, thus satisfying the first goal described above. Inter-rater reliabilities, sensitivities and specificities, however, vary among the currently used scoring systems, making some systems preferable over others. Additionally, several scoring systems' development was based on clinical observations and empirically invalidated assumptions (e.g., Royall et al., 1998; Shulman et al., 1986; Watson et al., 1993; Wolf-Klein et al., 1989). From this point forward, all discussed scoring systems will be referred to by the first author's name.

Shulman and his colleagues (1986) developed one of the first scoring systems designed to assess for the presence of dementia using the clock-drawing task. Using a pre-drawn circle, individuals are asked to draw a clock and to set the hand to three o'clock. The Shulman method uses a 6-point rating system, where "1" represents a perfectly executed clock drawing resulting from no impairment, and "6" represents the highest level of impairment. Using a cutoff score of greater than 2, Shulman et al. found that their scoring system positively identified 86% of Alzheimer patients (sensitivity), while correctly identifying 72% healthy controls (specificity). This scoring method was shown to have an inter-rater reliability of .74 (Shulman et al., 1986).

Three years later, Wolf-Klein and her colleagues (1989) created a clock-drawing scoring system based purely on number placement, where individuals drew in the clock face on a pre-drawn circle but where not required to set the hands to a given time. Like the Shulman scoring system, Wolf-Klein et al. (1989) used a scale of severity rating scheme, where clocks are scored on a scale from 1 (most impaired) to 10 (no impairment). A cutoff score of 7 and below is indicative of impairment with 86.7% sensitivity and 92.7% specificity. The subjective quality associated with a scale of severity of impairment system does not make the Shulman and Wolf-Klein scoring systems the ideal methods for dementia screening. When using a scale of severity system, individual scorers hold their own assumptions about what constitutes a "normal" or "severe" clock drawing. These assumptions can drastically differ from person to person and interfere with the objectivity and standardization of the scoring system. Many other scoring methods (e.g., Lam, Chiu, Ng, Chan, Chan, Li, & Wong; Libon et al., 1993; Spreen & Strauss, 1998; Sunderland et al., 1989) also employ a scale of severity system, thus subject to the same criticisms as Shulman and Wolf-Klein's scoring procedures (Tuokko & O'Connell, 2006).

Although no other scoring system relies purely on the subjective rating of the scorer, several systems base their methodology on assumptions that have not been empirically supported (e.g., Royall et al., 1998; Watson et al., 1993). Royall and his colleagues developed a scoring system to primarily assess executive functioning abilities in Alzheimer patients. In their study, individuals were given a pre-drawn circle, asked to draw in the face of a clock, and set the hands to 1:45. A 15-point scoring system was derived from the 107 individuals (62 healthy young adults, 45 patients with probable Alzheimer's disease) who completed the CDT. The 15-point system designates 1 point for organizational elements (e.g., the figure resembles a clock, 2 hands are present, the minute hand is longer than the hour hand) and 2 points for intact spacing on either side of the vertical 12-6 axis. No reason is given for assigning 2 points for spacing and only 1 point for all other errors. Additionally, several "errors" (e.g., use of only Arabic numerals, diameter is greater than 1 inch, hands represented as arrows) cannot be considered impairments, as the instructions did not provide specific directions to avoid these details. Furthermore, this scoring system requires the initial placement of the 'anchor' numbers 12, 3, 6, and 9, which the authors claim employs executive functioning abilities. No evidence suggests this assertion or links it as an important impairment for dementia screening. This scoring system yielded an inter-rater reliability of .94. Sensitivity and specificity were not directly determined.

The Watson scoring system also bases its error detection on potentially faulty assumptions (Watson et al., 1993). Using only calculated sensitivity and specificity percentages (87% and 82% respectively) from 40 Alzheimer’s patients and 36 non-demented patients attending a geriatric assessment clinic or rehabilitation program, Watson and her colleagues decided to place a higher weight on number placement errors made in the fourth quadrant of the clock face. According to the Watson scoring system, any number error made in quadrants one through three is assigned a score of one. A number placement error made in the fourth quadrant is assigned a score of four. An individual is deemed cognitively intact if his/her final score falls in the zero to three range, whereas an individual scoring between four and seven is likely to have dementia. This scoring system has been found to have an inter-rater reliability of .90. Due to the small and limited characteristics of their sample size, the generalizability of the Watson scoring method is questionable.

Only two scoring systems, Freedman and Tuokko, have taken a systematic and normative approach for deciding which errors best discriminate people with dementia from cognitively intact controls (Freedman et al., 1994; Tuokko et al., 1995; Tuokko & O’Connell, 2006). In 1994, Freedman and his colleagues developed a 15-point free-drawn clock scoring system and a 13-point pre-drawn clock scoring system based on critical items obtained from their normative study containing 348 cognitively intact controls. For the free-drawn condition, individuals were asked to draw a clock and set the hands to a quarter to seven. Critical items for this scoring system included errors in contour shape, number placement, hand location, and center designation. Once empirically validated errors had been established, Freedman et al. examined these errors in Alzheimer’s patients and found that a cutoff score of 12 on free-drawn clocks produced a sensitivity of 85% and a specificity of 89%. Tuokko and her colleagues

(1995) also used normative data to develop a dual scoring system, which provides both quantitative and qualitative analyses of errors produced on the clock drawing task. This system draws on error types from seven different categories: omissions, perseverations, rotations, misplacement, distortions, substitutions, and additions. Details for the Tuokko scoring procedure as well as the other commonly used methods described above are summarized in Table 1.

Table 1. Common CDT scoring system criteria and psychometric properties.

	Method					
	Shulman	*Wolf-Klein*	*Royall*	*Watson*	*Freedman*	*Tuokko*
Free- or Pre-Drawn	Pre	Pre	Pre	Pre	Free	Pre
Time Setting	3:00	None	1:45	None	A quarter to 7	10 past 11
Scale/Points	1 – 6	1 – 10	15 pts	1 – 7	15 pts	31 pts
Cut Score	2	7	N/A	4	12	2
Sensitivity	86%	86.7%	N/A	87%	85%	92%
Specificity	72%	92.7%	N/A	82%	89%	86%
Inter-rater Reliability	.75	N/A	.94	.90	N/A	.90 - .95

V. A Closer Look at the Clock Test (Tuokko et al., 1995)

To date, the Clock Test (CT) developed by Tuokko and her colleagues, provides clinicians and researchers with one of the most comprehensive and valid scoring methodologies. The CT consists of three component measures: clock reading, clock setting, and clock drawing. These three tasks, when combined together, have been shown to increase the sensitivity and specificity (94% and 93% respectively) in detecting older adults with dementia compared to using each condition separately (Tuokko et al., 1992; Tuokko, Beattie, Crockett, & Horton, 1989). Normative data was established from 1,753 adults participating in the Canadian Study of Health and Aging (CSHA, Canadian Study of Health and Aging Working Group, 1994). These participants, who ranged in age from 65 to 100, were deemed healthy and cognitively intact, as they scored above the cutoff point of 77 on the Modified Mini-Mental Status Examination (3MS, Teng & Chui, 1987). The clinical sample consisted of 269 individuals with probable dementia, mostly of the Alzheimer type.

In 1995, Tuokko and her colleagues created the *Clock Test: Administration and Scoring Manual*, which details administration and scoring procedures, provides illustrative examples, and includes a QuikScore™ form for assistance with scoring the Clock-Drawing Test. According to this manual, the clock-drawing component of the test begins by presenting the participant with a pre-drawn circle 2.8-in (7-cm) in diameter. The examiner then tells the participant to "imagine that this [pre-drawn circle] is the face of a clock. Put the numbers on the face" (Tuokko et al., 1995, p. 5). When the participant has finished filling in the clock face, the examiner instructs him/her to "now put the hands on the clock to indicate 10 past 11" (Tuokko et al., 1995, p. 5). The phrase "10 past 11" can be repeated as many times as needed, but no other variation of "10 past 11" (e.g., "11:10", "10 minutes after 11") can be given.

As mentioned above, the clock-drawing component of the CT has seven main scoring categories: omissions, perseverations, rotations, misplacement, distortions, substitutions, and additions. Within each main category, subtypes of possible errors exist to ensure standard and unbiased scoring across all examiners. Although a ceiling for error totals does not exist, the manual reports an approximate maximum of 31 points (Tuokko et al., 1995). Table 2 provides a detailed summary of each error category, subcategory, and assigned point value.

Table 2. Summary of error types by category from the Clock Drawing component of the Clock Test.

Error Type	Description of Error	Maximum Point Value
Omissions		
Neglect	No marks made within 2 consecutive 5-min. intervals.	1
Number	Number(s) 1 through 12 is left out.	12 (1 per omitted number)
Hands	Stem of a clock hand, from the center (or near center) is left out.	2
Perseverations		
Repetition of numbers	Number(s) have been repeated.	1 per repetition
Hand perseveration	2+ hands originate from the clock center.	1 per extra hand
Sequence of number perseveration	Sequential numbers are written following 12 (e.g., 13, 14, 15).	1 per added number
Rotations		
Clock face rotations	Numbers written counterclockwise; numbers systematically displaced.	1
Number rotations	Number(s) rotations more than 45°.	1 per rotated number
Reversal rotations	Number(s) drawn in reverse.	1 per reversed number
Hand rotations	Time drawn in mirror image.	1
Misplacements		
Misplaced numbers	Number(s) not within ½ cm of its proper place.	12 (1 per misplaced number)
Sequence	Numbers not in sequence.	1
Misplaced hand(s)	Hand(s) not within ½ cm of its correct time setting.	2 (1 per misplaced hand)
Misplaced outside	A number is misplaced outside the clock face.	1 per misplaced number
Distortions		
Unique	Clock face is used to draw a recognizable figure other than a clock (e.g., human face, apple).	1
Horizontal/vertical number placement	Numbers are arranged in a horizontal or vertical manner.	1
Circularity	Number placement does not follow the contour of the circle.	1
Substitutions		
Letters/words	Letters or words take the place of numbers.	1 per each substitution
Scribble	1+ scribbles in place of numbers.	1 per each substitution
Words	Words indicating time are substituted for hands (e.g., "ten past eleven" is written).	2 (1 per hand replaced by a word)
Additional number	Numbers other than 1 - 12 are written and not in sequential order.	1 per each occurrence
Relative lengths	When 2 hands drawn, lengths differ by more than ½ cm.	1
Additions		
Irrelevant words	Irrelevant word(s)/phrase(s) are written.	1
Irrelevant scribble	Irrelevant scribble(s) or line(s) are written.	1
Irrelevant figures	1+ figures in addition to the clock face is drawn.	1

Psychometrically, the Clock Drawing component of the Clock Test (Tuokko et al., 1995) proves to be a strong and valid measure. Inter-rater reliabilities have been examined for each of the seven main error categories and were found to be: $r = 1.00$ for omissions, $r = .96$ for perseverations, $r = .98$ for rotations, $r = .93$ for misplacements, $r = 1.00$ for distortions, $r = .82$ for substitutions, and $r = .73$ for additions (Kurzman, 1992). These high r-values show that the Clock Drawing scoring system produces the same results with different raters. Sensitivity and specificity statistics have also been found to be very high for the CDT. Using a cutoff score of two error points, the clock-drawing test was found to have a sensitivity of 86% and a specificity of 92% (Tuokko et al., 1995). In 2006, Tuokko and O'Connell provided a review of six quantified methods for assessment of qualitative clock drawing and found that of all the clock-drawing scoring systems, the Tuokko scoring system had the highest sensitivity and specificity. Furthermore, this scoring system is the only procedure to provide age-based cutoff scores to maximize sensitivity and specificity within each age group and to account for cognitive decline associated with normal and healthy aging (Tuokko & O'Connell, 2006).

VI. A Modified Scoring System for the Clock Test

The qualitative and quantitative Tuokko scoring system has shown to be highly predictive in detecting dementia (O'Rourke et al., 1997). It is one of the most detailed and comprehensive systems available today, with seven broad error categories, 24 error sub-categories, and an approximate maximum error score of 31 (Tuokko et al., 1995). This system, however, has been criticized for its complexity and impracticality in the clinical setting (Lin, Wang, Chen, Chiu, Kuo, Chuang, & Liu, 2003). An ideal clinical screening tool should: (a) be free from educational, linguistic, and cultural biases; (b) require little or no equipment; (c) classify individuals with high sensitivity and specificity; and (d) be quick and easy to administer and score (Borson et al., 2000). The clock-drawing component of the CT created by Tuokko et al. (1995) satisfy all these conditions, but one can argue the scoring system can be simplified to reduce scoring time.

In 2008, Lessig and her colleagues at the University of Washington sought to simplify the CDT scoring system by identifying the optimal number of critical errors needed to detect dementia by combining clock errors from three different scoring systems (Mendez et al., 1992; Tuokko et al., 1992; Shulman et al., 1993). Twenty-four separate errors were identified and applied to 536 clock-drawings from older adults enrolled in the University of Washington Memory Disorders Clinic and the Alzheimer's Disease Research Center Satellite (154 classified as healthy controls, 101 classified with subsyndromal cognitive impairment, 281 classified with dementia). This sample was ethnolinguistically and educationally diverse. An algorithm created by a stepwise logistic regression was used to reduce and simplify the 24 pre-selected errors down to six critical errors: wrong time, no hands, missing numbers, number substitution, repetition, and refusal to draw a clock. Sensitivity and specificity for dementia detection was found to be 71% and 88%, respectively. The results from this study show that an equally effective yet simplified clock-drawing scoring system can be derived from a more complex classification structure when using a clinic-based sample.

VII. Research Question

This Master's thesis replicates Lessig et al.'s (2008) study on the modification and simplification of the clock-drawing scoring system and examines its generalizability to a population-based Canadian subsample. It is hypothesized that the same clock-drawing errors found by Lessig and her colleagues will predict dementia with similar sensitivity and specificity when applied to a subsample of the Canadian population. No study has yet to generalize the newly developed Lessig scoring system to a population-based sample.

Chapter 2:
Method

Participants

Clock-drawings from 356 participants in the Canadian Study of Health and Aging (CSHA, CSHA Working Group, 1994) were analyzed. The CSHA began its intensive data collection in 1991 by conducting screening interviews with over 9,008 community-dwellers aged 65 years and older. The study sought to identify individuals with cognitive impairment by conducting clinical examinations (medical assessments including blood tests and physical measurements, interviews with caregivers, and neuropsychological assessments) for three groups: (a) those who screened positive for cognitive impairment; (b) a subsample of those who screened negative for cognitive impairment; and (c) 1,255 residents of care facilities. In total, the CSHA collected comprehensive neuropsychological data from approximately 1,879 participants ranging in age from 65 to 100 (Tuokko, Kristjansson & Miller, 1995). Of this sample, 493 individuals from British Columbia completed the clock-drawing component of the Clock Test (Tuokko et al., 1995). For this study, 356 individuals of the 493 who completed the CDT were identified - 276 classified as cognitively intact and 80 exhibiting any form of dementia as defined by the DSM-III-R criteria (e.g., possible Alzheimer's disease, probable Alzheimer's disease, vascular dementia, unclassified dementia). The remaining 137 individuals were classified with cognitive impairment no dementia (CIND)/mild cognitive impairment (MCI) and were excluded from the analyses. Previous research has shown that the CDT possesses very poor discriminative power when used to detect CIND/MCI (Ehreke, Luppa, Konig, & Riedel-Heller, 2010; Lee, Kim, Hong, Lee, Oh, & Cheong, 2008). Consequently, this demographic group was excluded from the study so as not to cloud the results. Collectively, participants ranged in age from 65 to 100 years (*M* = 79.28, *SD* = 6.98). Their education level ranged from 0 to 25 years (*M* = 9.18, *SD* = 4.01). Both groups (no cognitive loss and dementia) were fairly evenly split between males and females (53.5% females in the no cognitive loss group and 51.3% females in the dementia group). Table 3 presents a summary of the demographic characteristics of each group, individuals with no cognitive loss and those diagnosed with dementia, according to CSHA criteria.

Table 3. Age, education, gender, and race for individuals classified with no cognitive loss and dementia*.

	No Cognitive Loss (N = 276)	**Dementia** (N = 80)
Age	*M* = 77.74, *SD* = 6.79	*M* = 79.33, *SD* = 8.00
65 – 75 years	N = 60	N = 13
76 – 85 years	N = 69	N = 24
86 – 95 years	N = 23	N = 7
96 – 100 years	N = 0	N = 2
Education	*M* = 9.67, *SD* = 3.96	*M* = 9.10, *SD* = 4.16
0 – 5 years	N = 37	N = 14
6 – 8 years	N = 87	N = 25
9 – 12 years	N = 99	N = 25
13 – 16 years	N = 36	N = 10
16 + years	N = 17	N = 4
Percent Female	53.3%	51.3%
Dementia Type		
Probable Alzheimer's		N = 29
Possible Alzheimer's		N = 31
Vascular Dementia		N = 14
Unclassified Dementia		N = 4

*According to criteria from the *Diagnostic and Statistical Manual of Mental Disorders*, 3rd Ed. (American Psychiatric Association, 1987).

Procedure

During administration by CSHA examiners, participants were presented with a pre-drawn circle (2.8 in/7 cm in diameter) and told to: "Imagine this is the face of a clock. Put the numbers on the clock". Once the participant completed this initial step, the examiner then instructed the participant to: "Put the hands on the clock to indicate 10 past

11". A trained psychometrist from the Vancouver CSHA team first scored the clock-drawings using the Tuokko (1995) method. The trained psychometrist also scored the clock-drawings using the Shulman, Wolf-Klein, and Watson procedures in order to conduct an inter-scoring system comparison (Shulman et al., 1986; Shulman et al., 1993; Watson et al. 1993; Wolf-Klein et al., 1989; see below).

Data Analysis

Correlation analyses to determine co-variates: Initial analyses (correlations) examined the relationship between age, years of education, and dementia status to determine if these variables should be treated as co-variates. Previous research has demonstrated that education level can alter performance on the clock-drawing test (Ainslie & Murden, 1993) thus warranting closer examination of this variable in this study. Although age has traditionally been used as a co-variate in statistical analyses, emerging research suggests that correcting for age can, in fact, diminish a test's discriminative ability (Sliwinski, Lipton, Buschke, & Wasylyshyn, 2003). According to Sliwinski et al. (2003), age is an important contributor to the development of dementia and, as a result, is a necessary component to examine when using diagnostic norms.

Identifying significant errors: Lessig et al.'s (2008) 24 pre-selected clock-drawing errors were then matched with the errors from the Tuokko scoring system (Table 4). Nine errors could not be matched due to differences in administration and scoring procedures between the two tests. For example, Lessig et al.'s "clock-like figure", "refusal", and "incorrect shape" errors could not be matched to Tuokko et al.'s errors because the latter scoring system uses a pre-drawn clock thus making these three specific errors impossible to score. Several errors from the Tuokko scoring system were more detailed than errors

Table 4. Initial clock-drawing errors used in statistical analyses.

Lessig et al. (2008) Error Type	Tuokko et al. (1995) Error Type	Description
Incorrect time	Hand rotation; misplaced hands	Hands set incorrectly.
No hands	Hand/s omitted	No hands drawn.
Missing numbers	Number/s omitted	One or more numbers missing; includes tick marks in place of numbers.
Repeated numbers	Repetition of number/s	Same numbers appear more than once.
Substitution	Letters/words, Scribble, Words Substituted	Symbols/marks used in place of numbers; time written out rather than shown by hands.
Number orientation	Clock face rotation	Numbers counterclockwise.
Number order	Sequence misplaced	Incorrect number sequence.
Number/s outside of circle	Misplaced outside	Number/s placed outside the contour.
Clock-like figure	No match	Image, figure, symbols, or characters drawn do not resemble a clock.
Number spacing (major)	No match	Gross error of spacing.
Refusal	No match	Test form blank, or examiner recorded refusal to start or finish a partial attempt.
Incorrect shape	No match	Clock face is not round or not a closed figure.
Extra marks	Irrelevant words, scribble, figures	Uninterruptible extraneous marks/symbols/figures anywhere on the clock face.
Number perseveration	Number sequence perseveration	Any number past 12 is present.
Hands in center	No match	Hands not connected; do not radiate from center.
Number of hands	Hand perseveration; Hand/s omitted	Only one or more than two hands are drawn.
Hand length	Relative lengths	Both hands same length; hour hand not shorter than minute hand; hands extend past contour.
Number rotation	Number rotation/s	Numbers backward/upside down; rotated.
Face geometry	No match	Clock gestalt present, but internal geometry is rotated.
Number distance	Circularity	Numbers not equidistant from center and edge of face.
Number spacing (all)	Misplaced numbers	Any uneven spacing/gaps
Second try	No match	Participant started over
Military time	No match	Numbers indicating an entire 24-hour period (1, 2…23, 24).
Anchoring	No match	Only anchoring numbers present (12, 3, 6, 9).

identified by Lessig and her colleagues (e.g., letters/words, scribble, and words substitution; irrelevant words, scribbles, and figures; hand rotation and misplaced hands; hand perseveration and hand omission) and, as a result, these errors were combined to fit Lessig et al.'s error description. Fifteen final clock errors were obtained to represent the overlap between the Lessig and Tuokko scoring systems. Of the six critical errors Lessig and her colleagues (2008) found in their original study, only five were used in this study, as the sixth error (refusal to draw a clock) could not be identified. All error types from the Tuokko scoring system were dichotomized (0 = no error made, 1 = 1 of more error/s made) in order to match Lessig's binary scoring procedure.

An initial sequential logistic regression analysis was then conducted to predict cognitive status (cognitively intact vs. dementia) using SPSS. The 15 individual error types were treated as the independent variables and dementia status was used as the dependent variable. Lessig et al.'s (2008) five identified critical errors were entered in block 1 (incorrect time, no hands, missing numbers, repeated numbers, and substitution), followed by all remaining ten errors in block 2 (number orientation, number order, number/s outside the circle, extra marks, number perseveration, number of hands, hand length, number rotation, number distance, and number spacing).

Reduced errors: Once the errors that significantly discriminated between the cognitively intact participants and those with dementia were isolated, a second direct logistic regression analysis was conducted to assess for optimal specificity and sensitivity, with dementia status (0 = normal, 1 = demented) as the dependent variable and significant errors as predictors. All significant errors were entered in block 1. These critical errors were then combined to create a single composite score, which was then tested for specificity and sensitivity for discriminating cognitive intact participants from those with dementia (see next section).

Determining sensitivity and specificity: Receiver operating characteristic (ROC) curves, a decision-making tool used to evaluate the sensitivity and specificity at different cutoff points, was used to determine the particular sensitivity and specificity associated with the newly derived scoring system, which hereafter will be referred to as the Jouk scoring system. An ROC curve plots sensitivity (true positive rate) versus 1 – specificity (false positive rate). More specifically, it represents the probability that two scores, one obtained from disease-positive individuals and one from disease-negative individuals, will be accurately identified (Hanley & McNeil, 1982). As a result, for each cutoff score a diagnostic test will have its own associated sensitivity and specificity, which represent a trade-off as the cut scores change. When criteria parameters are loosened, a test's sensitivity will increase, but its specificity will decrease (Rosman & Kortsen, 2007). Typically, researchers developing clinical screening tools favor tests with high sensitivity (Essex-Sorlie, 1995) since it is desirable to avoid incorrectly identifying disease-positive individuals (false positives) in order to diagnose the disease in its early stages. However, with a disease like dementia where no cure exists, a diagnostic screening tool should also be highly specific (Essex-Sorlie, 1995) in order to avoid causing undue psychological harm by giving a diagnosis of dementia to a healthy individual. As a result, an excellent scoring system for the CDT is both highly sensitive and specific.

In addition to examining sensitivity and specificity to understand a test's discriminative ability, calculating the area under the ROC curve can also provide insight into a screening tool's utility. An area closer to 1 represents a diagnostic test with a high ability to correctly identify disease-positive individuals from disease-negative individuals (Bewick et al., 2004). Although the area under the ROC curve can provide information about a test's discriminative abilities, clinicians most often examine sensitivity and specificity since these two measures give a more detailed picture of a test's diagnostic capabilities.

Comparison with other clock-drawing scoring systems: The sensitivity and specificity of the Jouk scoring system was compared with the Lessig scoring system. To provide a more direct evaluation, the Jouk scoring system's sensitivity, specificity, and area under the ROC curve were also compared to the Shulman, Watson, Wolf-Klein, and Tuokko scoring system using CSHA data. For each scoring procedure error totals were converted to dichotomous variables (0, 1) according to their published cutoff scores, with "0" representing a healthy, cognitively intact individual and "1" representing a person with suspected dementia. The Shulman system scores the CDT on a scale from 1 to 6, with lower scores indicating less or no impairment (Shulman et al., 1993). A cutoff score of great than 2 suggests cognitive impairment (Shulman et al., 1986). The Watson method scores clock errors from 0 (no errors/least impairment) to 7 (most errors/most cognitive impairment), with a cutoff score of 4 or greater reflecting a person with dementia (Watson et al., 1993). The Wolf-Klein scoring system ranges from 1 to 10, with higher scores depicting less impairment and a cutoff score of 8 or greater suggesting a healthy individual with no cognitive impairment (Wolf-Klein et al., 1989). The Tuokko scoring method produces an error score ranging from 0 to approximately 31 and utilizes a cutoff score of 2 and below to indicate no cognitive impairment (Tuokko et al., 1995).

After dichotomizing all the error totals using the appropriate cutoff scores, the data were consistent with the Lessig and Jouk scoring systems and could then be easily compared.

Chapter 3:

Results

Correlation Analyses

Initial correlations examined the relationship between dementia status, age, and years of education to determine if these variables should be treated as co-variates. A Pearson correlation coefficient was computed to assess the relationship between age and dementia status. There was no correlation between the two variables, r = .095, N = 198, p = ns. A second Pearson correlation coefficient was computed to assess the relationship between years of education and dementia status. Once again there was no correlation between these two variables, r = -.059, N = 354, p = ns. Since both age and years of education were not significantly correlated with dementia status, they were not entered into the regression model as co-variates.

Significant and Non-significant Errors

A sequential logistic regression was conducted to identify which of Lessig's 15 original errors significantly predicted dementia status. Table 5 shows the regression coefficients, the corresponding standard errors, the odds ratios and their associated 95% confidence intervals, as well as the chi-square tests and Nagelkerke R^2. In block 1 of the analysis, Lessig's five identified critical errors (incorrect time, no hands, missing numbers, repeated numbers, and substitution) were entered into the regression equation and significantly predicted dementia status, χ^2 (5, N = 356) = 74.62, p = .00, Nagelkerke R^2 = .29. Examination of the odds-ratios (see Table 5) showed that all but one of Lessig's original errors (number substitution) significantly contributed to predicting dementia status. When all the remaining ten errors (number orientation, number order, number/s

Table 5. Summary of initial logistic regression analysis for all 15 clock-drawing errors predicting dementia status (N = 356).

	Block 1				Block 2			
Error Type	*B*	*SE B*	*Odds-Ratio*	*CI of Odds-Ratio*	*B*	*SE B*	*Odds-Ratio*	*CI of Odds-Ratio*
Incorrect Time	.69	.34	2.00**	1.03 – 3.90	.48	.38	1.61	.77 – 3.34
No Hands	1.00	.32	2.71**	1.44 – 5.08	-.46	.79	.63	.13 – 3.00
Missing Numbers	1.36	.41	4.92**	1.74 – 8.76	1.37	.46	3.95**	1.60 – 9.81
Repeated Numbers	1.60	.39	3.90**	2.29 – 10.57	1.00	.46	2.73**	1.11- 6.72
Number Substitution	.69	.49	1.98	.76 – 5.19	.76	.52	2.13	.77 – 5.92
Number Orientation					1.23	.56	3.41**	1.14 – 10.22
Number Order					-.10	1.06	.90	.11 – 7.25
Number/s outside circle					1.13	.72	3.09	.77 – 12.61
Extra Marks					.70	.35	2.00*	1.01 – 4.01
Number perseveration					1.33	1.06	3.80	.48 – 30.01
Number of hands					1.27	.77	3.55	.78 – 16.09
Hand Length					.23	.41	1.26	.56 – 2.82
Number Rotation					.36	.38	1.43	.68 – 2.99
Number Distance					1.61	.52	4.99**	1.82 – 13.68
Number Spacing					.09	.37	1.09	.53 – 2.25
χ^2		74.62				34.71		
Nagelkerke R^2		.29				.40		

SE = Standard Error; *CI* = Confidence Interval at 95%
* $p < .05$; ** $p < .01$

outside the circle, extra marks, number perseveration, number of hands, hand length, number rotation, number distance, and number spacing) were entered in block 2 of the regression equation, they showed a significant improvement in the prediction of dementia status, χ^2 (10, N = 356) = 34.71, *p* = .00, Nagelkerke R^2 = .40.

Examination of odds-ratio showed that five errors (missing numbers, repeated numbers, number orientation, extra marks, and number distance) significantly discriminated cognitively intact participants from participants with dementia (see Table 5). Two of Lessig's originally identified errors, incorrect time and no hands, no longer proved to be significant errors using this data. However, three other errors (number orientation, extra marks, and number distance) were added to the set of significant/critical errors. Closer scrutiny of the odds-ratios showed that those exhibiting missing number errors (0 versus 1) are at approximately a four-times greater risk of having dementia compared to those who did not make this error. Individuals who repeat numbers when drawing the clock versus those who do not make this error are at a 2.73-increased risk for dementia. Those who display number orientation errors are at a 3.41-increased risk of having dementia compared to individuals who correctly oriented the numbers on the clock face. Individuals who make extra marks on the clock face (versus those who do not make this error) are at twice the risk of being demented. Finally, those exhibiting number distance errors are almost at a five-times increased risk of having dementia compared to individuals who do not make this error.

Reduced Errors

A second direct logistic regression analysis was conducted to evaluate the newly identified errors' impact on dementia status. Table 6 shows the regression coefficients, the corresponding standard errors, the odds ratios and their associated 95% confidence intervals, as well as the chi-square test and Nagelkerke R^2. The five significant errors identified from the initial logistic regression (missing numbers, repeated numbers, number orientation, extra marks, and number distance) were entered in block 1 of the regression equation and significantly predicted dementia status χ^2 (5, N = 356) = 92.69, p = .00, Nagelkerke R^2 = .35. Examination of the odds-ratios verified that all five errors significantly discriminated cognitively intact individuals from participants with dementia with greater predictive power than the previous analyses (see Table 6). Closer inspection of the odds-ratios showed that those exhibiting missing number errors (versus those who correctly included the numbers "1" through "12") are at approximately a five-times increased risk of being demented. Those demonstrating repeated number errors are at a 4.81-increased risk for having dementia compared to individuals who do not make this error. Individuals who make number orientation errors (versus those who do not make this error) are at a 4.55-times increased risk of having dementia. Those who place extra marks on the clock face are at a 2.64-increased risk of being demented compared to those who do not make this error. Finally, individuals displaying number distance errors (compared to those who do not make this error) are at an approximate five-times increased risk of having dementia.

These five critical errors (missing numbers, repeated numbers, number orientation, extra marks, and number distance) were then combined to create a composite score, which was used to calculate the sensitivity and specificity of this reduced scoring system using ROC curves for each of the six possible cutoff points (see next section).

Table 6. Summary of the logistic regression analysis for five significant clock-drawing errors predicting dementia status (N = 356).

			Block 1	
Error Type	*B*	*SE B*	*Odds-Ratio*	*CI of Odds-Ratio*
Missing numbers	1.64	.40	5.13**	2.36 – 11.18
Repeated numbers	1.57	.41	4.81**	2.15 – 10.73
Number orientation	1.52	.52	4.55**	1.66 – 12.50
Extra marks	.97	.32	2.64**	1.40 – 4.97
Number distance	1.64	.49	5.16**	1.99 – 13.40
χ^2		92.69		
Nagelkerke R^2		.35		

SE = Standard Error; *CI* = Confidence Interval at 95%
* $p < .05$; ** $p < .01$

Sensitivity and Specificity of the Reduced Scoring System

An ROC curve was plotted (sensitivity versus 1 – specificity for all possible cutoff scores) for the reduced scoring system to evaluate its overall sensitivity and specificity for detecting dementia. The area under the ROC curve was .80 (*SE* = .03), indicating a moderate trade-off between the false negative (sensitivity) and false positive (specificity) rates for the six possible cutoff scores. After examining the sensitivity and specificities for each of the five possible cutoff points, a cut score of 1 produced the most balanced sensitivity and specificity values at 81% and 68%, respectively (see Table 7).

Table 7. Sensitivities and specificities associated with each possible cutoff score for the new scoring system.

Cutoff Score	**Sensitivity (%)**	**Specificity (%)**
0	100	0
1	81	68
2	44	94
3	20	99
4	7.5	100
5	1.3	100

Comparison to the Lessig Scoring System

The results from the ROC curve analysis discussed above were compared with the results published by Lessig and her colleagues in 2008. The Jouk scoring system's sensitivity of 81% was 10% higher than the sensitivity of the Lessig scoring system, which was 71%. The Lessig scoring system, however, had a higher specificity (88%) than the Jouk scoring system (68%) (see Table 9).

Comparison to Other Commonly Used Scoring Systems

Using CSHA data, ROC curves were drawn for the Shulman, Tuokko, Watson, and Wolf-Klein scoring systems to evaluate their respective area under the curve, sensitivity and specificity (see Figure 7). Table 8 shows the resulting area under the ROC curve for each scoring method. The Jouk scoring system had the largest area under the curve (.80, *SE* = .03), followed by the Wolf-Klein system (.72, *SE* = .03). The Shulman and Tuokko scoring procedures' area under the curve fell in the middle of all the scoring systems (both .70, *SE* = .03). The Watson scoring method had the smallest area under the curve (.63, *SE* = .04).

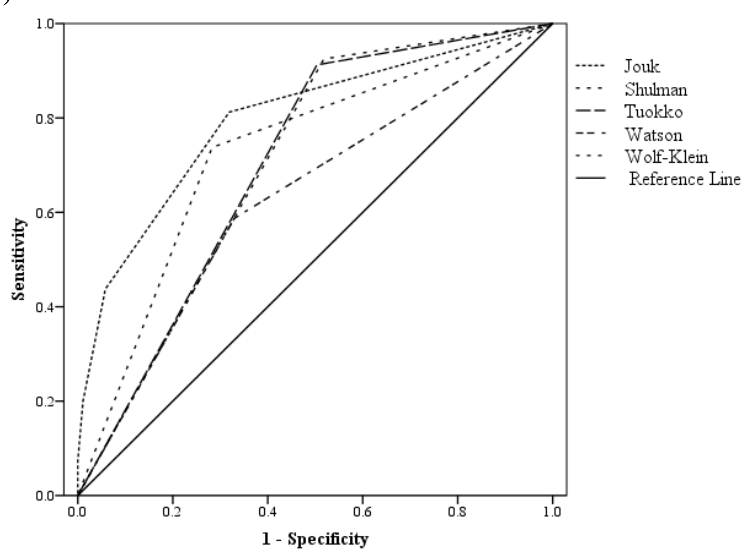


Figure 7. ROC curves for the five clock scoring procedures derived from CSHA data.

Table 8. Area under the ROC curve and associated standard error for each of the five clock scoring procedures.

Scoring Procedure	Area Under the Curve	Standard Error
Jouk	.80	.03
Shulman	.70	.03
Tuokko	.70	.03
Watson	.63	.04
Wolf-Klein	.72	.03

The sensitivities and specificities of the five scoring systems derived from CSHA data are presented in Table 9. The Shulman and Tuokko scoring systems had the highest sensitivities at 93% and 91%, respectively. The Jouk and Wolf-Klein scoring procedures fell in the middle at 81% and 74%, respectively. The Watson method had the lowest sensitivity at 59%, performing just above chance level for correctly identifying individuals with dementia. With regard to specificity, the Wolf-Klein scoring procedure had the highest specificity at 72%, followed by the Jouk scoring method at 68%. The Watson scoring system's specificity closely trailed the Jouk method at 67%. Both the Shulman and Tuokko procedures had the lowest specificities at 48% and 46%, respectively.

Table 9 also presents the sensitivities and specificities published by the original authors to provide another method of comparison.

Table 9. Sensitivity and specificity calculated at the published cutoff scores for the five clock scoring procedures derived from CSHA data and also published by the original authors.

Scoring Procedure	Cutoff Score	CSHA Sensitivity (%)	CSHA Specificity (%)	Published Sensitivity (%)	Published Specificity (%)
Jouk	1	81	68	---	---
Lessig	---	---	---	71	88
Shulman	2	93	48	86	72
Tuokko	2	91	46	86	92
Watson	3	59	67	87	82
Wolf-Klein	7	74	72	86.7	92.7

Chapter 4:

Discussion

Fairly consistent with the hypothesis, the Jouk scoring method for the CDT was found to have a similar sensitivity and specificity compared to the Lessig scoring procedure using one less critical error. The Jouk scoring system used five critical errors (missing numbers, repeated numbers, number orientation, extra marks, and number distance) with a cutoff score of 1 to produce a sensitivity of 81% and a specificity of 68%. The Lessig scoring system had a sensitivity of 71% and a specificity of 88%. Even though the specificity was better for Lessig scoring procedure, the Jouk scoring system's sensitivity surpassed the Lessig system. With regard to the other traditionally used scoring methods, the Jouk procedure had one of the most balanced sensitivities/specificities when using a population-based sample (81% and 68% respectively). The Wolf-Klein scoring system also produced a balanced sensitivity and specificity when using the CSHA data (74% and 72% respectively). Although the Shulman and Tuokko methods had two of the highest sensitivities when using population-based data (93% and 91% respectively), their specificities using this same data were very low (48% and 46% respectively). However, when the scoring systems derived from CSHA data in this study were compared with the original studies conducted by the authors who developed the scoring systems, the sensitivities and specificities were generally much higher and more balanced.

The differences in these results could have stemmed from inherent qualities of the two samples. Many of the originally published scoring systems, including the Lessig scoring method, were derived from samples obtained from dementia-specific clinics. Individuals attending these clinics have already been "pre-selected" for their cognitive impairments, thus creating homogeneity among the sample. The homogenous characteristics of this sample increase the sensitivity of the CDT. Conversely, individuals examined from a representative subset of the general population, like those tested in CSHA and used in this study, comprise a more heterogeneous sample, thereby decreasing the sensitivity of the CDT. Pinto and Peters (2009) also found similar results when conducting a general review of the administration and utility of the CDT. They also attributed the discrepancy in sensitivity and specificity among the original published results and subsequent comparative studies to differences in the populations and control groups used for each study.

In general, this study showed that the Jouk method provided one of the better procedures for scoring the CDT if using a population-based sample. Out of all the scoring systems examined, the Jouk scoring procedure had the highest area under the curve. It has been suggested that the area under the ROC curve provides a more unbiased method for determining the diagnostic utility of a screening test as it shows the relationship between sensitivity and specificity (Pinto & Peters, 2009). Moreover, when sensitivity and specificity were individually compared among the different scoring systems examined in this study, the Jouk and Wolf-Klein procedures most evenly balanced sensitivity and specificity. According to Essex-Sorlie (1995), a screening tool for dementia should be both highly sensitive and specific. It can be argued that if the CDT is to be used in dementia clinics or other clinical settings the test should have a higher specificity so as to not falsely identify healthy individuals as having an incurable disease. Screening tests with higher sensitivities may therefore be best suited for research settings, where it is important to accurately classify participants in disease-positive/experimental and disease-negative/control groups. As a result, the Jouk scoring system for the CDT may prove to be more useful in a research setting where a population-based sample is used. Further research, however, is needed to confirm this hypothesis.

When individual error types from the reduced Jouk and Lessig scoring systems were compared, two common errors were found – missing numbers and repeated numbers. These errors, as well as the other three errors detected from the Jouk scoring method (number orientation, extra marks, and number distance), are fairly obvious errors to detect from a single glance. This proves to be a strength of the Jouk scoring system since identifying these errors can be done quickly, easily, and without much deliberation.

Additionally, it was noted that the majority of the errors comprising the Jouk scoring system (missing numbers, number orientation, and number distance) represented a spatial planning deficit. This result supports other findings that spatial planning errors strongly predict dementia status (Blair, Kertesz, McGonagle, Davidson, & Bodi, 2006; Lee et al., 2009). Lee and his colleagues (2009) conducted a quantitative and qualitative analysis to define the characteristics of the error types made on the CDT by patients with Parkinson's disease, subcortical vascular dementia, and Alzheimer's disease. They found that spatial planning errors most sensitively detected all forms of dementia, even in the early stages. Blair and his colleagues (2006) conducted a similar study, although they limited their sample to include individuals with Alzheimer's disease and fronto-temporal dementia. They observed that patients with Alzheimer's disease made more spatial planning errors compared with patients with fronto-temporal dementia. Collectively, these results suggest that spatial planning errors represent a core feature in the detection of dementia, although further research is needed to disentangle the discrepant findings concerning the different types of dementia.

This study included several limitations that should be considered when examining the results. First and foremost, due to restrictions in the original data, not all 24 errors identified by Lessig and her colleagues were used in the initial regression analysis. One error in particular, refusal to draw a clock, was found to be a critical error in the Lessig scoring system, but could not be included in the initial regression. Quite possibly, if using a free-drawn clock, this error could have also been identified as a critical error in the Jouk system thereby increasing the procedure's sensitivity and specificity. In addition, although the sample of participants used in this study already comprised a more representative subset of the general population compared to the clinic-based participants used in the Lessig study, this sample was not ethnically diverse nor as representative as might be desirable. Even though the CDT has been shown to be free of cultural bias (Leung, Lui, & Lam, 2005), a stronger case for the reduced Jouk system could have been made if the results also demonstrated these findings.

Future studies should consider and rectify the limitations described above as well as examine larger topics that will ultimately strengthen the clinical utility of the Jouk scoring system. Previous studies have paired the CDT with other cognitive assessment tools like the MMSE, verbal fluency tests, or even similar clock tasks like clock-reading and clock-setting to increase sensitivity and specificity (Borson, Scanlan, Chen, & Ganguli, 2003; Schramm, Berger, Muller, Kratzsch, Peters, & Frolich, 2002; Tuokko et al. 1995). The Jouk scoring method could also be combined with another screening measure to increase its discriminative ability. Additionally, the same data reduction methods used in this study should be applied to various types of dementia (e.g., vascular dementia, fronto-temporal dementia, Alzheimer's disease) in order to examine if unique error types are associated with different forms of dementia. Lastly, inter-rater reliabilities for the Jouk scoring system should be calculated to determine what type of scorer is needed (novice research assistant versus skilled psychometrist). Given that the Jouk scoring procedure has a cutoff score of 1 and only 5 errors to score (most of which are spatial planning errors) this system may be easier to score by individuals of all training levels compared to other, more complex scoring methods. When using a complex scoring system, it has been shown that more experienced scorers who have had more time training and practice developing their rating skills should be used (Nishiwaki, Breeze, Smeeth, Bulpitt, Peters, & Fletcher, 2004). As a result, the simplicity of the Jouk scoring system may not only save significant amounts time that would have otherwise been spent identifying more errors, but also money since less training is needed to learn how to apply the Jouk method.

Chapter 5:
Conclusion

As the prevalence of dementia will reach unprecedented levels in the coming years, the need for quick, reliable, and accurate diagnostic tools has never been more important. This thesis project sought to simplify the scoring system for the Clock Drawing Test with the aim of creating a more streamlined, but sensitive dementia-screening test. Additionally, this project broadened the CDT scoring system to a representative sample to mimic conditions a general clinician will encounter when working among a wide-ranging population and not at a dementia/memory clinic. The findings support the current literature that the CDT scoring system can indeed be reduced while still maintaining fairly high sensitivity. The lower specificity found in this study, however, suggests that the Jouk scoring system may be better suited for use in research settings, where it is more important to classify individuals in their condition group by avoiding false-negatives.

In general, the findings underscore the importance of evaluating all aspects of a scoring system's performance, as certain qualities of the system may make it be more or less appropriate for the task at hand (e.g., providing a clinical diagnosis versus answering a research question). Since no screening measure can be perfectly valid, it is also important to use other assessment tests, clinical interviews, observations, and caregiver reports before arriving at a diagnosis. As this thesis demonstrated, a simplified scoring system for the CDT can be used as a dementia screening measure, although caution must be exercised so as not to solely base a diagnostic decision on the results from this screening test alone.

References

Ainslie, N. K. & Murden, R. A. (1993). Effect of education on the clock-drawing dementia screen in non-demented elderly persons. *Journal of the American Geriatrics Society, 41,* 249 – 252.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*, (4th ed., text revision). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders*, (3rd ed., revision). Washington, DC: American Psychiatric Association.

Baddeley, A. & Della Sala, S. (1998). Working memory and executive control. In A. C. Roberts, T. W. Robbins, L. Weiskrantz (Eds.), *The Prefrontal Cortex: Executive and Cognitive Functions* (pp. 9 – 21). New York, NY: Oxford University Press.

Battersby, W. S., Bender, M. B., Pollack, M., & Kahn, R. L. (1956). Unilateral "spatial agnosia" ("inattention") in patients with cortical lesions. *Brain, 79,* 68 – 93.

Bauer, R. M., Grande, L., & Valenstein, E. (2003). Amnesic disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical Neuropsychology, Fourth Edition* (pp. 495 – 473). New York, NY: Oxford University Press.

Bewick, V., Cheek, L., Ball, J. (2004). Statistics review 13: Receiver operating characteristic curves. *Critical Care, 8(6),* 508 – 512.

Blair, M., Kertesz, A., McGonagle, P., Davidson, W., & Bodi, N. (2006). Quantitative and qualitative analyses of clock drawing in frontotemporal dementia and Alzheimer's disease. *Journal of the International Neuropsychological Society, 12(2),* 159 – 165.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The mini-cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15,* 1021 – 1027.

Borson, S., Scanlan, J. M., Chen, P., & Ganhuli, M. (2003). The Mini-Cog as a screen for dementia: Validation in a population-based sample. *Journal of the American Geriatrics Society, 51,* 1451 – 1454.

Cahn-Weiner, D. A., Sullivan, E. V., Shear, P. K., Fama, R., Lim, K. O., Yesavage, J. A., Tinklenberg, J. R., & Pfefferbaum, A. (1999). Brain structural and cognitive correlates of clock drawing performance in Alzheimer's disease. *Journal of the International Neuropsychological Society, 5,* 502 – 209.

CSHA Working Group. (1994). Canadian Study of Health and Aging: Study methods and prevalence of dementia. *Canadian Medical Association Journal, 150,* 899 – 913.

Critchley, M. (1966). *The parietal lobes* (2nd ed.). New York, NY: Hafner Press.

Di Pellegrino, G. (1995). Clock-drawing in a case of left visuo-spatial neglect: A deficit of disengagement. *Neuropsychologia, 33(3),* 353 – 358.

Ehreke, L., Luppa, M., Konig, H. H., & Riedel-Heller, S. G. (2010). Is the Clock Drawing Test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. *International Psychogeriatrics, 21(1),* 56 – 63.

Essex-Sorlie, D. (1995). *Medical biostatistics and epidemiology*. Norwalk, Connecticut: Appleton and Lange.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189 – 198.

Freedman, M., Leach, L., Kaplan, E., Winocur, G., Shulman, K. I., & Delis, D. C. (1994). *Clock drawing: A neuropsychological analysis*. New York, NY: Oxford University Press.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive neuroscience: The biology of the mind* (2nd ed.). New York, NY: W. W. Norton & Company, Inc.

Geschwind, N. (1965) Disconnection syndromes in animal and man. *Brain, 88,* 237 – 294.

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia, PA: Lea & Febiger.

Hanley, J. A. & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*, 29 – 36.

Head, H. (1926). Aphasia and kindred disorders of speech. New York: Macmillan.

Jackson, J. H. (1874). On the nature of the duality of the brain. *Medical Press and Circular, 17(19),* (reprinted in *Brain*, 1915, *38,* 80 – 103).

Kim, Y-S., Lee, K-M., Choi, B. H., Sohn, E-H., & Lee, A. Y. (2009). Relationship between the Clock Drawing Test (CDT) and structural changes of brain in dementia. *Archives of Gerontology and Geriatrics, 48(2)*, 218 – 221.

Kingstone, A. & Gazzaniga, M. S. (1995). Subcortical transfer of higher order information: More illusory than real? *Neuropsychology, 9 (3), 3*21 – 328.

Kleist, K. (1912). Der gang und der gegenwartiage stand der apraxieforshung. *Neurologie und Psychiatrie, 1,* 342 – 452.

Kurzman, D. (1992). *The construct validity of the Clock Test in normal and demented adults*. Master's thesis, University of Victoria, BC.

Lam, L. C. W., Chiu, H. F. K., Ng, K. O., Chan, C., Chan, W. F., Li, S. W., & Wong, M. (1998). Clock-face drawing, reading, and setting tests in the screening of dementia in Chinese elderly adults. *Journal of Gerontology, 53B,* P353 – P357.

Lee, K. S., Kim, E. A., Hong, C. H., Lee, D-W., Oh, B. H., & Cheong, H-K. (2008). Clock Drawing Test in mild cognitive impairment: Quantitative analysis of four scoring methods and qualitative analysis. *Dementia and Geriatric Cognitive Disorders, 26,* 483 – 489.

Lessig, M. C., Scanlan, J. M., Nazemi, H., & Borson, S. (2008). Time that tells: Critical clock drawing errors for dementia screening. *International Psychogeriatrics, 20(3),* 459 – 470.

Leung, J. C. W., Lui, V. W. C., & Lam, L. C. W. (2005). Screening for early Alzheimer's disease in elderly Chinese patients using the Chinese Clock Drawing Test. *Hong Kong Journal of Psychiatry, 15(1),* 14 – 17.

Libon, D. J., Malamut, B.L., Swenson, R., Sands, L. P., & Cloud, B. S. (1996). Further analyses of clock drawing among demented and nondemented older subjects. *Archives of Clinical Neuropsychology, 11,* 193 – 205.

Libon, D. J., Swenson, R. A., Baronski, E. J., & Sands, L. P. (1993). Clock drawing as an assessment tool for dementia. *Archives of Clinical Neuropsychology, 8,* 405 – 415.

Lin, K-N., Wang, P-N., Chen, C., Chiu, Y-H., Kuo, C-C., Chuang, Y-Y., & Liu, H.C. (2003). The three-item clock drawing test: A simplified screening test for Alzheimer's disease.

Luria, A. (1966). *Higher cortical functions of man*. New York: Basic Books.

Manos, P. J. (1999). Ten-point clock test sensitivity for Alzheimer's disease in patient with MMSE scores greater than 23. *International Journal of Geriatric Psychiatry, 14,* 454 – 458.

Mayer-Gross, W. (1935). Some observations on apraxia. *Proceedings of the Royal Society of Medicine, 28,* 1203 – 1212.

Mendez, M. F., Ala, T., Underwood, K. L. (1992). Development of scoring criteria for the clock drawing task in Alzheimer's disease. *Journal of the American Geriatrics Society, 40,* 1095 – 1099.

Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H., Matsuda, M. (2005). Neural correlates of impaired performance on the clock drawing test in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 19,* 390 – 396.

Nishiwaki, Y., Breeze, E., Smeeth, L., Bulpitt, C. J., Peters, R., & Fletcher, A. E. (2004). Validity of the Clock-Drawing Test as a screening tool for cognitive impairment in the elderly.

O'Rourke, N., Tuokko, H., Hayden, S., & Beattie, B. L. (1997). Early identification of dementia: Predictive validity of the clock test. *Archives of clinical neuropsychology, 12(3),* 257 – 267.

Peters, R. & Pinto, E. M. (2008). Predictive value of the Clock-Drawing Test. *Dementia and Geriatric Disorders, 26,* 351 – 355.

Pinto, E. & Peters, R. (2009). Literature review of the Clock Drawing Test as a tool for dementia screening. *Dementia and Geriatric Cognitive Disorders, 27(3)*, 201 – 213.

Poppelreuter, W. (1917). *Die psychischen schadigungen durch kopfschuss im kriege 1914 – 1916 die storungen der niederen und hoheren schleistungen durch verletzungen des okzipitalhirns*. Leipzig: Voss.

Powell, D. H., Kaplan, E. F., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). *Microcog: Assessment of cognitive functioning (version 2.1)* [computer software], The Psychological Corporation: San Antonio, TX.

Rosman, A. S. & Korsten, M. A. (2007). Application of summary receiver operating characteristics s(ROC) analysis to diagnostic clinical testing. *Advances in Medical Sciences, 52,* 76 – 82.

Rouleau, I., Salmon, D. P., & Butters, N. (1996). Longitudinal analysis of clock drawing in Alzheimer's disease patients. *Brain and Cognition, 31,* 17 – 34.

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.

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

Schramm, U., Berger, G., Muller, R., Kratzsch, T., Peters, J., & Frolich, L. (2002). Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population. *International Journal of Geriatric Psychiatry, 17,* 254 – 260.

Shua-Haim, J., Koppuzha, G., Shua-Haim, V., & Gross, J. (1997). A simple score system for clock drawing in patients with Alzheimer's disease. *American Journal of Alzheimer's Disease, 12(5),* 212 – 215.

Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry, 15,* 548 – 561.

Shulman, K. I., Gold, D. P., Cohen, C. A., & Zucchero, C. A. (1993). Clock drawing and dementia in the community: A longitudinal study. *International Journal of Geriatric Psychiatry, 8,* 487 – 496.

Shulman, K. I., Herrmann, N., Brodaty, H., Chiu, H., Lawlor, B., Ritchie, K., & Scanlan, J.M. (2006). IPA survey of brief cognitive screening instruments. *International Psychgeriatrics, 18(2),* 281 – 294.

Shulman, K. I., Sheldetsky, R., & Silver, I. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry, 1,* 135 – 140.

Sliwinski, M., Lipton, R., Buschke, H., & Wasylyshan, C. (2003). Optimizing cognitive test norms for detection. In Petersen, R. C. (Ed), *Mild cognitive impairment: Aging to Alzheimer's disease* (pp. 89 – 104). New York, NY: Oxford University Press.

Smith, A. D., Gilchrist, I. D., Butler, S. H., Muir, K., Bone, I., Reeves, I., & Harvey, M. (2007). Non-lateralized deficits of drawing production in hemispatial neglect. *Brain and Cognition, 64,* 150 – 157.

Solomon, P. R., Hirschoff, A., Kelly, B., Relin, M., Brush, M., DeVeaux, R. D., & Pendlebury, W. W. (1998). A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease. *Archives of Neurology, 55(3),* 349 – 355.

Spreen, O. & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed.). New York, NY: Oxford University Press.

Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock drawing in Alzheimer's disease: A novel measure of dementia severity. *Journal of the American Geriatrics Society, 37,* 725 – 729.

Teng, E. L. & Chui, H. C. (1987). Modified Mini-Mental Status (3MS) examination. *Journal of Clinical Psychiatry, 48,* 314 – 318.

Tuokko, H., Hadjistavropoulos, T., Miller, J. A., & Beattie, B. L. (1992). The clock test: A sensitive measure to differentiate normal elderly from those with Alzheimer disease. *Jornal of the American Geriatrics Society, 40,* 579 – 584.

Tuokko, H., Hadjistavropoulos, T., Miller, J. A., Horton, A., & Beattie, B. L. (1995). *The Clock Test: Administration and scoring manual*. Toronto: Mental Health Systems.

Tuokko, H., Kristjansson, E., & Miller, J. (1995). Neuropsychological detection of dementia: An overview of the neuropsychological component of the Canadian Study of Health and Aging. *Journal of Clinical and Experimental Neuropsychology, 17(3),* 352 – 373.

Tuokko, H. & O'Connell, M. E. (2006). A review of quantified approaches to the qualitative assessment of clock drawing. In Poreh, A. M. (Ed.), *The quantified process approach to neuropsychological assessment,* (pp. 173 – 206). Philadelphia, PA: Taylor & Francis.

Van der Horst, L. (1934). Constructive apraxia: Psychological views on the conception of space. *Journal of Nervous and Mental Disease, 80,* 645 – 650.

Watson, Y. I., Arfken, C. L., & Birge, S. J. (1993). Clock completion: An objective screening test for dementia. *Journal of the American Geriatrics Society, 41,* 1235 – 1240.

Williams, T. F. (1992). Report on panel crafting guidelines to help screen for dementia. *APA Monitor,* p. 4.

Wilson, B., Cockburn, J., Halligan, P.W. (1987). Behavioural Inattention Test. Titchfield, UK: Thames Valley Test Company.

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