THE CLOCKME SYSTEM: COMPUTER-ASSISTED SCREENING TOOL FOR DEMENTIA

A Dissertation
Presented to
The Academic Faculty

by

Hyungsin Kim

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
College of Computing

Georgia Institute of Technology
May 2013

THE CLOCKME SYSTEM: COMPUTER-ASSISTED SCREENING TOOL FOR DEMENTIA

Approved by:

Dr. Ellen Yi-Luen Do, Advisor
School of Interactive Computing &
College of Architecture
Georgia Institute of Technology

Dr. Jim Foley
School of Interactive Computing
Georgia Institute of Technology

Dr. Mark Guzdial
School of Interactive Computing
Georgia Institute of Technology

Dr. Janet S Cellar
School of Medicine & Emory
Alzheimer's Disease Research Center
Emory University

Dr. Michael J. Jacobson
Centre for Computer Supported
Learning and Cognition
University of Sydney

Date Approved: December. 6, 2012

To my family for their endless love and support

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep appreciation for my advisor and mentor, Ellen Yi-Luen Do. I thank you for everything you have done for me, both academically and personally. You first gave me the opportunity to work with you and the freedom to explore diverse research topics that interested me. You completely taught me every single step to becoming a proficient researcher. I learned how to write academic papers and grant proposals, and how to make good presentations. Your expert guidance led my intellectual development. I especially thank you for your enormous amount of time and attention devoted to reading the many drafts of my works in progress and more amazingly, your comments were both incredibly quick and incisive. Furthermore, you believed in me and gave me strength and confidence when I doubted myself. It is hard for me to express my gratitude for everything I had learned from you here. You deserve my most sincere and special thanks for literally countless reasons. You are the reason for where I am today. Truly, you are the *best* role model I could ever imagine. I hope to continue our collaboration throughout the rest of my career.

Second, I next want to thank my dissertation committee members, for each and every one of whom I could not have more esteem or affection: Jim Foley, Mark Guzdial, Janet Cellar, and Michael J. Jacobson. Thank you to Jim, who helped me to construct several ideas during the early stage of my dissertation work. You, a great scholar and a wonderful person, have always amazed me with your depth of thinking. My gratitude also goes to Mark, who introduced me to a new College of Computing Ph.D. program, Human Centered Computing, at Georgia Tech. I learned this program from you and now have finished my doctoral study. Your supportive comments and advice helped me make my dissertation complete. I would also like to thank Janet for your suggestions and keen critiques that have made a profoundly positive impact on my work. In particular, you have always been willing to share your immense expertise in the field of Alzheimer's Diseases and dementia screening. As a result, your feedback has shaped my work to be more relevant. I am very grateful to have you on my committee. Michael, my lifetime advisor and mentor, I have been extraordinarily fortunate to have you on my committee. You have always encouraged me throughout my doctoral studies in numerous ways. Moreover, I thank also Melody Moore Jackson. I had the opportunity to know you when I worked as your teaching assistant for three semesters. Since then, you have become my mentor, and you always encouraged me to persevere with my studies despite any difficult situation. It has been a true joy and honor to know and work with you.

Third, this dissertation work lies at the intersection of the field of Human Computer Interaction and Medicine. To make my work more significant in medical practice, I was truly fortunate to work with researchers and doctors at the Emory Alzheimer's Disease Research Center (ADRC). I thank all those who gave me invaluable feedback about the ClockMe System; moreover, I thank those who provided direct advice in my research: Dr. Allan Levey, Dr. James Lah, Dr. David Gutman, Dr. Felicia Goldstein, Dr. Ted Johnson, and Dr. David Loring. I appreciate your time and the opportunities you gave me to make this interdisciplinary work happen. I would also like to thank the medical specialists at the Emory ADRC, who supported me in conducting all user studies with the patients: Marie Walters, Jonathan Huh, and Julie Kozarsky. Marie and Jonathan especially spent several hours evaluating more than one hundred clock drawings. I know how hard it was to do this with your busy schedule at the clinic. I deeply appreciate your time and effort spent collecting patients' surveys during the five-month deployment study. Without you, this field study could not have been done.

Fourth, I would like to thank all of the study participants in my research for their feedback, their patience, and their time. Most importantly, you encouraged me by emphasizing that the work I was doing was very meaningful to you. You gave me a great deal of motivation, which made me continue working with the strong belief that my research is truly important for many people. Working with the actual potential users of the system was frequently challenging in various ways; nevertheless, I found much joy and many rewards in being an HCI researcher.

Fifth, I am grateful for the generous funding during the course of my doctoral training and for my dissertation research. This study was initiated by the support of Georgia Tech's Health Systems Institute Seed Grant. This research is also based upon work supported by the National Science Foundation's Smart Health and Wellbeing under Grant Number 1117665. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The final year of this dissertation work was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000454. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. In addition, I was also supported by a feasibility study grant at the Korea Institute for the Advancement of Technology (KIAT)'s Global Industry Academia Cooperation Program.

Sixth, I thank the Association for Computing Machinery (ACM) Student Research Competition (SRC) sponsored by Microsoft Research. This dissertation work was evaluated by three rounds of a panel of judges and attendees from the CHI SRC 2011; I won first prize. I was then selected for the SRC Grand Finals 2012, which coincided with the ACM Turing Award; I won third prize. I was very pleased with these recognitions, which gave me the opportunity to meet researchers in the field who had similar interests. I gained much insightful feedback from this experience, which greatly improved my research work.

Seventh, I thank all of the members of the ACME lab, past and present, my fellow HCC and HCI Ph.D. students, and talented undergraduate and master's students who joined as a part of this dissertation work: Chih-Pin Hsiao, Paula Gomez, Szu-Chia Lu, Nicholas Davis, Tina Seunghyun Lee, Matt Swarts, Yan Xu, Lijun Ni, Jeonghwa Yang, Hyunjean Lee, Young Suk Cho, Anupam Guha, Wassa R Panont, Sammer Yadav, Sravya Kotte, Sandeep Segireddy, and Madhuri Tondepu. I very much enjoyed our eating-out club and teatime meetings, which rejuvenated me and gave me a break from my isolated dissertation work. With the friendship and camaraderie of these wonderful people, my graduate studies were not only tolerable, but highly enjoyable. I have truly appreciated your emotional support throughout this sometimes daunting process.

Finally, I thank my parents, Jinhyun Kim and Junghee Hong, for their unconditional support and encouragement to pursue my dream. You always believed in me. Thank you for everything. I am also grateful to my late grandmother, Sunhee Kim, who gave me all kinds of support in every way. You never doubted that I could successfully finish my studies. I wish I could have completed my studies before you

passed away. To my brother, Hyoungsoo Kim, thank you for sharing your life as a Ph.D. student at Georgia Tech. I never expected that when I started my studies that we would have our own labs in the same research building, TSRB. It was fun going to school together. Last, but not least, I truly appreciate my husband, Hakkyun Kim, for his extensive support, love and patience. Without you, I could not have completed this work and my degree. I would like to close my acknowledgement to thank you, my six-month old baby, Alexander Jaemoo Kim for giving me a new joy and pushing me to finish this journey. Over the past six months, you have been a very good baby boy by giving Mommy enough time to write this dissertation. I love all of you deeply.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADRC ....Alzheimer's Disease Research Center

ADRD .... Alzheimer's Disease and its Related Disorder

CDT ....The Clock Drawing Test

MCI ....Mild Cognitive Impairment

SUMMARY

Due to the fastest growing senior population, age-related cognitive impairments, including Alzheimer's disease, are becoming among the most common diseases in the United States. Currently, prevention through delay is considered the best way to tackle Alzheimer's disease and related dementia, as there is no known cure for those diseases. Early detection is crucial, in that screening individuals with Mild Cognitive Impairment may delay its onset and progression. For my dissertation work, I investigate how computing technologies can help medical practitioners detect and monitor cognitive impairment due to dementia, and I develop a computerized sketch-based screening tool.

In this dissertation, I present the design, implementation, and evaluation of the ClockMe System, a computerized Clock Drawing Test. The traditional Clock Drawing Test (CDT) is a rapid and reliable instrument for the early detection of cognitive dysfunction. Neurologists often notice missing or extra numbers in the clock drawings of people with cognitive impairments and use scoring criteria to make a diagnosis and treatment plan. The ClockMe System includes two different applications - (1) the ClockReader for the patients who take the Clock Drawing Test and (2) the ClockAnalyzer for clinicians who use the CDT results to make a diagnosis or to monitor patients.

The contributions of this research are (1) the creation of a computerized screening tool to help clinicians identify cognitive impairment through a more accessible and quick-and-easy screening process; (2) the delivery of computer-collected novel behavioral data, which may offer new insights and a new understanding of a patient's cognition; (3) an in-

depth understanding of different stakeholders and the identification of their common user needs and desires within a complicated healthcare workflow system; and (4) the triangulation of multiple data collection methods such as ethnographical observations, interviews, focus group meetings, and quantitative data from a user survey in a real-world deployment study.

CHAPTER 1

INTRODUCTION

Screening is one of the most effective strategies to fight against illness and disease. People at risk for cancer, one of many life-threatening diseases, have benefitted through routine screening processes. Screening tools, such as mammography and colonoscopy, have been able to catch the early symptoms of specific types of cancer that are treatable (41, 64, 92). Moreover, the World Health Organization (WHO) has emphasized the importance of screening in their paper on the principles and practice of screening for diseases (104).

The senior population is increasing dramatically. Age-related cognitive impairment, such as dementia, is becoming one of the prominent public health challenges in the US (5). Dementia is a set of symptoms, including poor mental functioning associated with confusion, forgetfulness, and difficulty in concentrating, as well as poor functional abilities such as difficulty in completing complex work-related tasks and daily activities (61, 65, 71, 78, 96).

Alzheimer's disease is the top cause of dementia in the US (3). Unlike other diseases that are physically visible, early detection of dementia is rarely easy. In fact, fewer than 50 % of Alzheimer's cases are diagnosed, and only approximately 25% are treated, even after several years of progressive cognitive decline (94).

In 2003, the Alzheimer's Foundation of America proclaimed the second Tuesday of November as National Memory Screening Day (76). The goal of this initiative is to promote the early detection of Alzheimer's disease and related disorders (ADRD) and to encourage timely intervention and treatment. It is often difficult to detect the early stages of cognitive impairment because (1) it is hard to differentiate cognitive impairment from normal cognitive degeneration due to aging (71); (2) there is limited opportunity for seniors to meet with specialists, such as neurologists or neuropsychologists, unless they have serious observable symptoms (66); and (3) the disease usually develops progressively; thus, making a diagnosis at the appropriate moment is challenging, as it normally requires continuous monitoring through everyday activities (24, 31, 38, 65). Furthermore, the disease may also be under-recognized and under-diagnosed due to the lack of a quick, efficient tool to distinguish normal aging from very mild cognitive impairment (43, 44, 66, 90, 98).

In this dissertation, I investigated how computing technologies can advance our understanding of detecting and monitoring cognitive impairment due to dementia. To this end, I designed, implemented, and evaluated a computer-assisted screening tool for dementia, called the ClockMe System.

1.1. Thesis Statement and Research Questions

Computing technologies have shown great potential for support when administering neuropsychological assessments, analyzing and interpreting data, and assisting in clinical decision-making for screening (2, 13, 20, 22). I designed and developed a computerized neuropsychological assessment tool to assist in screening, monitoring, and diagnosing dementia.

The following is my thesis statement:

"A computer-assisted screening tool for dementia will be able to provide a more accessible and quick-and-easy screening process. The tool will offer a better

opportunity for medical practitioners to identify cognitive impairment. The tool will also be able to increase the consistency of neuropsychological assessments. Furthermore, the tool will provide novel behavioral data, which may be able to offer new insights into understanding cognition."

A computer-assisted dementia screening tool through automated sketch analysis would provide medical practitioners with an efficient, effective, and valuable way to screen people at risk for cognitive impairment. Automatically captured and stored data through such a tool would provide an effective way to identify people with dementia and would replace the tedious effort and error of human scoring with a consistent scoring practice and analysis. Furthermore, capturing and monitoring data through computerized systems can provide additional types of data such as airtime, pausing tendency, millisecond-level completion time, planning strategy, and patterns of exerting pressure. These novel behavioral data could be applied to the process of making or modeling diagnoses, resulting in more evidence-based decision-making.

This dissertation has three research questions, which I answered through the iterative design, development, and evaluation process of the ClockMe System.

Research Question 1 (RQ1)

What do doctors need? What are the design requirements of a computerized tool for neuropsychologists and neurologists to do an analysis, screening, and diagnosis for dementia?

Research Question 2 (RQ2)

Can the computerized tool yield the same results as the paper-and-pencil Clock Drawing Test?

- Do patients feel equally comfortable using both methods?
- Do patients perform the same?
- Do technicians score both drawings the same way?

Research Question 3 (RQ3)

Can a computerized tool perform better? Would a computerized tool provide more information? Can this tool increase the consistency of cognitive impairment assessment? Can this tool provide new diagnostic insights for medical practitioners about their patients by providing concrete data beyond that offered by the paper-and-pencil test?

My first research initiative for this dissertation work started as an investigation of the currently available types of neurological assessment tools by conducting a literature review. I consulted with medical technicians, clinical nurse practitioners, neuropsychologists, and neurologists at the Emory Alzheimer's Disease Research Center (ADRC) about their screening assessments. I also conducted a series of observational studies of the current practice of dementia screening sessions at Emory ADRC and I interviewed older adults who attended the sessions. Through the observational study I conducted, I learned about the Clock Drawing Test (CDT), which is one of the most popularly used tools in current dementia screening. The advantage of the CDT as a screening tool is that it is very quick and easy to administer. However, despite its simplicity, there are several issues. One issue can be the way in which it is administered, that is, manually. A patient who wants to take a test needs to have a one-on-one meeting with a clinician. After that, the clinician needs to spend time evaluating the test. Furthermore, there are no universally agreed-upon scoring criteria available. With this lack of consensus, different clinicians may evaluate the same drawing differently. Furthermore, I found that there are many interesting features available, especially when I observed how a patient constructs the clock. Unfortunately, in current practice, technicians most commonly conduct the testing, and physicians are only given the final performance scores without ever having observed the patient. Thus, clinicians lose an opportunity to review many features of qualitative data that could inform further diagnostic insights. Kaplan, one of the most famous neuropsychologists, argues that the current neurological assessment methods should be replaced by what she calls the Boston Process Approach (17, 26, 74). Despite its value, there are many difficulties in conducting a process approach, such as a lack of expertise and time: not all clinicians have the necessary skills to conduct a process analysis, which is very time consuming. In brief, designing a computerized test can enhance these weaknesses and address those issues.

These formative studies led me to design the ClockMe System. I was able to answer Research Question 1 by drawing the design requirements to develop a computerized Clock Drawing Test. Then, I implemented the ClockMe System and conducted a preliminary study to investigate whether it was well designed for deployment in a real clinical environment. The results from the preliminary study showed the positive potential of a computerized assessment and helped me answer Research Question 2. Furthermore, these results also encouraged me to conduct a longitudinal study with a larger number of participants. Thus, I deployed and evaluated the ClockMe System at the

Emory Alzheimer's Disease Research Center (ADRC) over a five-month period. The deployment study first provided me with a definitive answer to Research Question 2, which was that the computer drawings were equivalent to the paper-and-pencil clock drawings. Second, through the user study and interviews with medical practitioners, I was able to answer Research Question 3, in that the computerized assessment had more potential to improve the current screening practice. Table 1 shows a summary of the thesis studies in terms of the research questions, completed studies, and their outcomes.

1.2. Dissertation Overview

The dissertation is structured as follows. In Chapter 2, I provide a review of the background and a discussion of related work. Chapter 3 describes a formative study of the current dementia screening practices to inform the design of the ClockMe System. Then, Chapter 4 reports two approaches to technological implementation with respect to the sketch recognition of the proposed system. In Chapter 5, I present the design and implementation of the ClockMe System, together with the preliminary study. Chapter 6 includes the evaluation results of the ClockReader System. Finally, in Chapter 7, I end this dissertation with my conclusion and suggestions for future research.

Table 1. Summary of Thesis Studies

Research Question	Studies Completed	Outcomes
RQ1	Formative studies conducted at Emory ADRC • 24-session observations of cognitive screening practices • Two focus group interviews with medical practitioners • Data analysis of the 50 Clock Drawing Test datasets previously collected by Emory ADRC • Computer-experience surveys together with their demographic surveys • Drawing comparisons between paper-and-pencil and Tablet PC-based drawings	Described in Chapter 3. Resulted in (52) and the related publication (37, 48).
RQ2	Preliminary study of the ClockReader Deployment and Evaluation • Two drawings (both computer-clock and paper-clock) per individual patient evaluated by medical technicians • User surveys, interviews, and clock drawings with patients	Described in Chapters 5 & 6. Resulted in (53) and the related publication (15, 36, 50, 54).
RQ3	Deployment study of the ClockMe Evaluation • Interviews, surveys, and user study with medical practitioners and medical technicians • Statistical analysis of correlations of the computer-collected data with other patients' medical data	Described in Chapter 6.

CHAPTER 2

BACKGROUND AND RELATED WORK

This chapter provides background information for this dissertation regarding dementia, the Clock Drawing Test, and computerized cognitive assessment tools.

2.1. Dementia

Dementia is becoming one of the most common diseases in the United States due to the highly increasing aging population (3). Unlike diseases that are physically visible, early detection of dementia is rarely easy. Dementia is defined as a progressive and irreversible clinical syndrome that includes poor mental functioning associated with confusion, forgetfulness, and difficulty in concentrating, as well as poor functional abilities such as difficulty in completing complex work-related tasks and daily activities. (61, 65, 71, 78, 96).

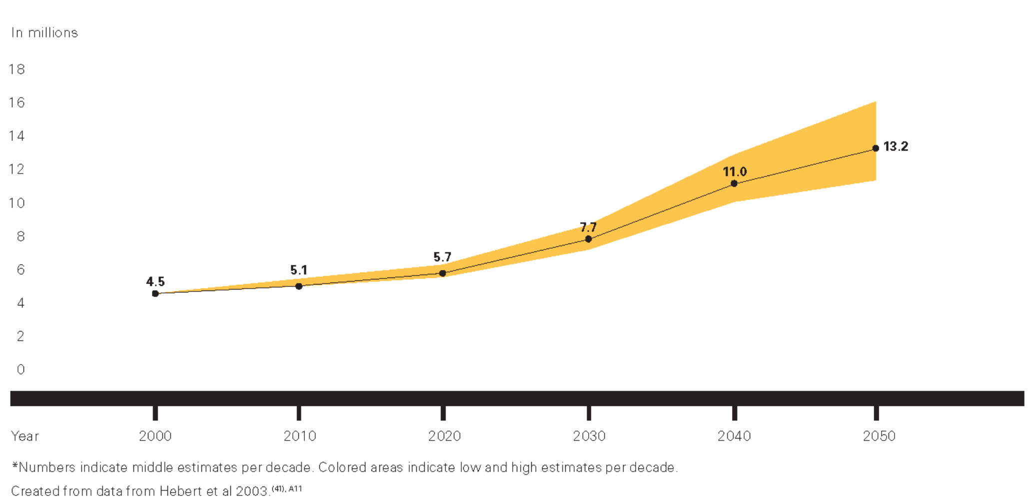
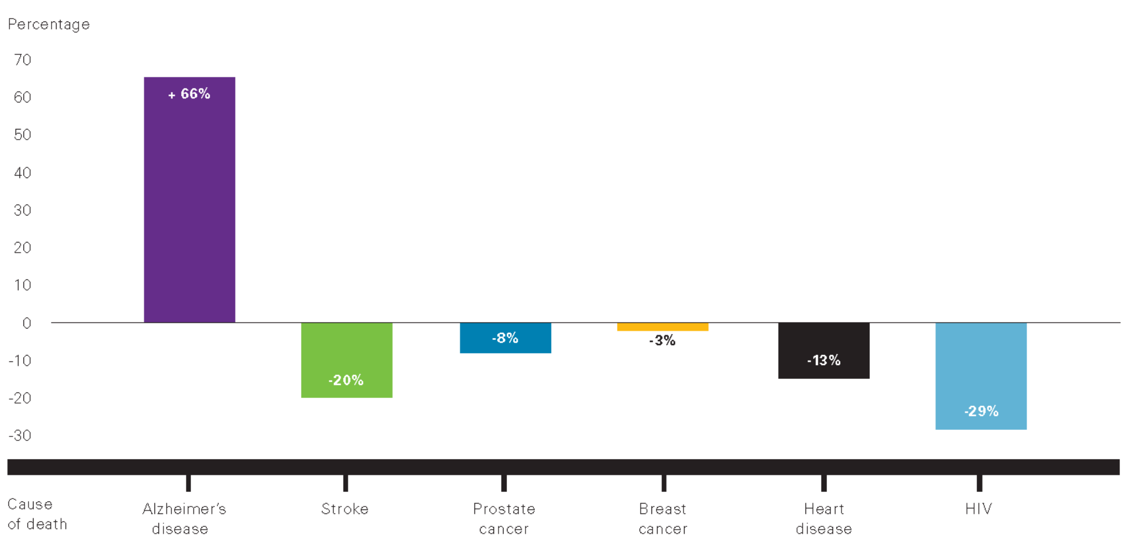


Figure 1. Projected Number of People Aged 65 and Over in the U.S. Population with Alzheimer's Disease (in Millions) Using the U.S. Census Bureau Estimates of Population Growth, reproduced by the Alzheimer's Association 2011 Report (3)

Figure 1 shows the projected number of people aged 65 and over in the U.S. with Alzheimer's disease using the U.S. Census Bureau estimates of population growth (39). By 2031, the first wave of baby boomers will reach age 85, and the number of people aged 65 and older with Alzheimer's disease is projected to be around 7.7 million, which is a 50 percent increase from the number of people currently estimated to have the disease (3, 39).

The most common forms of dementia are Alzheimer's disease and vascular dementia, with more than 60-65% of dementia cases due to Alzheimer's disease (3, 30). Representative symptoms for people with Alzheimer's are recent or episodic memory deficits (18). The memory deficits can either be verbal (why one went to the grocery store) or visual-spatial (how to find one's way back home after shopping) or some combination of the two (6, 12, 29, 93).



a National Center for Health Statistics. *Deaths: Final Data for 2000.*(80)
b National Center for Health Statistics. *Deaths: Preliminary Data for 2008.*(75)

Figure 2. Percentage Changes in Selected Causes of Death (All Ages) between 2000 and 2008, reproduced from the Alzheimer's Association 2011 Report (3)

Figure 2 above shows the increased number of deaths due to Alzheimer's disease compared to other diseases (3). Between 2000 and 2008, the number of deaths caused by Alzheimer's increased approximately 66 percent. Interestingly, over the same period, the number of deaths caused by other diseases such as stroke and heart disease decreased 20 percent and 13 percent, respectively.

People with Mild Cognitive Impairment (MCI) have a high risk factor of developing Alzheimer's disease (12, 60, 68, 70, 102). According to the American Academy of Neurology, individuals with MCI are defined as those who have a measurable memory deficit, but without the other key characteristics of dementia, which include confusion (71). There are different types of MCI. Some people with MCI will have cognitive impairment in multiple areas of cognitive function in addition to memory. Others may have normal functioning in their memory, but some other cognitive functions can be abnormal. Figure 3 shows the progression of cognitive impairment in people with MCI as a pre-dementia stage.

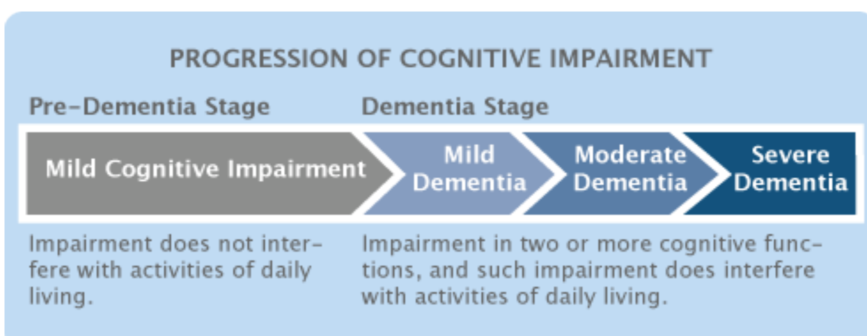


Figure 3. Progression of Cognitive Impairment from Prevented AD.com (77)

The Quality Standards Subcommittee of the American Academy of Neurology reviewed the evidence-based medical data on the benefits of early dementia detection

(71). The subcommittee's goal is to develop scientifically sound, clinically relevant practice parameters for detecting Mild Cognitive Impairment. They recommend that since patients with MCI are at the greatest risk of developing dementia, individuals with MCI should be further evaluated and regularly monitored (71). In particular, screening individuals with Mild Cognitive Impairment (MCI) for dementia is crucial. MCI is thought to be an intermediate (or transitional) clinical state between normal aging and the very earliest stage of Alzheimer's disease (68). Therefore, to prevent dementia, it is critical to screen people with MCI.

2.2. The Clock Drawing Test

The Clock Drawing Test (CDT) has been used for decades as a stand-alone neuropsychological screening test (30). The CDT is also a component of several widely used neuropsychological assessment batteries: the 7-Minute Screen, the CAMCOG (Cambridge Cognitive Examination), and the Spatial-Quantitative Battery in the Boston Diagnostic Aphasia Examination (62). The CDT focuses on visual-spatial, constructional, and higher-order cognitive abilities, including executive functioning (17). The CDT accesses the human cognitive domains of comprehension, planning, visual memory, visual-spatial ability, motor programming and execution, abstraction, concentration, and response inhibition (42, 96). The major value of the CDT for medical practitioners is that the CDT provides a concrete visual reference of a patient's cognitive impairment. As a result, this provides good information that can be used to detect cognitive dysfunction.

Figure 4 shows an example of visual neglect in a clock drawing from a stroke patient. The patient clearly omits the left side of the available clock face when drawing the clock features. Even though the patient could verbally express that the clock face had a left side, he failed to notice that the drawing was incomplete. Drawing tasks can play an important role in helping clinicians localize specific impairments of brain lesions; in this case, the right hemisphere.

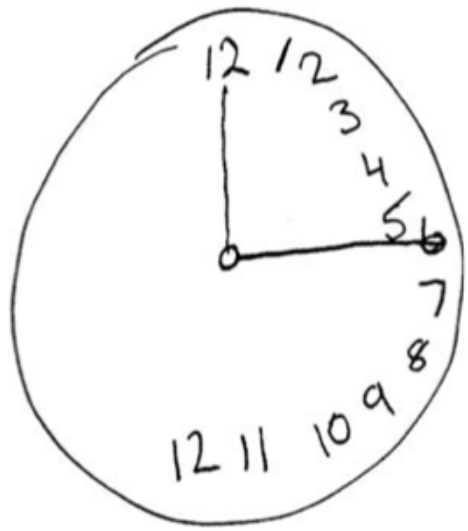


Figure 4. An example of a Clock Drawing from a patient with hemi-spatial neglect. Reproduced from Smith et al. (91)

The Clock Drawing Test (CDT) is one of the simplest, and most commonly used screening tools to detect cognitive impairment in seniors (45, 67). Simply asking people to draw a clock can easily discriminate those with dementia from those without it (30). In general, clock drawings from people with cognitive impairment frequently show missing or extra numbers, or misplaced clock hands (26).

As an illustrative example, the drawings in Figure 5 clearly show degradation of cognition in the three figures with abnormal clock drawings: two from patients with Alzheimer's disease, and one from a patient with suspected frontotemporal dementia. The patients were asked to draw a clock by putting the numbers on the clock and setting the time at ten past eleven.

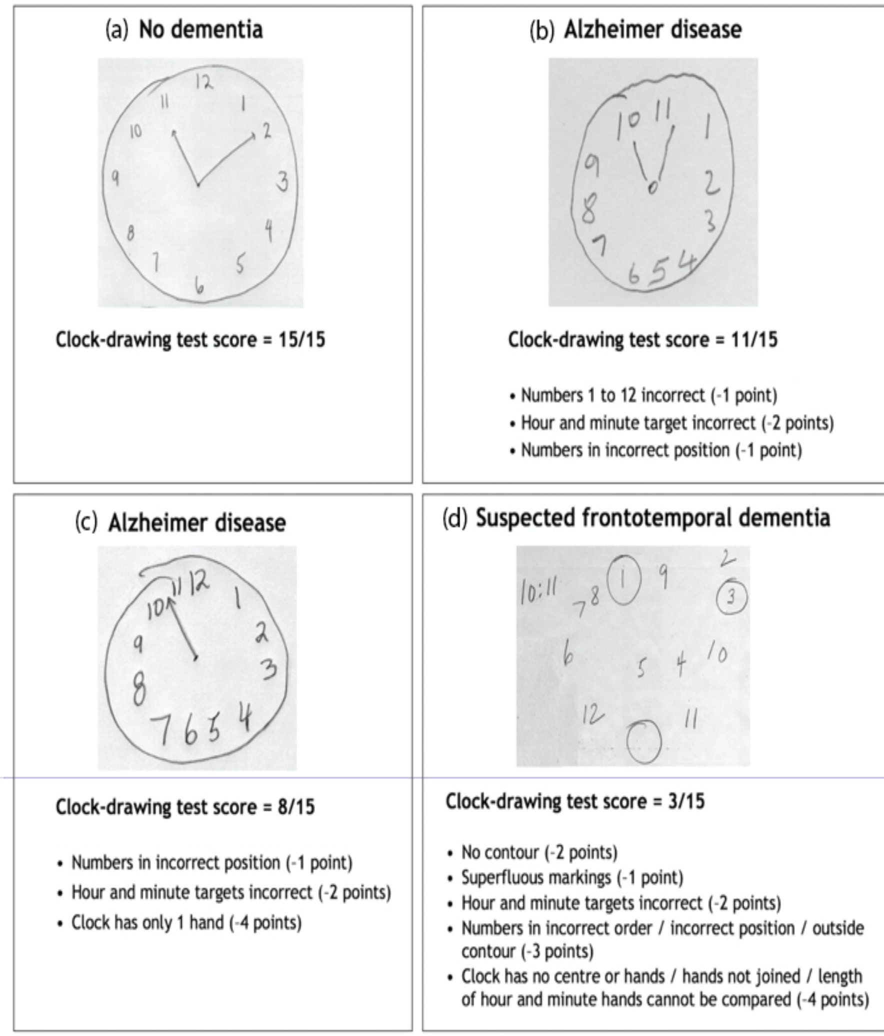


Figure 5. Clock drawing examples of patients without dementia, as well as those with Alzheimer's disease and suspected frontotemporal dementia. Reproduced from Feldman et al. (23)

Freedman's 15-point scoring criteria for free-drawn clocks were used to evaluate the drawings. Patients whose drawing meets all of the criteria will obtain the maximum score of 15. Let's look carefully at four clock drawings, one by one. The top left drawing (a) shows a clock from a person without dementia. It satisfies all of the criteria. Unlike the clock from a person without dementia, the drawings from people with dementia did not satisfy several evaluation criteria. The top right drawing (b) shows a clock from a patient with Alzheimer's disease. It is missing the number 12, and the numbers are not in the correct position. It also indicates the time (11:10) incorrectly. The bottom left drawing (c) also shows a clock from a patient with Alzheimer's disease. The drawing has incorrectly indicated the hour and minute with only one hand. It also includes numbers incorrectly positioned. Lastly, the bottom right clock (d), from a suspected frontotemporal dementia patient, is significantly misdrawn, and thus it does not satisfy most of the criteria. Please see Table 2 for detailed scoring criteria used for the analysis.

Table 2. Freedman's Free-Drawn Clock Evaluation Criteria (26)

Criteria
Contour of the circle
1. Acceptable
2. Not too small, overdrawn or reproduced repeatedly
Numbers
3. Only numbers 1-12 present
4. Only Arabic numbers used
5. Numbers in correct order
6. Numbers drawn without rotating paper
7. Numbers in correct position
8. Numbers all inside contour

Table 2 continued

Hands
9. Two hands present
10. Hour target number indicated
11. Minute target number indicated
12. Hands in correct proportions
13. No superfluous markings
14. Hands relatively joined
Center
15. Center is present (drawn or inferred)

Currently, the clinical administration of the CDT in a hospital environment is performed by medical practitioners such as neuropsychologists, neurologists, and clinical nurse practitioners (31, 88). The patient is asked to draw a clock with a pencil on a provided sheet of paper and to follow instructions given by the clinicians. There are two variations of the test. One test asks patients to draw a clock in a pre-drawn circle. This test focuses on the spatial distribution of the numbers, as well as on the hands of the clock. The other type of test does not provide any circle on the paper. Therefore, patients are asked to draw a freehand circle on the paper. In some cases, patients are shown a picture of a clock drawing and are asked to copy that onto the paper. Patients are then asked to set different times for the clock, such as 11:10, 1:45 or 3:00 (26, 96).

There are numerous scoring systems available. Each of the scoring systems places differing emphases on visual-spatial, executive, quantitative, and qualitative issues (95). Qualitative errors can provide valuable information for medical practitioners to understand the relationship between the different drawing errors and the etiology or progression of diseases. For example, clocks drawn by patients with right frontal lesions tend to show difficulty with number position. Clocks drawn by patients with left frontal lobe damage usually show reversal of the minute and hour hand proportions (26). The CDT is one of the most versatile tests available to directly map neural damage to the impairment of visual-spatial behavior. Overall, the CDT is accepted as an ideal cognitive screening test, based on widespread clinical use (72). Among published studies, the CDT achieves a mean sensitivity of 85% and a specificity of 85% [1] (87).

2.3. Computerized Cognitive Assessment Tools

A great deal of research in developing computerized cognitive assessment testing has been conducted since the advent of personal computers (13, 22, 47, 84, 105). Many research studies attest that the potential of computerized assessments lies in their precision of measurement, shorter assessment time, standardization, automatic scoring, minimization of subjectivity, and reduction of the impact that examiners may have on the participants (105).

Studies show that computing technologies have many advantages over paper-based tests, from administering behavioral assessments to assisting clinical decision-making processes (11, 58). First, technology can support the administration of the assessment process by automating data collection. Second, it can also increase the

---

[1] Sensitivity and specificity are two statistical measures for validating a screening test. Sensitivity is the probability of testing positive if the disease is truly present. Specificity is the probability of screening negative if the disease is truly absent (69). In medicine, test sensitivity refers to the test's ability to correctly identify people with the disease, whereas test specificity relates to the test's ability to correctly identify those without the disease (55, 88).

precision of measurement. The precision of time measurement is necessary to assess a person's cognitive condition with respect to reaction times and information processing speed (13, 14, 21). The millisecond levels of time-stamping data can be captured and stored. Third, automatic scoring can support the data analysis process and can decrease the subjectivity of test-examiners' interpretations, all of which can increase scoring standardization. Moreover, computing technologies, such as pen-based interaction, can be collected as novel behavioral data, which previously could not be measured or captured. For example, pen-based interaction can support collecting stroke-level data from a drawing test. Such pen-based interaction can provide quite a few insights for clinicians because many neuropsychological assessments heavily rely on drawing performance.

The CANTAB (Cambridge Neuropsychological Test Automated Battery) system is one of the best-known computer batteries designed to comprehensively evaluate cognitive abilities (63). The CANTAB system can cover a diverse age range from 4 to about 90 years. It can also be used to test patients for dementia, Parkinson's disease, depression, schizophrenia, HIV and learning disabilities in children (96). The battery consists of 13 sub-tests, including motor skills, visual attention, memory, and working memory; it takes approximately 90 minutes to complete. Because it is so long, the CANTAB is not suitable for dementia screening (33). With this limitation, it has been used almost exclusively in research settings (33, 34). Please see Table 3 for detailed tasks included in the CANTAB system.

Table 3. CANTAB Tasks and Approximate Time per Test (96)

Test	Approximate Time
Motor Screening	3 min
Matching to Sample Visual Search	9 min
Delayed Matching to Sample	10 min
Pattern Recognition Memory (Immediate/Delayed)	5 min each
Spatial Recognition Memory	5 min
Paired Associate Learning	10 min
Spatial Span	5 min each
Spatial Working Memory	8 min
Big Little Circle	3 min
Intra/Extra Dimensional Shift	7 min
Rapid Visual Information Processing	7 min
Reaction Time	5 min
Stockings of Cambridge	10 min

CNSVS (Central Nervous System Vital Signs), another computerized neuropsychological test battery, consists of seven conventional neurocognitive tests: verbal memory, visual memory, finger tapping, symbol digit coding, the Stroop test, a shifting attention test, and a continuous performance test (32, 35). Administering the test requires about half an hour. CNSVS has been used to test patients with brain injuries, dementia, and Attention Deficit Hyperactivity Disorder (ADHD). Figures 6 and 7 show the results of reaction times and attention errors among different subject groups such as normal, MCI, and demented. The figures show that the demented group clearly has the longest reaction time, as well as the highest attention errors. One study on the reliability and validity of CNSVS shows that it is appropriate as a screening tool. However, it is still not sensitive enough to replace formal neuropsychological testing because it is not diagnostic (32, 35). Gualtieri's study also emphasizes that CNSVS can play a critical role as an intermediate tool for neuropsychologists (33). This means that CNSVS can provide more precise information than a patient's subjective complaints. Moreover, when used as an evaluation tool in the physician's office at clinics, physicians were able to screen patients who required referral to a neuropsychologist for further examination (5, 33). This evidence shows the strong potential of the clinical utility of using a computerized assessment as a screening tool in a doctor's office.

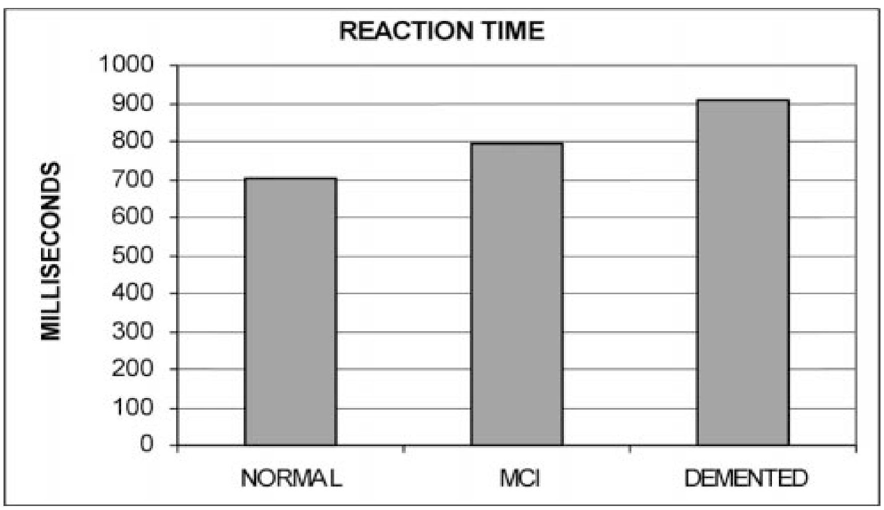


Figure 6. Reaction Times in three different groups: Normal, Patients with MCI and Demented. Reproduced from Gualtieri CT (34).

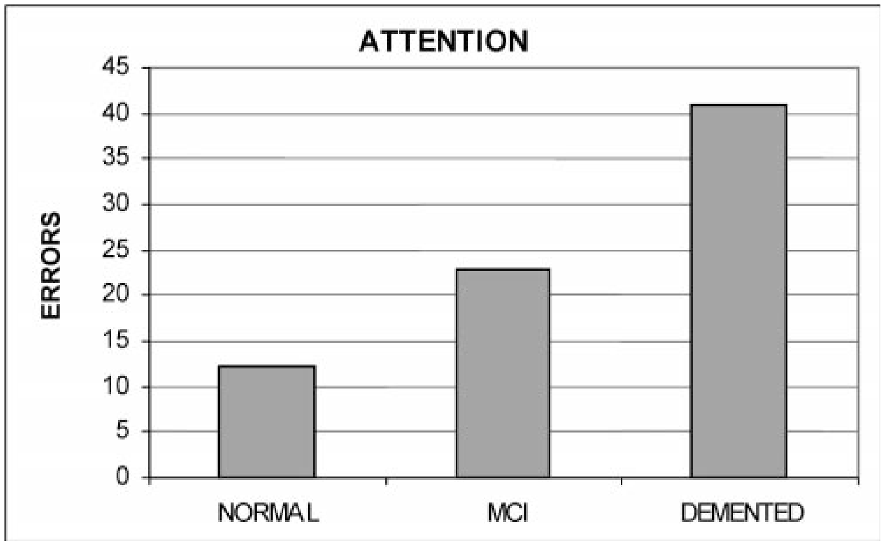


Figure 7. Attention Errors in three different groups: Normal, Patients with MCI and Demented. Reproduced from Gualtieri CT (34).

Besides the CANTAB and CNSVS, many computerized tests have been investigated and developed (20, 22, 47, 55, 56, 58, 59, 63). However, none of the tests include drawing tasks. Drawings of all kinds have been used as a fundamental source for neuropsychologists to understand their patients (46, 91). Lezak argues that drawing tasks should take a central position in neuropsychological assessment because they are a rich source of information in understanding the presence (or absence) of cognitive abilities (61). However, until now, most computerized assessments have not included a sketch-based drawing analysis. Rather, the assessments have simply modified the drawing tests into other types of tests that can be easily implemented by keyboard and mouse interactions. This lack of drawing test implementation shows that despite many trials on developing computerized assessments, they are still not used in real clinical environments. Rather, computers are still used mostly to support lab-based experiments.

CHAPTER 3

FORMATIVE STUDY:

UNDERSTANDING USERS AND TECHNOLOGY DESIGN SPACE

In order to gain an in-depth understanding of the technological potential, I conducted a formative study before designing and implementing the ClockMe system. This chapter presents my formative study to explore the technological opportunities for cognitive impairment screening. The findings from the study, as well as the implications for design, will be described here.

3.1. Overview

The main purpose of the formative study is to understand the technological potentials to support dementia screening by observing how medical practitioners conduct and administer the Clock Drawing Test (CDT) in their screening routines.

The formative study consists of five different research activities: (1) observations of 24cognitive screening sessions in clinical practices; (2) surveys of computer experience and demographic information of the session participants; (3) a preliminary usability comparison test between paper-and-pencil and Tablet PC-based drawings; (4) two focus group meetings with medical practitioners at the Emory University School of Medicine and Emory Alzheimer's Disease Research Center (ADRC); and (5) a data analysis of the 50 Clock Drawing Test datasets previously collected by Emory ADRC research.

The Emory ADRC researchers, including neurologists, neuropsychologists, clinical nurse practitioners, and medical technicians supported all of the research activities, such as the data collection and analysis. The Emory IRB, together with the Georgia Tech IRB, also approved our research. This enabled us to access Emory ADRC's resources, such as patients and their medical records.

3.2. Understanding How the Clock Drawing Test is Conducted in Practice

My approach started from gaining contextual knowledge (8), such as understanding the screening process, interactions between clinicians and patients in the screening process, and interactions with materials. I conducted ethnographical observational studies at the Emory ADRC. During March to August 2010, I was fully immersed as an observer at the Emory ADRC. First, I learned how to conduct different dementia screening methods from the doctors and medical technicians. Then, I thoroughly observed how medical technicians worked with patients and how they interacted with people such as other medical technicians, doctors, and patients. When they moved from one place (in the session) to another (where the data were stored) with the data, I shadowed their movements and made my own field notes. By keeping notes, I was able to understand their work process at a highly detailed level.

In brief, the medical technicians stayed in their own offices to prepare the materials for cognitive screening prior to having sessions with the patients. When a patient came to the clinic, they met each one in a single room for an approximately two-hour session. When a patient came to the clinic, they met each one in a single room for an approximately two-hour session. The session includes (1) several paper-and-pencil tests: The Clock Drawing Test, Mini-Mental Status Examination, Boston Naming, Digit Span, Geriatric Depression Scale, Word List Memory, WMS-R Logical Memory, BVMT-R & Delay, Trials A and B Testing, Category Fluency, and WAIS-R Digit Symbol; (2) interviews with an individual patient for past and current medical history; and (3) interviews with the patient's study partner. All Honor volunteers must bring a "study partner" to accompany them to their research visit. The study partner must be someone who has at least 10 hours of contact with an individual patient each week. The role of the partner is to explain about a patient's memory and thinking by answering a set of questions. The paper-and-pencil tests take approximately one-and-a-half hours to complete, and interviews are conducted in the remaining time period. Of course, the time varies, based on the patient's communication condition. Specifically, patients spent less than 1 minute to complete the CDT. Later, experienced technicians spent about one hour to score the paper-and-pencil tests; technicians reported that an individual Clock Drawing Test scoring usually takes 3 to 5 minutes.

With the support from the Emory ADRC, I was also fortunate to have the privilege to watch the full process of how the CDT, as well as other cognitive assessments were conducted during the individual patient's screening session. In most cases, I sat behind the medical technicians, which enabled me to shadow their work without causing any interference. By being close to the technicians, I was easily able to see interactions between the technician-patient dyad. Furthermore, at the end of the day, I had time to debrief my observations with the technicians. This debriefing greatly helped me clarify whether I correctly understood their practices. During the observations, I frequently tried to understand their working environment preferences with respect to technology acceptance and integration. For example, if they used Information

Technology in their practice, I wanted to know how they might want to change the technology or what things they liked, as well as what they did not like. Interestingly, after a two-week observation period, the medical technicians and I quickly became more comfortable with each other's presence.

Here are the lists of specific questions to which I found answers through my observations:

- **Who administers and analyzes the CDT?**
- **How is the CDT collected and stored?**
- **What is used to evaluate the individual CDT?**
- **How would medical practitioners use the CDT results when they diagnose a patient?**

At Emory ADRC, technicians conduct the dementia screening tests. These medical technicians are trained to administer neuropsychological testing to patients, interpret these test results, and prepare the reports for clinicians. The current practice of the Clock Drawing Test (CDT) is administered with traditional analog media. Patients are asked to draw a clock by using a pencil on a given sheet of paper with a pre-drawn circle. The instructions are as follows: "*Pretend the circle is like a clock face. Put in all of the numbers and set the time at ten after eleven*."

Afterward, the medical technician would need to spend some time analyzing and scoring the tests. To score the tests, the technicians at Emory use the 13-point evaluation criteria developed by Freedman et al.(26). Having a full score of 13 indicates that the drawing is intact and the patient has no cognitive impairment. First, the technician examines the drawing to determine whether all of the numbers are present with only Arabic numbers. Then, the technician looks at whether the numbers are in the correct order, together with the correct position. Lastly, they see whether the numbers are all inside the circle. As for items associated with the hands, all that matters is the depiction of time. Whether the two hands are present and are indicated based on the instructions, such as ten after eleven with the correct proportions, are examined. Also, there should not be any superfluous markings. The center should indicate the joining point of the two hands. As shown in Table 4, the clock consists of 13 critical items. These 13 items can be grouped with respect to numbers (6 items), hands (6 items), and a center (1 item) (57, 62, 82).

Table 4. An Evaluation Criteria of the Clock Drawing Test (26)

Numbers
Criterion 1: Only numbers 1 – 12 are present (without adding extra numbers or omitting any)
Criterion 2: Only Arabic numbers are used (no spelling, e.g., “one, two,” no Roman numerals)
Criterion 3: Numbers are in the correct order (regardless of how many numbers there are)
Criterion 4: Numbers are drawn without rotating the paper
Criterion 5: Numbers are in the correct position (fairly close to their quadrants & within the pre-drawn circle)
Criterion 6: Numbers are all inside the circle

Table 4 continued

Depiction of Time
Criterion 7: Two hands are present (can be wedges or straight lines; Only 2 are present)
Criterion 8: The hour target number is indicated (somehow indicated, either by hands, arrows, lines, etc.)
Criterion 9: The minute target number is indicated (somehow indicated, either by hands, arrows, lines, etc.)
Criterion 10: The hands are in correct proportions (if the subject indicates which one is which after "finishing," have him/her fix the proportions until he/she feels they are correct)
Criterion 11: There are no superfluous markings (extra numbers or errors on the clock that were corrected, but not completely erased, are not superfluous markings)
Criterion 12: The hands are relatively joined (within 12mm; this does not need to happen in the middle of the circle)

Center
Criterion 13: A Center (of the pre-drawn circle) is present (drawn or inferred) at the joining of the hands

Once the medical technician completes the analysis of the individual patient's CDT, he or she stores the CDT drawing together with the analysis sheet in an individual patient's folder. Later, when a clinician sees the patient, he or she refers to the CDT results from the folder. The clinician examines the total score of the CDT and considers it with other screening tests results as a part of the diagnostic process.

From the observation of the paper-and-pencil memory testing sessions, I also identified three issues that can be supported by technologies. First, the drawings are predominantly analyzed at the end of the tests as a final product. Second, medical specialists are required to do many things at the same time, such as measuring the time and observing the patient simultaneously. As a result, some critical moments can easily be missed. Third, the way in which data are collected and stored is not appropriate for monitoring any changes of cognitive condition.

3.3. Understanding Data from the Clock Drawing Test

To understand the overall common error patterns of the Clock Drawing Test (CDT), we collected and analyzed 50 datasets of patients who were registered as a research participant at Emory ADRC. The 50 datasets were collected from 30 unique patients who took the CDT from 2007 to 2009. To understand the longitudinal CDT scoring changes, some of the tests were progressively done on the same patients throughout multiple years. A total of 30 participants' CDTs were collected. For example, three sets of CDTs were collected from four different patients. Two sets of CDTs were collected from twelve patients. Moreover, we only collected a single CDT from the last fourteen patients. Table 5 below shows the number of CDTs from an individual patient, as well as the total number of patients. For the purpose of this analysis, we collected 50 CDT samples from 30 unique individuals.

Table 5. Summary of the 50 Datasets by the number of CDTs per individual patient

Number of CDTs per Individual patient	Number of Participants	Total Number
1	14	14
2	12	24
3	4	12
	30	50

Then, based on Freedman's 13-point scoring criteria, we scored them. The results show that the two most frequent errors are related to (1) Criterion 10: Proportions of the hands; and (2) Criterion 5: Correct position. Since we analyzed the CDTs based on the collected data, Criterion 4 was not evaluated. Table 6 shows the overall incorrect instances per Freedman's CDT Criteria.

Table 6. The number of incorrect instances per Freedman's CDT Criteria

Criteria	The number of incorrect instances
1. 1~ 12 are present.	13
2. Only digits are present.	2
3. Correct order	4
4. No rotation of the paper	N/A
5. Correct position	26
6. Inside circle	7
7. Two hands are present.	13
8. Hour indicated correctly	14
9. Minute indicated correctly	19
10. Correct proportions of the hands	28
11. No superfluous markings	6
12. The hands are relatively joined.	10
13. A center is present.	19

Let me explain how the criteria are evaluated with specific drawing examples in Figures 8 and 9. First, Figure 8 below shows two clock drawing examples from two different patients. The drawings clearly show that the patients were not able to draw the numbers in the correct position (Criterion 5). The CDT on the right side also misses the digit 12.

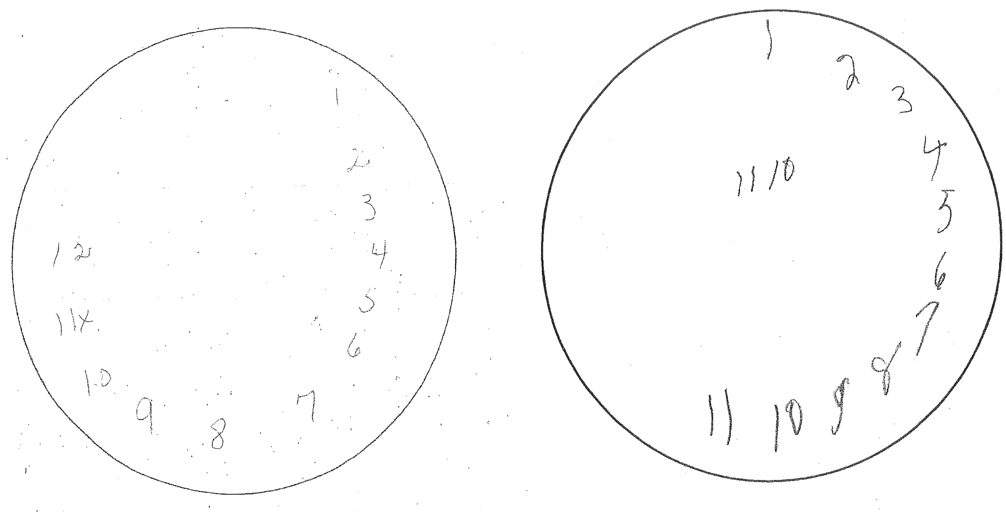


Figure 8. Two clock drawing examples showing that the numbers are not correct in position.

Second, Figure 9 below shows two clock drawings from two different patients, demonstrating a clear difference in the proportions of the two hands. The CDT on the left side shows the correct proportions of the hands. The hour hand is relatively shorter than the minute hand. However, the CDT on the right side includes approximately the same lengths for both the hour and minute hands. Thus, Criterion 10 was not satisfied for the CDT on the right side.

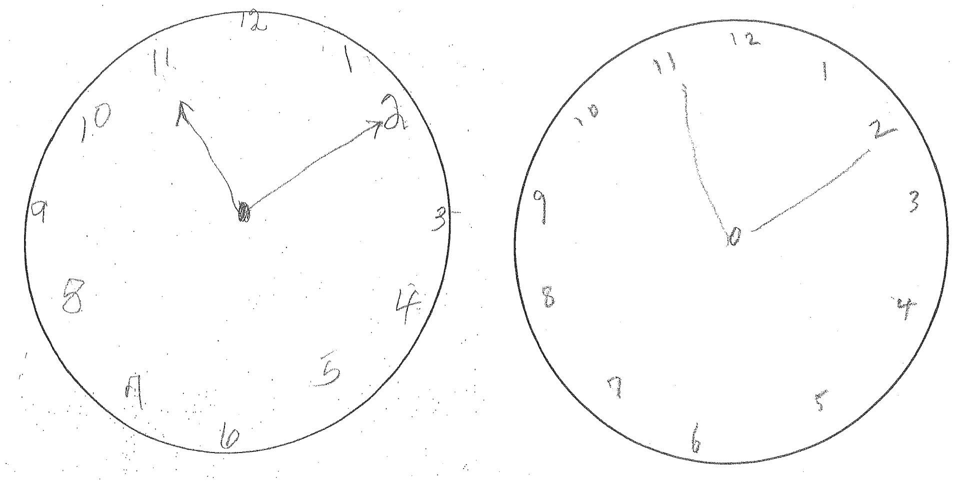


Figure 9. Two clock drawing examples showing correct proportions of the hands (clock on the left side) and incorrect proportions of the hands (clock on the right side).

Another interesting finding from the 50 CDT data we analyzed is that many patients' CDTs show errors with respect to the time setting at ten after eleven. Figure 10 shows six different clock drawings in which the patients were asked to set the time at 11:10. Only the first clock drawing in the top left column (a) shows the correct time setting. The other five clocks show incorrect time settings to indicate what they believed was ten after eleven. The data are consistent with other studies' results, which show that patients with Alzheimer's disease frequently set the time at 10 minutes to 11 rather than 10 minutes after 11 (1, 9, 97, 100). Freedman et al. also argue that this particular time setting placement is difficult for people with Alzheimer's disease (26). Moreover, some other studies show that stimulus-bound responses are more common among Alzheimer's disease patients compared to normal elderly individuals and patients with frontal lobe dementia (9, 62, 83).

Clearly, time setting errors are an important sign for medical practitioners to pay special attention to in their patients. Even though the total CDT score shows a good range (for example, 10 to 13), this incorrectly indicated minute may demonstrate that the patients need to conduct other assessments for more detailed examinations.

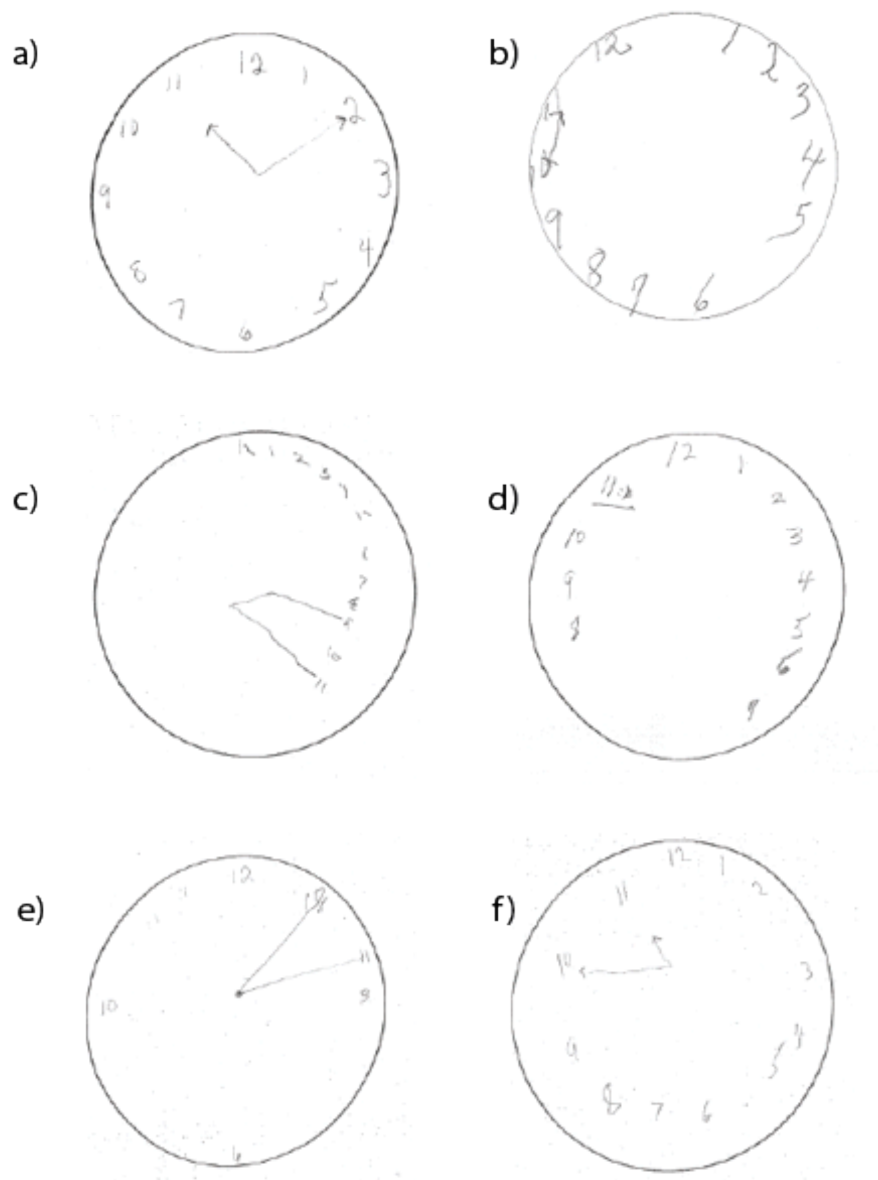


Figure 10. Six different clock drawings that show a time setting at 11:10. Only the clock in the top left-hand corner shows the correct time. All of the other clocks are dementia patients from the ADRC sample. Time setting errors are the most prominent errors of dementia.

Figure 11 below shows eight CDTs from four different patients during a two-year-period. The four CDTs on the left side all show the correctly indicated time setting. All of the patients were asked to repeat the CDT one year later; all four repeat tests show incorrectly indicated time settings. Although all of the digits are present (not for Patient3, Year 2) with the correct order and position, the time setting was the most apparent error. In general, time setting seems to be the most difficult task for patients with dementia because of the high demand function. Or incorrect clock hand placement may be one of the first signs of dementia in abnormal clock drawings. Thus, monitoring CDT results can be a sensitive indicator of a patient's deteriorating cognition.

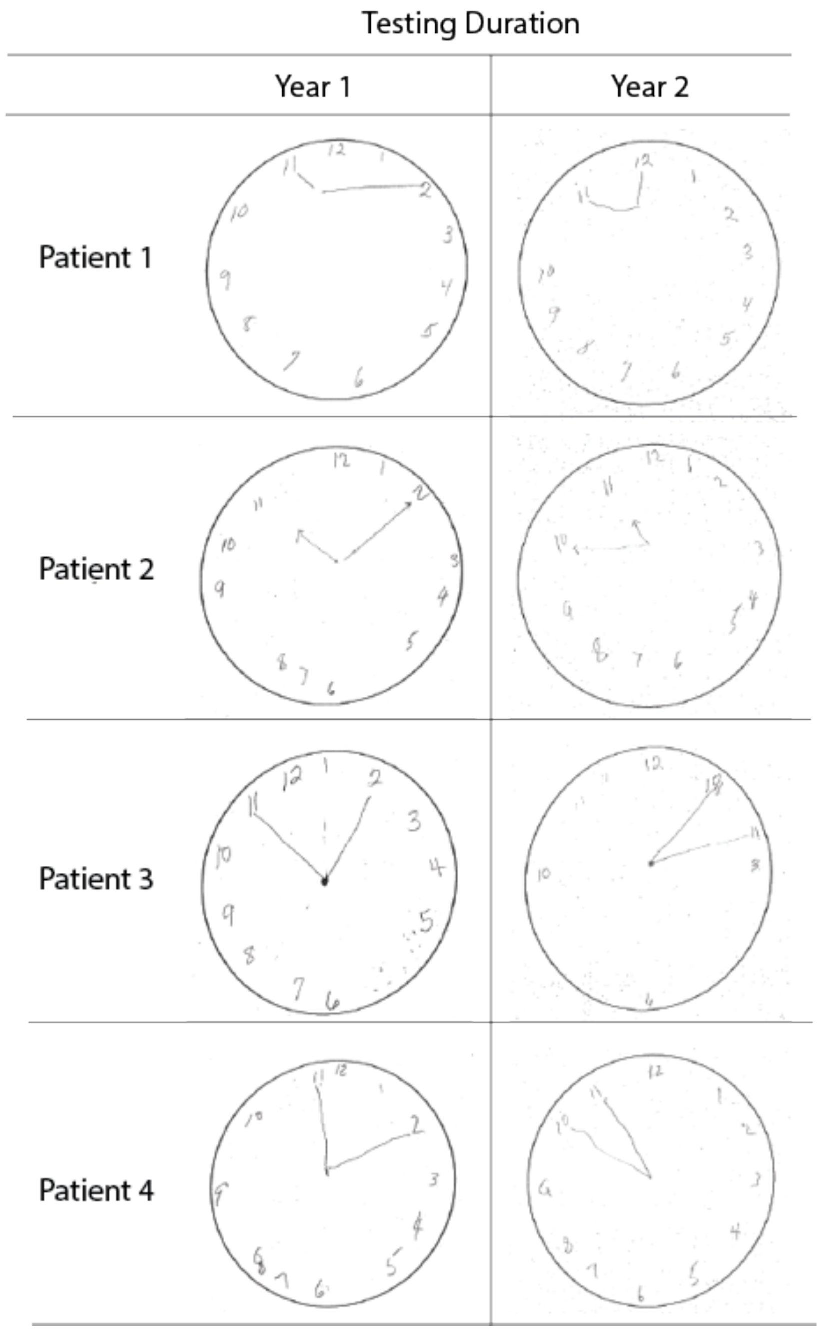


Figure 11. Two clock drawing samples each from four patients over a two-year period.

Lastly, Figure 12 below shows three successive clock drawings by two patients from 2008 to 2009. The drawings clearly show the degradation of the patients' cognition. There is clear evidence of visual-spatial disturbance. Interestingly, the patients could not use the space of the clock face evenly. We can see that the shapes of the digits are the most resilient compared to the positions of the digits. Even with a seemingly high degree of impairment (poor use of the clock face), approximately all of the digits are clearly recognizable.

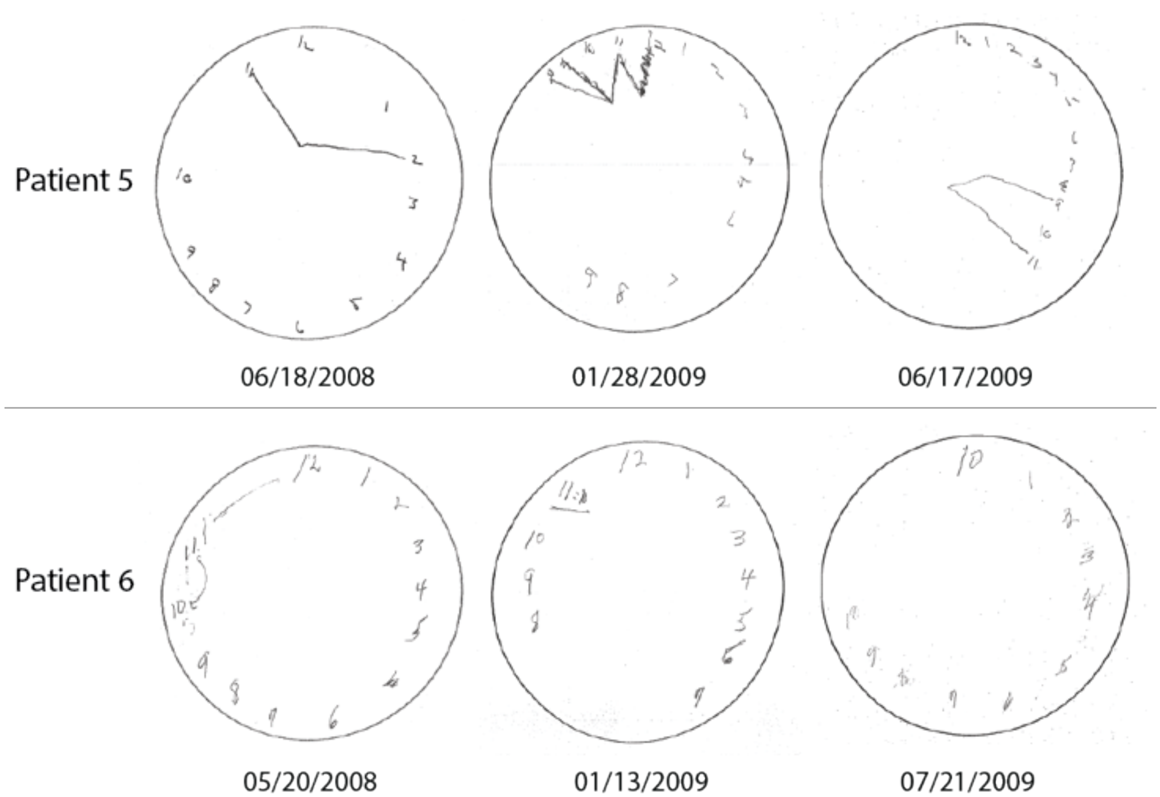


Figure 12. Two sets of three clock drawings clearly showing the cognitive deterioration of two patients over a two-year period.

3.4. Understanding Patients

Understanding patients is critical because they are potential users of the system. Therefore, I conducted a survey of likely users to obtain (1) their demographic information; (2) their computer familiarity; and (3) a comparison of their performing a free-drawing task using both a Tablet PC computer and a paper-and-pencil medium.

The study participants were volunteers recruited from the Honor Research Registry (Honor) of the Clinical Research in Neurology (CRIN) database. The CRIN research registry database is maintained by the Department of Neurology, Emory University School of Medicine and Emory ADRC in Atlanta. As a long-term study initiated by Emory ADRC, Honor studies enable medical practitioners to follow people over their lifetime to learn about memory and thinking. Thus, Honor registry volunteers have provided useful knowledge on the brain changes of people with normal aging, as well as people with disease.

In order to enroll in Honor registry, participants must meet the following inclusion criteria: over age 70 with no memory problems; any age with diagnosed mild cognitive impairment; any age with very mild memory or thinking problems; or any age with mild to moderate Alzheimer's disease. Honor volunteers must participate in at least one Honor study each year. They should bring their study partner with them when they visit Emory ADRC because the study partner is also involved by being asked about the volunteer's day-to-day functioning. Each Honor study takes approximately three hours. It starts with collecting the volunteer's past and current medical history and his or her family medical history, interviews with the volunteer's study partner, and an interview with the volunteer. Then, the volunteer completes a paper-and-pencil memory test and has blood drawn for genetic research.

For my study, 24 older adult volunteers (average age = 75 years old; range= 59~ 98) participated. The participants were 11 females and 13 males. The group included 10 individuals who were older than 78. Educational levels were diverse: Six participants had a high school diploma; nine participants were college graduates; nine participants had graduate-level education. Table 7 shows a brief summary of the participants.

Table 7. Number of Participants by Gender, Education, and Age

Gender	Number of Participants
Female	11
Male	13
Total	24

Education	Number of Participants
High school	6
College	9
Graduate school	9
Total	24

Age	Number of Participants
50~ 59	1
60~69	6
70~79	12
80~89	2
90~99	3
Total	24

The computer experience survey form consisted of five questions: (1) whether participants are familiar with using a computer; (2) how they rate their computer literacy; (3) what the most uncomfortable thing they feel about using a computer; (4) what they would prefer if they could use a computer differently; and (5) what their most common activities are using a computer. Please see Appendix A for the survey form I used for the study.

The results show that 17 out of the 24 participants were familiar with using computers. The last 10 participants reported that either they had never used a computer or they were not familiar with using computers. Despite their computer experience, most participants considered themselves as having low- or medium-level computer literacy. Fourteen participants out of 24 considered themselves as having low literacy, with no computer experience. The last ten participants considered themselves as medium-computer literate, with limited computer experience. Below, Table 8 shows the results of the computer experience survey.

Table 8. Results of the Computer Experience Survey

1. Are you familiar with using a computer?	Number of Participants
Yes	17
No	7
Total	24

2. Computer Literacy	Number of Participants
High	0
Medium	10
Low	14
Total	24

Table 8 continued

3. What is the most uncomfortable thing about using a computer?	Number of Participants
Reading text on the screen	0
Using a mouse	10
Typing on a keyboard	14
Total	24

4. What would you prefer if you could use a computer differently? Describe it. (For example, I would like to use a computer with voice-activated control)	
No Comments	19
Voice-Activated Control	5
Total	24

5. What are your most common activities using your computer?	Number of Participants
Document/Word Processing	3
Internet	8
Games	2
No Comments	11
Total	24

To see whether using a Tablet PC would affect a participant's drawing, we conducted a free-drawing task comparison study, using both a Tablet PC computer and a paper-and-pencil medium. By asking participants to draw anything freehand, such as a house or flower, we tried to understand the quality of a drawing based on a different medium. In order to avoid any carryover effect, we conducted the study with two different orders of presentation of the two tasks: one group used the Tablet PC with a stylus first, while the other group performed the drawing with a pencil and paper first.

If a participant felt uncomfortable, there may have been differences in the freehand drawing tasks. None of the freehand drawing tasks showed any differences between the two media. Even the oldest participant, a 98-year-old grandmother who had never used a computer in her life, said that it was very fun to draw something on the Tablet PC. Below, Figure 13 shows her drawings of a house and a flower. She also commented that drawing with a stylus was much easier because there is less friction compared to the paper and pencil.

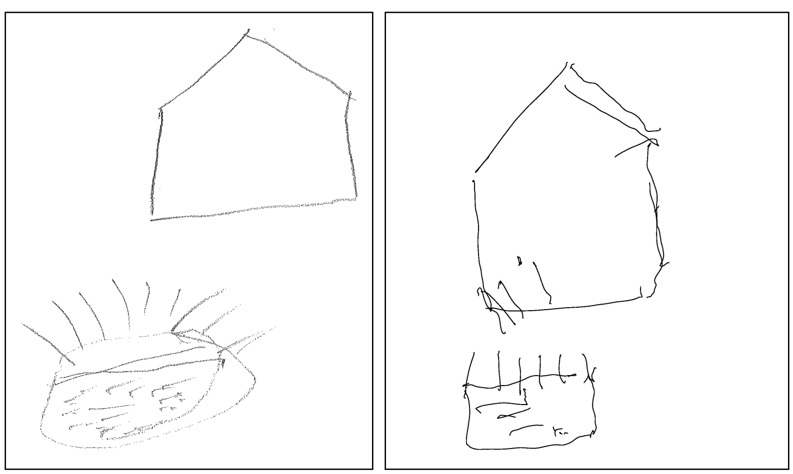


Figure 13. Two drawings from a 98-year-old participant: A paper-and-pencil drawing (left) and a drawing using a stylus on the Tablet PC (right).

Figure 14 below shows another example from a 79-year-old male participant. He drew a house and a flower and also wrote the digits from 1 to 10. The two drawings are almost identical. He had also never used a computer before and identified himself as having low-computer literacy.

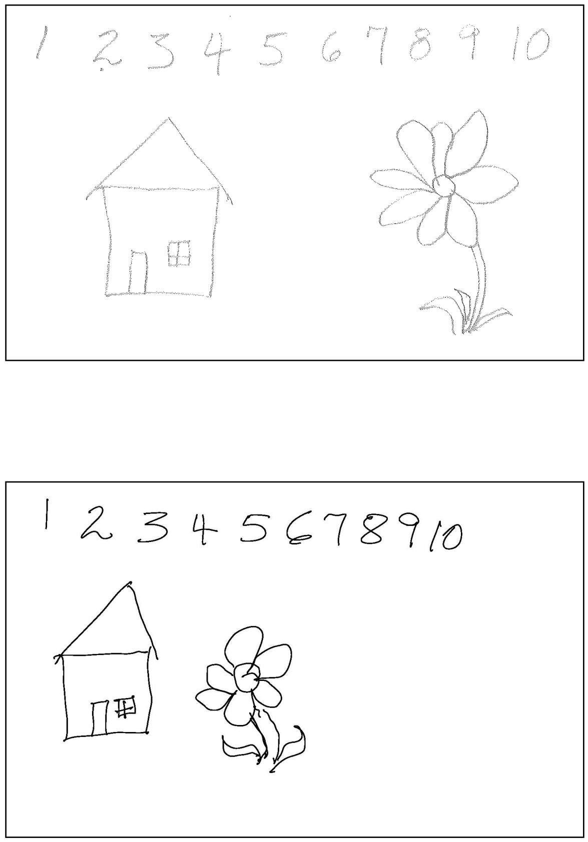


Figure 14. Two drawings from a 79-year-old participant: A paper-and-pencil drawing (top) and a drawing using a stylus on the Tablet PC (bottom).

Overall, the results of the freehand drawing tasks using the Tablet PC are positive. None of the participants expressed difficulties in using a stylus, and all of them successfully completed the drawing task.

In brief, considering the participants' ages, it is not surprising to see the results of the computer experience survey. However, unlike their self-reported lack of computer literacy, the comparisons of drawings (Tablet PC vs. paper and pencil) indicate that there were no critical difficulties for older adults in using a Tablet PC. This finding also supports the results of previous studies on the advantages of using a digital pen compared to other indirect manipulation devices, such as a mouse and keyboard (49, 53, 89). Pen-based interaction is easily accessible to people with little or no experience with computers (4, 52). Users can leverage their previous writing experience without having to learn how to manipulate the keyboard and mouse. Contrary to common perception of older adults and technology use, the Tablet PC environment is amenable for them to perform drawing tests for cognitive screening.

I would like to close this chapter with my reflection of methodological triangulation. As an HCI researcher, it is very important to know one's users, their needs, and the contexts of system use, such as the environment. By triangulating multiple sources of data, I was able to fully understand how the patients and medical technicians would use the system in practice. For example, in order to understand how a patient perceived the use of the computerized screening tool, I gathered data from interviews, observations, and questionnaires. From in-person interviews, I could understand the patients' feelings about using the system from *their own voices*. Through my own observations of the patients using the system, I was able to see and compare *their drawings* by using the computerized tool compared to the drawings by the paper-and-pencil method. Lastly, from the questionnaire results of the demographic information, as well as the computer use, I could gain contextual understanding of an individual patient's background. In brief, data triangulation made me confident about the results I found from the usability study. It provided concrete evidence of system use from three different methods such as the patients' drawings, *their own voices*, and their experience reports on the survey; furthermore, all three sources showed consistent answers. Yet, sometimes triangulation would help me reconcile some contradictory results from the user study. For example, medical technicians and practitioners liked to have a computerized screening tool because of the advantage of using computerized data collection methods and automatic scoring. From the survey completed by the medical practitioners and technicians, they all reported that 1) they liked to use it; 2) they experienced no difficulties in using it; and 3) they would continue to use it further in their practice. However, interestingly, when I had focus group meetings, as well as in-person interviews, many of the technicians reported their preference to include a Print function with the system. This finding shows that the technicians preferred to maintain the conventional practice, even though they were willing to use a new computerized tool. To resolve this conflicting result, I included a Print function to meet the users' preference. I believe that one of the goals of this new system deployment is to gradually change the current practice of work rather than to expect practitioners to radically accept new technologies.

CHAPTER 4

APPROACHES TO SKETCH RECOGNITION TECHNOLOGY

This chapter presents how recognition technologies can be used in developing the automatic Clock Drawing analysis in two different ways: (1) Offline sketch recognition through the Machine Learning algorithm; and (2) Online sketch recognition with the Microsoft Tablet PC SDK (Software Development Kit).

4.1. The Automatic Clock Drawing Analysis with Offline Sketch Recognition

The purpose of our first technology implementation is to develop an automatic Clock Drawing Analysis system using offline handwriting recognition. To implement offline handwriting recognition, we used different machine learning algorithms. The advantage of using the offline recognition method is that it can be used to analyze existing data, such as previously completed pencil-and-paper CDTs. We used the offline recognition method to analyze a large number of clock drawings that had been scanned and stored at the ADRC to study the early stages of different types of dementia.

The automatic Clock Drawing Analysis system was based on three steps. First, a medical technician would scan a completed CDT. Second, the offline recognition algorithm processing would run the scanned drawing data. Third, the data were automatically analyzed and the technician received an automatic final score report.

The offline sketch recognition algorithm was written in MATLAB® using image processing to recognize characters. We wrote the code first to recognize handwritten digits and then recognize the clock hands. We utilized several different machine-learning techniques to recognize handwritten numbers. The algorithm first identifies where the blobs in the figure are, and then if those blobs are digits, the system then processed and recognized each digit. The development process consisted of three stages: image pre-processing, digit recognition, and post-processing.

At stage 1, the image-pre-processing module of the system runs the MATLAB® program with the scanned patients' clock drawing images. The MATLAB® program we designed works with the following rules.

- Clean the image and polarize it by converting every bit to a "1" or "0," based on a threshold.

- If the blobs have nearly similar maximal and minimal width, draw square bounding boxes and store each as an image.

- If the blobs have much greater length to negligible width, assuming that they are meant to depict the clock hands, skeletonize the hands, identify their elevation with respect to the horizontal, and store the hands as separate images.

- Take the digit blobs, resize and clean them again.

Figure 15(left) shows a scanned image of a clock drawing. Figure 15 (right) shows the results of the extraction program. With successful extraction, the program cleans up the output and resizes each digit. The image (right) was inversed to use MATLAB®'s inbuilt functionality to detect connected components.

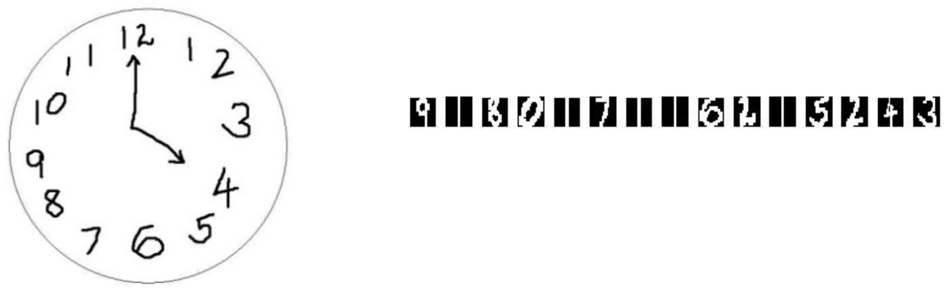


Figure 15. Scanned image (left) and the results after running the extraction program (right)

Stage 2 focuses on digit recognition. After the first stage, the system knows the locations of the digits and can then classify them by using a machine-learning algorithm. We found that neural networks have the best performance from our experiments with machine-learning techniques, such as kNN (k Nearest Neighbor), neural networks, and decision trees on the MNIST dataset (http://yann.lecun.com/exdb/mnist). However, it is incredibly slow to train. To optimize for speed and high accuracy, we chose kNN because the kNN algorithm has a reasonable success rate (85 ~95%) if a large enough (10,000 instances) dataset is used for comparison(36).

After each digit is recognized, the system needs to distinguish a single digit from double-digit numbers, considering each digit's centroid. With centroids and proximity, the system identifies the position of each digit. For example, there are five 1's such as 1, 10, 11, and 12. The 1's that are in close proximity to a 0, a 1, and a 2 are designated as 10, 11, and 12. The proximity is defined as Euclidean distance between centroids being 50% lesser than average. Figure 16 shows the detected centroids of each digit (left) and clock hands (right).

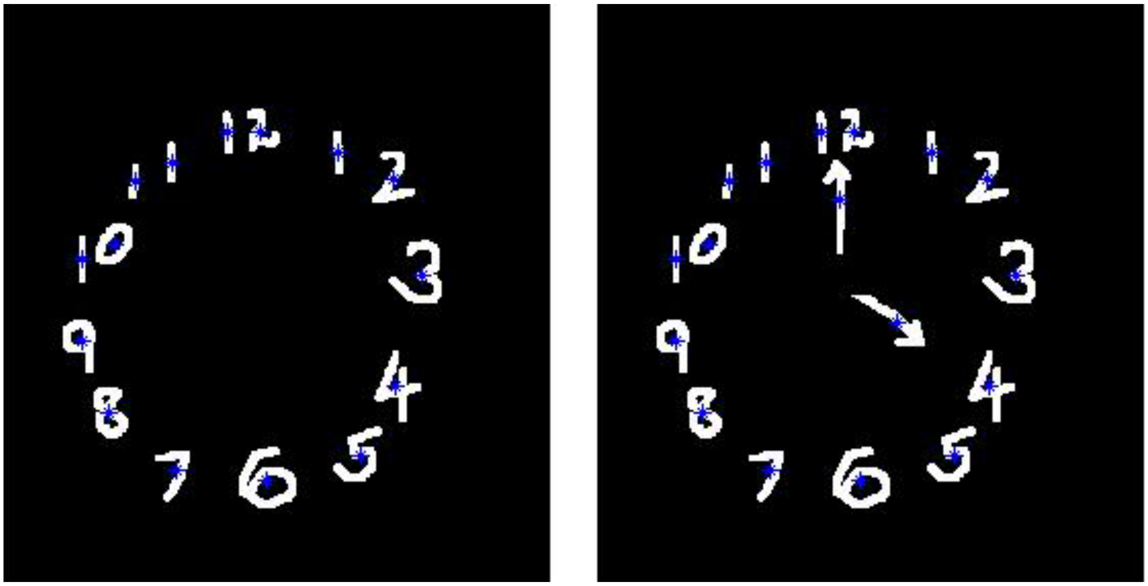


Figure 16. Clock drawings showing the detected centroids of each digit (left) and clock hands (right)

Clock hands can be detected by their geometrical features. For example, two disconnected "blobs" with their major axes longer than 100% of the average major axes of all the blobs that the software has to recognize are probably the two hands. Since the system assumes the larger sketches to be the hands, the lengths of their major axes (measured by the functionality provided in MATLAB®) are known, which is used to identify the minute (longer) and hour hands (shorter). The system identifies the angles of the hands in relation to the horizontal axis to determine the pointing directions. The normal position of the center is defined as anywhere within 20% distance of the radius from the center. Stage 3, the post-processing handles the outputs of the screening test and exports the results into an Excel file. We implemented the 13-point evaluation criteria developed by Freedman et al. (25) for offline sketch recognition system development.

4.2. The Automatic Clock Drawing Analysis with Online Sketch Recognition

Our second approach to technology implementation was to develop an automatic Clock Drawing Analysis system using online handwriting recognition. To develop online handwriting recognition, we used a Tablet PC platform for the input. A patient will be asked to draw a clock with a stylus while running the program. Then through the online sketch recognition algorithms, the drawing will be collected and recorded in a proper data structure. The character recognition method using a Tablet PC would collect dynamic process data from each stroke level and the overall sequence of the drawing and arrangements for analysis.

The advantage of using the online recognition method is that it would collect dynamic process data from each stroke level and overall sequence of the drawing and arrangements for analysis. Another advantage is that Tablet PCs, a representative platform in pen-based computing, provide users with a natural way to interact with a computer (73, 89). Similar to using a pencil to write on paper, one can use a stylus to draw on top of the tablet screen.

The system with the online sketch recognition algorithm is developed in C# programming language and is supported by the "Microsoft Windows XP Tablet PC Edition Software Development Kit 1.7" and "Microsoft Visual Studio 2010." For the first release, the running environment of the program is limited to the Microsoft Windows platform, equivalent to or better than the "Windows 2000 Service Pack 4" with the "Microsoft.Net Framework 3.5 Service Pack 1."

The architecture of the system consists of three components: a GUI for drawing a clock, a character recognition engine, and a data storing utility. Figure 17 shows a screen shot of the initial prototyping system, and the box on the right shows the recognized digits for each stroke.

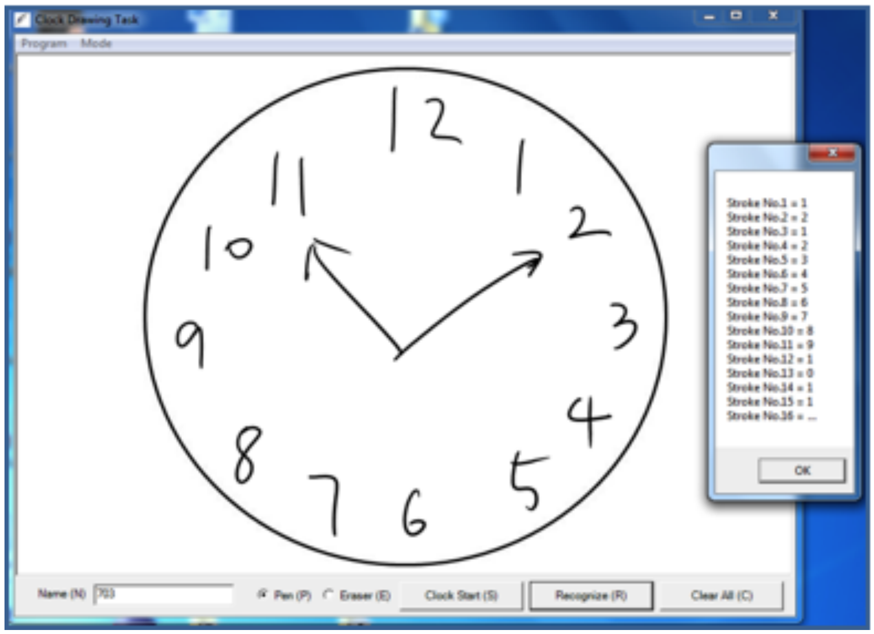


Figure 17. A screen shot of the automatic Clock Drawing System

A test participant starts the entire data flow of the system using a Tablet PC. Once the patient draws the clock as requested, the handwritten recognizer will analyze each number stroke along the inside of the circumference of the given circle and will return the best matched character. Two clock hands that start from near the central point of the circle will have their validity determined by calculating the distance and internal angle between them. It will also collect every end point, intersected point, Bezier point, pressure value and timestamp in each packet point, and airtime of the stroke for the successive possible analysis afterwards. The collected data will be transmitted to database storage and in the hard disk of the Tablet PC simultaneously, just in case the communication is interrupted for any reason. Figure 18 illustrates the data flow overview of the system. Below we will explain each one.

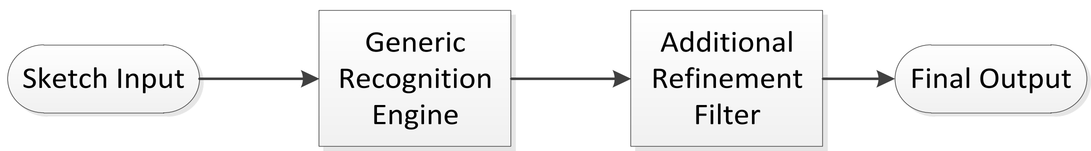


Figure 18. The System Terminal Application Data Flow

- **Sketch Input**: The participant is asked to draw a clock and set the time (e.g., at 10 after 11. Every coordinate of the cusps and intersections of each stroke (even if it represents a character) will be stored in the memory.

- **Generic Recognition Engine**: The processor captures a rectangular-shaped dynamic recognition region for the strokes and recognizes them to the best-matched characters by the Microsoft Handwritten Recognition Engine. The two clock hands will be identified by their relative locations, angles, and the size of the strokes.

- **Additional Refinement Filter**: After the first recognition result is generated, an additional refinement filter enhances recognition accuracy by correcting misrecognized characters. The most frequent errors are misclassifications of numerical digits as alphabet characters. We implemented the Context-Bounded Filter Algorithm (CBFA) (15, 51) to convert the recognized characters into appropriate numeric digits using constraint predicates.

- **Final Output**: When the recognition process is over, the program analyzes the relative position between each number and scores, and stores them into a database with given schemas.

4.2.1. Context-bounded Refine Filter Algorithm (CRFA)

To address the inherent ambiguity of the handwritten data of sketch recognition, we use the context in order to improve a recognizer's accuracy. We implemented a context-specific recognition approach, called the Context-bounded Refine Filter Algorithm(15, 51). Context-bounded recognition relies on the recognizer's processing specifically in a given situation. In the system, the context would be bounded for drawing a clock, which means that people would mostly use alphabetical numbers and lines for depicting clock hands. This allows us to develop several features of the algorithm.

- Algorithms for recognizing digits and clock hands distinguished from unnecessary strokes
- Algorithms for recognizing the numbers from 1 to 12, together with each number's coordinates: the coordinate helps to distinguish a single digit from double-digit numbers considering the sequence of each number's before-and-after position
- Algorithms for automatically calculating the CDT score results, based on pre-programmed criteria

- Algorithms for excluding unnecessary strokes during the process of evaluation

A stroke is a computerized drawing element composed of a series of packet points from a tablet computer when users put down the stylus on the screen until they lift it up. Figure 6 shows that there are three categories for strokes to represent the numbers in a clock drawing: one stroke for a single digit, multiple strokes for a single digit (Figures 19-A and 19-C), and two single-digit characters for a double-digit character (Figure 19-B).

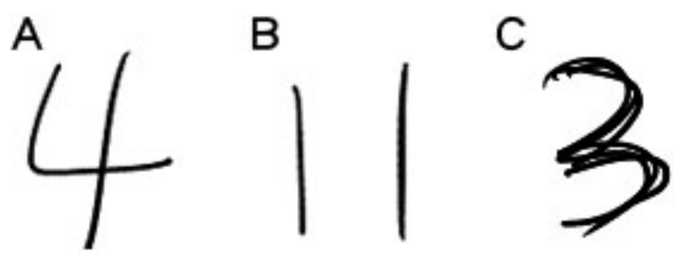


Figure 19. Three different numbers constructed by multiple strokes

These categories are determined by whether they are close to each other and by the contextual information in the clock. According to the location and the size of the strokes in each category, we create bounding rectangles as the recognition region for the generic recognition engine to make the best decisions. However, the recognition engine cannot always give us accurate results because it excludes the contextual understanding of handwritten data. For example, it would misrecognize 6 as 9 or 1 as l. For these cases, the Context-Bounded Filter Algorithm (CBFA) comes in and helps refine the results. CBFA uses the relative locations between each number and circle to decide whether the results from the generic recognition engine are correct.

In order to calculate the relative locations, we to convert the ink coordinate system used in the Ink Serial Format (ISF) file to the pixel coordinate system in which the circle is drawn. This conversion also benefits the calculation of the scores.

4.2.2. Data Structure

The relative locations, relations between strokes, and recognition results are stored in a separate database, along with the original ISF file after a user presses the "recognize" button in the system. The goal is to represent the contextual information and collect the data in a clock when the participant is drawing a clock. In the database, the Clock_Stroke entity, which contains each packet point, bounding rectangles, and properties, such as timestamps, is the most rudimentary object in this database. However, compared to the Stroke object in the ISF file, it has the relationship attributes to indicate the related strokes with it. For example, when two strokes compose the single digit "4," such as Figure 19-A, we add the id key of the second stroke in 4 to the "MergeTo" attribute of the first stroke. Likewise, we also add the id key of the first stroke to the "MergeTo" attribute of the second stroke. In the case of a two-digit number, we simply add the id of the associated stroke(s) in one digit to the "CombineTo" attribute of the stroke in the other. With these relationships, the bounding rectangles of multiple strokes for one digit and two digits for one number can be calculated from the bounding rectangle of one single stroke when analyzing and visualizing data in the ClockReader System. Figure 20 shows the structure of the database.

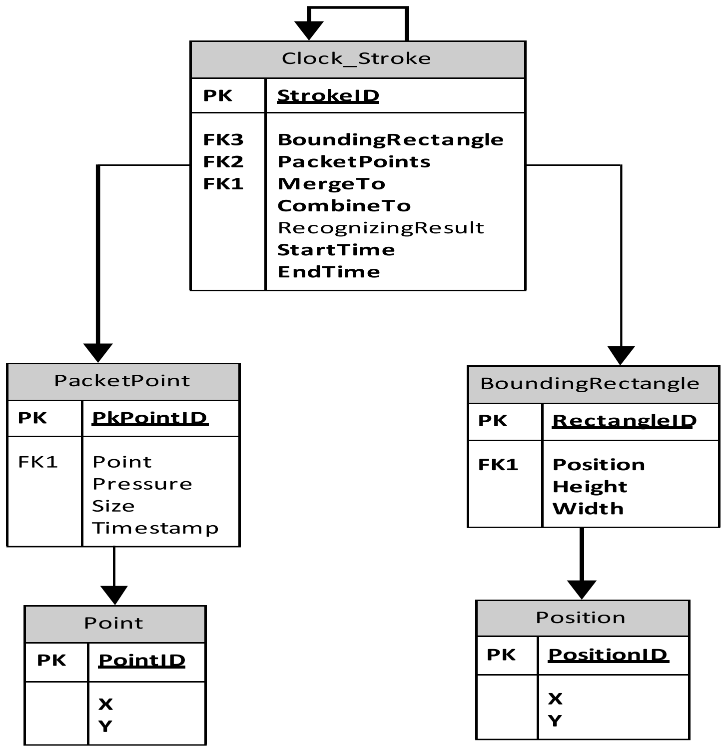


Figure 20. Database for storing the contextual information relating to the circle

4.3. Summary

We implemented two different prototypes of an automatic Clock Drawing Analysis System. In brief, the first approach entails asking a patient to draw a clock on a given piece of paper, and a medical technician would scan the drawing. Then, through offline sketch recognition algorithm processing, the scanned drawing is automatically analyzed, and the final score is reported. The second approach involves using a Tablet PC platform for the input. A patient is asked to draw a clock with a stylus on the running program. Once the patient finishes drawing a clock, through the online sketch recognition algorithms, the drawing will be collected, recorded, and analyzed. Each approach has its own advantages and can be used effectively in a different situation. For example, some clinics may tend to stick with the traditional paper-and–pencil-based assessment. However, they can adopt the computerized scoring method through offline recognition technology. By doing so, this would increase the accuracy and consistency of data analysis and interpretation. Others may choose to adopt advanced technology such as online recognition technology, which would offer benefits such as data collection, analysis and interpretation, as well as data management.

After discussing these two approaches with researchers at Emory ADRC, I decided to develop a prototyping system on a Tablet PC platform. Beyond simple scoring and analyzing the static data, we can leverage the computational features of the online recognition method to enhance the understanding of a patient's drawing caused by dementia. In the next section, I will introduce the ClockMe System, which is implemented based on the online sketch recognition method.

CHAPTER 5

DESIGN AND IMPLEMENTATION OF THE CLOCKME SYSTEM

This chapter introduces the overview of the ClockMe System, the user interface design of the ClockReader and the ClockAnalyzer, and the unique functional features of the ClockMe system.

5.1. Overview of the ClockMe System

The overarching goal of the ClockMe System is to develop an automated analysis of the Clock Drawing Test. First, the system records and recognizes a patient's freehand drawing data. Then based on the scoring criteria, the system automatically analyzes the drawing and reports the score. The system can reduce the labor needed for manual scoring and can improve the consistency of scoring. Furthermore, capturing and monitoring data through computerization provides additional dimensions (such as pressure, airtime, and the drawing sequence) in understanding cognitive impairment for medical practitioners.

The ClockMe System includes two main parts: **The Patient Side** – The ClockReader Application and **the Clinician Side** – The ClockAnalyzer Application. By using the ClockReader application, older adults can conduct the dementia screening test. The ClockAnalyzer application enables medical practitioners to review and monitor their patients' CDT results. Figure 21 shows an overview of the ClockMe System.

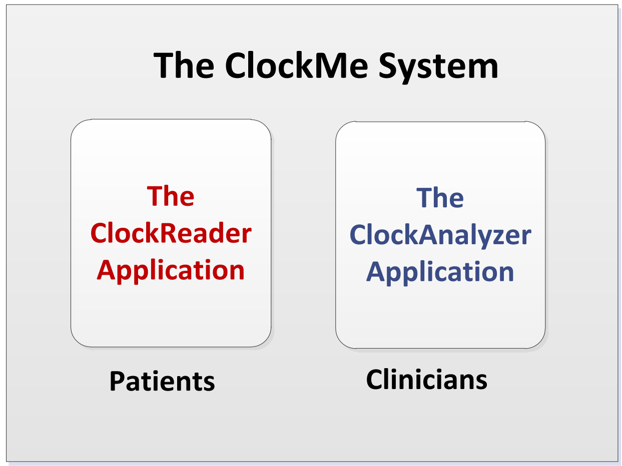


Figure 21. Overview of the ClockMe System

The ClockMe System with the online sketch recognition algorithm is developed in C# programming language and is supported by the "Microsoft Windows XP Tablet PC Edition Software Development Kit 1.7" and "Microsoft Visual Studio 2010." For the first release, the running environment of the program is limited to the Microsoft Windows platform, equivalent to or better than the "Windows 2000 Service Pack 4" with the "Microsoft.Net Framework 3.5 Service Pack 1."

5.2. User Interface Design of the ClockMe System

Each application in the ClockMe System has a different User Interface (UI). It can also run independently as a separate application. However, the ClockAnalyzer Application will share all of the data collected from the ClockReader Application. In the following sub-sections, I describe each application's UI.

5.2.1. User Interface of the ClockReader Application

We developed the ClockReader Application by implementing it on a Tablet PC platform to replace the traditional paper-and-pencil style of testing. The design goal of the ClockReader User Interface was to provide patients with an environment that is similar to paper-and-pencil-based testing.

As can be seen in Figure 22, the UI consists of two main parts: the Drawing area and the Action area. In the Drawing area, similar to the traditional CDT, a user would see a pre-drawn circle in the middle of the Tablet PC screen. The bottom part of the UI, what we call the Action area, includes an automatically generated–ID, one radio button, and two submit buttons. The radio buttons are designed to provide users with an input option such as "Pen." The default setting would be "Pen."

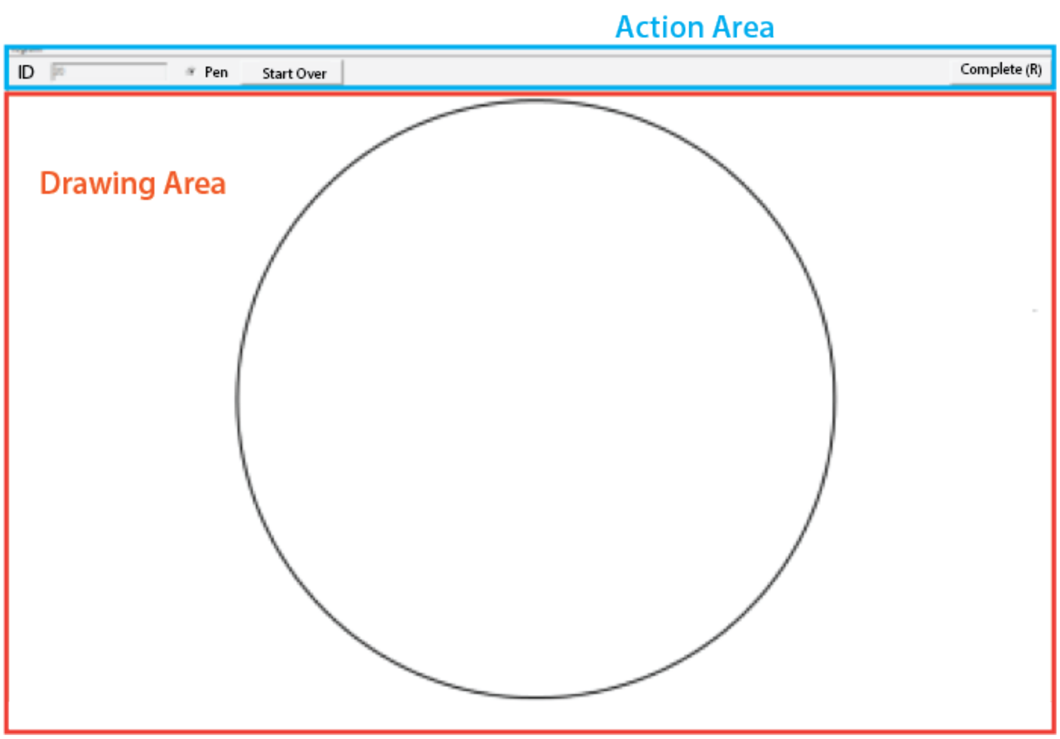


Figure 22. The User Interface of the ClockReader Application

The two submit buttons are “Start Over," and “Complete.” the “Start Over” button would delete all of the strokes in the Drawing area. If users want to redraw a clock, they need to press the “Start Over” button. The “Complete” button would (1) record the ending time of the test; (2) initiate running the sketch recognition algorithm; and (3) send all of the data to a server.

5.2.2. User Interface of the ClockAnalyzer Application

The ClockAnalyzer Application is a tool for medical practitioners. The ClockAnalyzer is designed to support the medical practitioner’s decision-making by performing several different analyses of automated data collection. For example, the ClockAnalyzer will help doctors analyze the data by automatically scoring the criteria, as well as by visualizing graphs from the existing data. The goal of the ClockAnalyzer UI design is to effectively provide a data analysis and an interpretation through multiple visualization methods to medical practitioners.

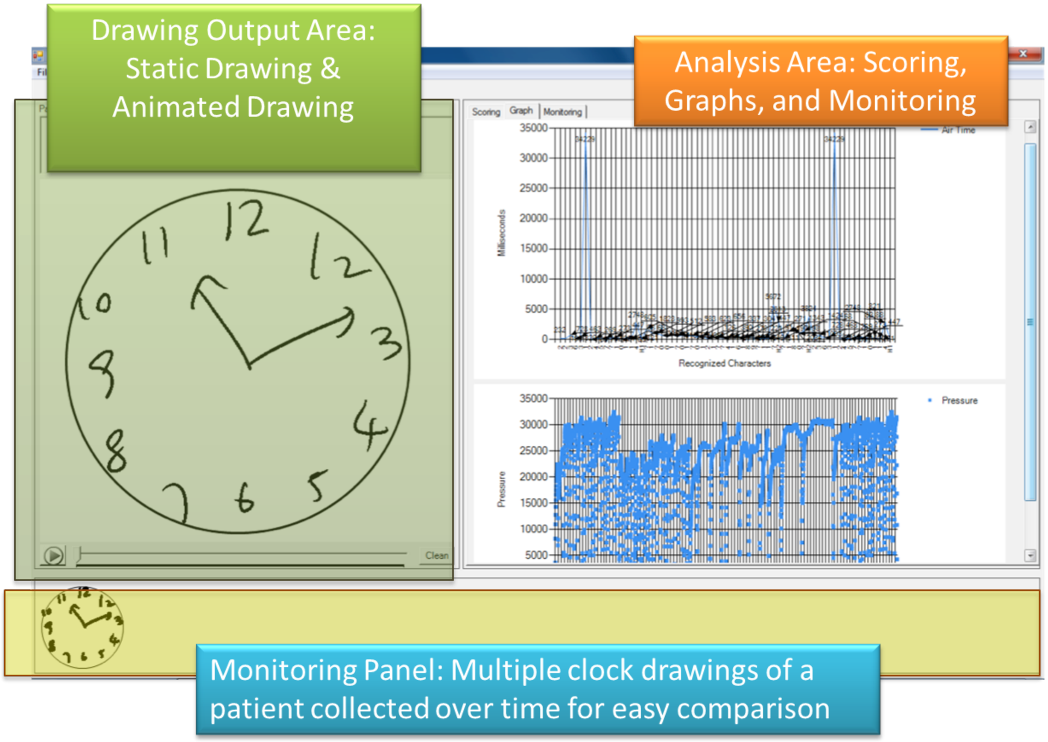


Figure 23. The Default User Interface of the ClockAnalyzer Application

As can be seen in Figure 23 above, the user interface of the ClockAnalyzer Application consists of three main parts: the Drawing Output area, the Analysis area, and the Monitoring panel. The drawing output area is in the top left window. We provide two different drawing outputs: static drawing and active animated drawing. When a user opens a patient's file, the user would see the static clock drawing from the CDT. The static drawing becomes an animation when a user clicks the "Arrow" button on the bottom of the window. The "Arrow" button plays the role of a video playback button. To reset the animation to its initial status, a user needs to click the "Clean" button.

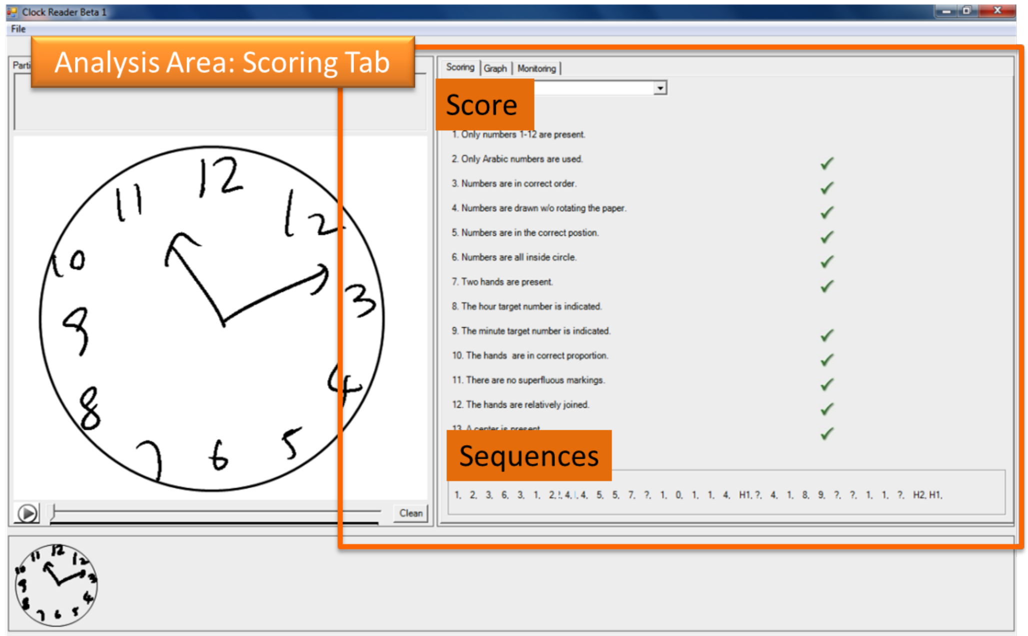


Figure 24. The User Interface of the ClockAnalyzer: Scoring Tab

As can be seen in Figure 24 above, the Analysis area lies in the top right side of the UI. The Analysis area includes three different tabs: Scoring, Graph, and Monitoring. Each tab shows different information. The Scoring Tab shows two types of data: Scores and Sequences. The CDT results are evaluated by Freedman et al.'s 13-point CDT criteria. If a patient's CDT drawing applies each criterion, the green-color check is shown. The Sequences show how a user constructs a clock with a specific sequence. The timestamps of each stroke for each number can help clinicians inspect the patient's drawing strategies. While a patient without dementia could use a strategic way to draw a clock (e.g., 12, 3, 6, 9, or 1, 2, 3, 4), a patient with dementia may exhibit a haphazard manner of arranging the numbers properly. For example, in Figure 24, a user drew a clock by first writing 12, 3, 6, 9, 1, and then filled out the remaining numbers.

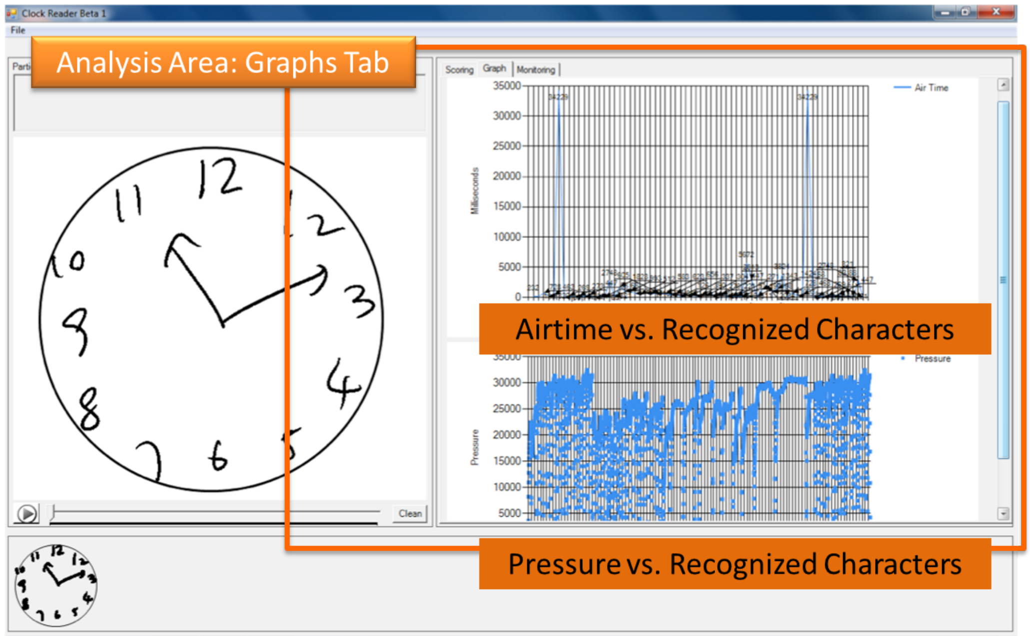


Figure 25. User Interface of the ClockAnalyzer: Graph Tab

The Graph Tab in Figure 25 above shows two different graphs: Millisecond Airtime vs. Recognized Characters and Pressure vs. Recognized Characters. The airtime and pressure data are two novel types of data that the computerized system can provide. Airtime is measured by the sensor to calculate how many milliseconds a pen is up in the air versus a pen on the screen. The pressure is also measured by the pressure sensor embedded in the Tablet PC. For more information about using novel data for clinical purposes, please refer to the next section: Features of the ClockMe System as Automated Novel Data Collection.

The third tab, the Monitoring Tab, relates to the bottom panel. This bottom panel shows the Monitoring Panel, which includes the patient's clock drawing history: Multiple clock drawings from the patient collected over time (e.g., weeks, months or multiple years) for easy comparison. The Monitoring tab will show a graph of historical CDT results, which provide a quick overview of the patient's cognitive condition.

5.3. Features of the ClockMe System: Automated Data-capturing Methods

In this section, I introduce the three automatic data-capturing features of the ClockMe System. Our proposed automatic data-capturing methods are based on one of the leading neuropsychologists, Kaplan's process-oriented approach (17). Many researchers within her school of thought argue that despite the heavy use of drawing tasks in neuropsychological assessment, they are not fully utilized in supporting neuropsychologists' decision-making processes (7, 74, 75, 91). These drawings are predominantly analyzed at the end of the tests as a final product, and the process is ignored (16).Werner (1956) also emphasizes the importance of the process approach by saying, "Every cognitive act involves 'microgenesis.'" Werner also states, "Close observation and careful monitoring of behavior enroute to a solution (process) is more likely to provide more useful information than can be obtained from right or wrong scoring of final products (achievement)" (17).

When I observed the current practice of CDT at Emory ADRC, I identified that the drawing analysis is usually conducted after the completion of a task, and clinicians rarely examine or analyze the process of how the drawings were constructed because of the difficulty in data collection. With the computerized Clock Drawing Test, however, clinicians will be able to derive more data from the CDT, which will help them gain further insights in their data analyses and interpretations.

Feature 1: Automatic Process Capture

The system would capture and analyze a patient's "planning strategy."

We implement the functions of record and playback of the drawing sequence for clinicians to interpret the planning strategies of patients (e.g., 12, 3, 6, 9, or 1, 2, 3, 4). This would provide clinicians with new useful information that was not previously available from paper-and-pencil tests.

For example, the ClockMe System includes a function to save the final output drawing as an Ink Serialized Format (ISF). Whenever a person completes the clock-drawing task, the ISF image is created and will be saved in the database. This would be a helpful resource included in the interface for clinicians because it shows different drawing strategies. Figure 26 shows two clock drawings in an ISF image format. Even though the final output indicates the same time, they are constructed from two different planning strategies.

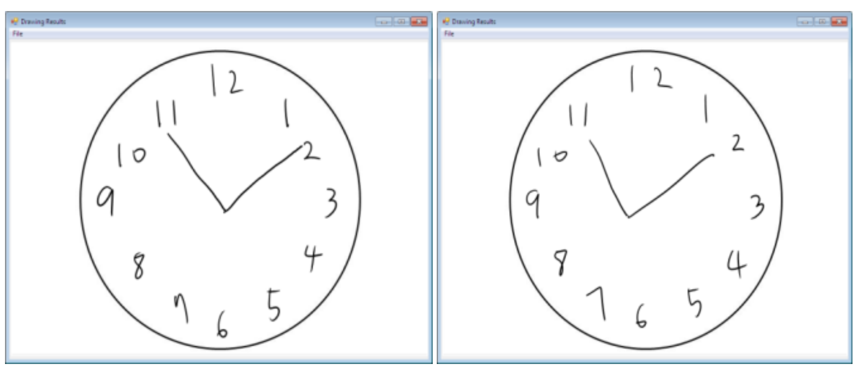


Figure 26. The ISF image of Clock Case 1 (left) and the ISF image of Clock Case 2 (right)

Figures 27 and 28 show the process of two different drawing constructions in Figure 26. Figure 27 represents four steps as to how a clock is constructed in Clock Case 1 in Figure 28. The patient first drew four anchoring numbers, such as 12, 3, 6, and 9, and then filled out the remaining numbers.

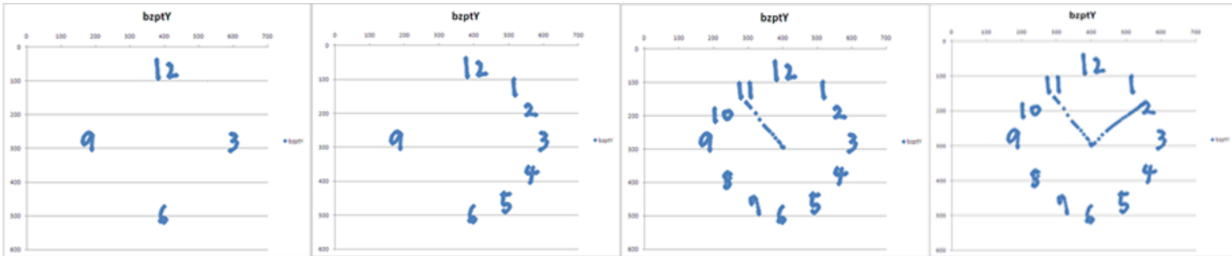


Figure 27. Reconstructed Clock Case 1

The clock drawn in Figure 28 shows that the clock was constructed first with the number 12, and then with the sequential order from 1, 2, 3, and up to 11. Then the patient drew the two clock hands.

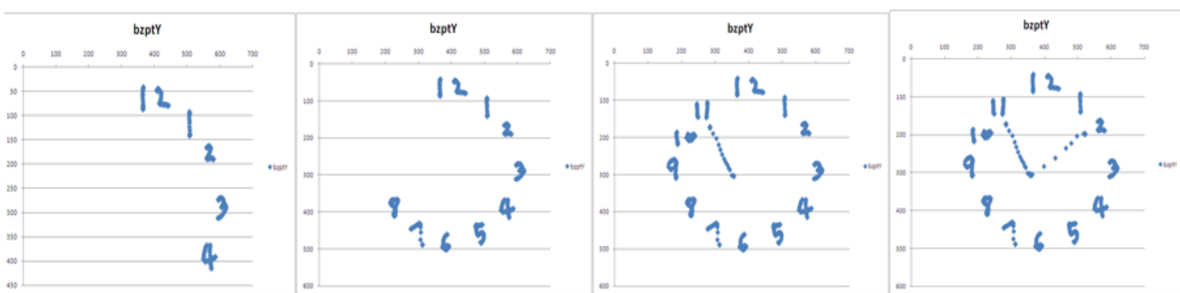


Figure 28. Reconstructed Clock Case 2

Feature 2: Automatic Time Capture

2.1. The system would capture and analyze "airtime," which is the time when the patient is not-drawing – the time of pausing.

When patients draw a clock, at a certain point, they may hesitate to draw, perhaps due to memory problems. Or perhaps they may experience difficulty in recalling a specific number on the clock face. Airtime could be a useful indicator of abnormal unstable cognition for clinicians. For example, if a person has a hard time recalling all of the digits, the airtime graphs would show vertical lines, without much fluctuation. The airtime graph can provide time-related patterns of the patient's drawing. That information can show critical moments of the initiation of drawing and time setting for the clock. None of the existing criteria for the Clock Drawing Test take this factor into consideration. However, airtime could be a good index.

Figure 29 shows an example of an airtime graph, which indicates some patterns of the drawing. The x-axis indicates a stroke order written when a patient draws a clock. The y-axis indicates the airtime of each stroke drawn. The first stroke was written in 4.3 milliseconds. In this case, the 13th and 14th strokes pertain to the clock hand drawings. The 13th stroke took 5 milliseconds to write down. This is the longest time for the patient to think before she actually wrote down something. This implies that the patient needed more time to draw a clock hand in order to set the time. Overall, patients have two critical moments in which they spend quiet time thinking before drawing something: initiation of drawing and setting the time.

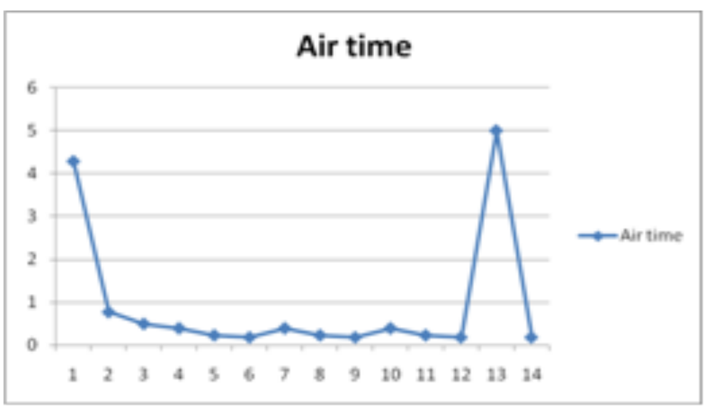


Figure 29. Air time during the Clock Drawing Test

2.2. The system would report "time to completion" of an individual patient's clock drawing test session.

Most psychological assessment measures how long it takes an individual to complete a task. Measuring time while administering a task is usually difficult, especially when the task takes a relatively short time. One of the findings I identified from the observation study in Chapter 3 clearly shows that medical technicians were too busy to prepare and administer the task. They simply did not have enough time to measure how long the patients took to complete the clock drawing task. Despite this current situation, several researchers, such as clinical nurse practitioners and neuropsychologists who attended our focus group meeting, commented that if the system can capture the time to completion, this information would be very useful. They hypothesized that time to completion could help identify the early signs of cognitive impairment. For example, some people may draw a clock correctly; however, they may take a significantly longer time to complete the task in the following year. The completion time will provide opportunities for medical practitioners to test their hypotheses, as well as investigate correlations among other patients' clinical conditions.

Feature 3: Automatic Pressure Capture

The system would measure pressure information during the drawing process.

Studies on the handwriting process (19, 27, 28, 79, 80, 85, 101, 103) demonstrate that a kinematic analysis of handwriting provides important information about the handwriting process among older people (40). Their experimental results show that being a higher-age adult was consistently associated with longer on-paper and in airtime, as well as it was associated with lower speed and lower pressure (81). Furthermore, analyzing pressure data in the drawing may help clinicians advance their understanding of the patient's drawing and relate such data to other diagnoses.

5.4. Preliminary Study

This section provides the results from a preliminary study. The purpose of this preliminary study was, prior to the deployment study, to understand how accurately the system recognizes a patient's drawing and to assess the usability of the Tablet PC-based ClockReader System.

5.4.1. Participants

We recruited the same 24 subjects who had participated in our formative study from the Honor Research Registry in Clinical Research in Neurology (CRIN) at the Emory Alzheimer's Disease Research Center in Atlanta. The participants were 11 females and 13 males. Even though their average age was 75 years, the groups included 10 individuals who were older than 78 years. The oldest participant was a 98-year-old, and the youngest participant was a 59-year-old. Their educational levels were diverse. Six participants had high school diplomas. Nine participants were college graduates. The last nine participants had a graduate-level education. Their computer literacy levels were also mixed. Seven participants had never used a computer before. On the other hand, some of them regularly used computers as a daily routine. For more detailed information about the participants, please see Section 3.4 in Chapter 3.

5.4.2. Sketch Recognition Accuracy

Sketch recognition is an integral part of designing an automatic Clock Drawing Test. Here is the question we tried to answer from our preliminary study: "How accurate is the sketch recognition with our filtering algorithm?"

We asked our study participants to draw a clock in order to understand the handwritten digit accuracy rates of the system. The entire process is captured by the ClockReader System, as well as by the Camtasia Studio desktop recording program. The external recording program would help us investigate several unexpected errors. With the movie generated by the program, we would be able to observe the in-depth context of an individual patient's drawing activity.

Overall, the average recognition accuracy rate was 84 %. This was calculated by determining what percentage of the digits was correctly recognized out of a possible total number of 15 digits. Our digit recognition algorithm detects each digit individually within two-digit numbers, such as 10, 11, and 12. Thus, the total digits are 15 instead of 12. The most misrecognized numbers are either 7 or 8. They were recognized as 1 or 0. Figure 30 shows the results of recognition accuracy with the 20 subjects. During the study, the system accidentally crashed four times, which prevented us from collecting the recognition data.

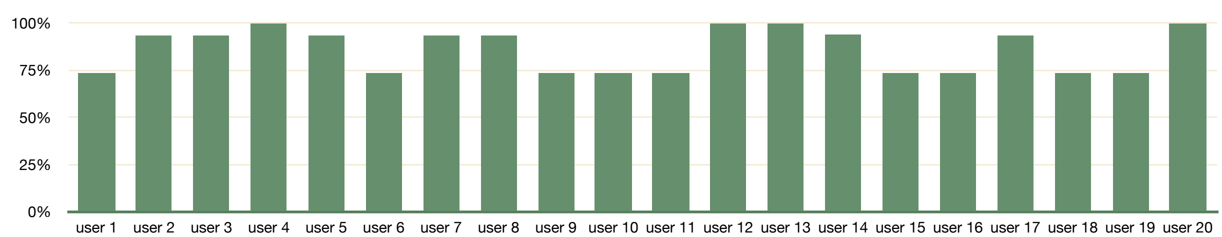


Figure 30. Results of Recognition Accuracy

We added additional filtering to track the order of strokes in each digit so as to find out how many times the stroke changes its direction and toward which direction the stroke changes. For example, number 8 is often misrecognized as number 0 due to the similarity in the shapes of both numbers. The system can algorithmically identify the two numbers by comparing the number of direction changes between those two numbers. We can see that the number 8 has 4 curved segments that are convex to the left, right, left, and right, and the number 0 has 2 curved segments that are convex to the left and right in the order of sequence. These could be unique signatures for each number. We also implemented this additional algorithm for the numbers 1, 7, 4, and 9, which are often misrecognized as other numbers, in order to reduce the possibility of misrecognition.

5.4.3. Usability Comparison

. To deploy a computerized Clock Drawing Test in a real clinical environment, the system should be usable for patients, especially those who are not familiar with using a computer. Here is the question we tried to answer from our preliminary study: "When a patient uses our Tablet PC-based ClockReader System, does the system yield the same results as the traditional paper-and-pencil-based Clock Drawing Test?"

We conducted a usability comparison study with 24 older adult volunteers. The usability test consisted of 1) a computer familiarity survey; 2) two drawing tasks: one using the ClockReader and the other using a pencil and paper; and 3) a post-interview.

The computer familiarity survey results of computer familiarity showed that all participants considered themselves as having low- to medium-level computer literacy. None of them considered themselves as having high computer literacy, which is not surprising, considering the participants' average ages. Participants were seniors who did not use computers during their working years. Personal computers were not popular until most of them had entered retirement.

However, unlike their self-reported computer literacy, all participants successfully completed the Clock Drawing Test using the ClockReader application. In the post-interview, we also asked them whether using the stylus to draw a clock on the surface of the computer was difficult. None of the participants expressed difficulties in using a stylus. Also, from the results of the two different drawings that each participant did, it was evident that there were no critical difficulties for older adults in using a Tablet PC. Overall, the usability testing results of the Tablet PC-based ClockReader System were positive. We conclude that the computerized Clock Drawing Test is acceptable for seniors who do not have experience with computers. Figure 31 shows how a user interacts with the system.

Figure 31. A picture showing how a patient draws a clock on the Tablet PC

We asked two medical technicians to evaluate the drawings to understand whether differences existed, both between different evaluators and between the two different drawings per individual patient. Figure 32 shows the percentage graph of disagreement between two evaluators per each criterion. Table 8 shows the frequency of disagreement per each criterion.

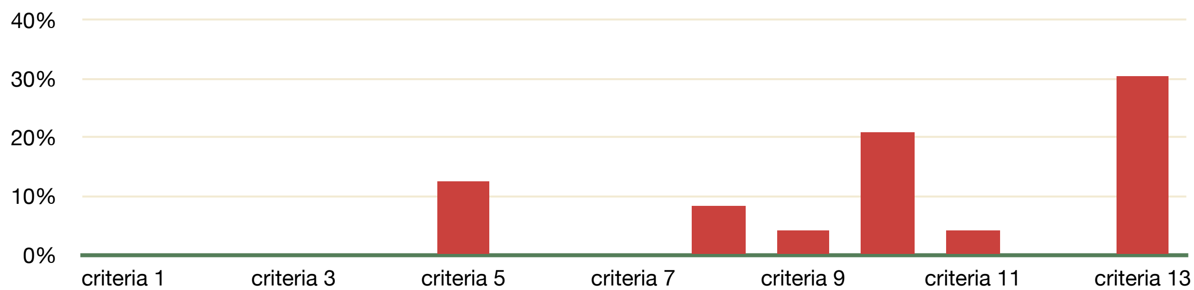


Figure 32. Percentage of Disagreement between Evaluator 1 and 2 per each Criterion

Table 9. Frequency of disagreement of each criterion of the Clock Drawing Test

Criteria	Frequency of disagreement
1. 1~ 12 are present	0
2. Only digits are present	0
3. Correct order	0
4. No rotation	0
5. Correct position	3
6. Inside circle	0
7. Two hands are present	0
8. Hour indicated correctly	2
9. Minute indicated correctly	1

Table 9 continued

10. Correct proportions of the hands	5
11. No superfluous markings	1
12. The hands are relatively joined	0
13. A center is present	7

As can be seen in the data from both Table 9 and Figure 32, the highest disagreement is with respect to the criterion about a center. Most disagreements are due to interpretations of the depiction of time (criteria 7 to 12: four out of six) compared to the criteria with respect to the numbers (criteria 1 to 6: one out of six).

Now, let's carefully examine two drawing sets from four different users to explore how different the evaluators analyzed them.

Figure 33. Two clock drawings from a 92-year-old male participant: A paper-and-pencil drawing (left) and a drawing using a stylus on the Tablet PC (right)

First, Figure 33 above shows two clock drawings from a 92-year-old male participant. Both drawings are almost identical. The only difference shows the shakiness of the clock hands from a drawing using the ClockReader application. In fact, the participant has a tremor. Since there is less friction on the surface of the computer, a participant's drawing using the application can apparently show tremor problems more easily. The paper-and-pencil drawing, on the other hand, does not show any signs of tremor. This shows the potential use of a computerized sketch-analysis to identify the early stages of Parkinson's disease.

Table 10. Comparison of the total score and the criteria evaluated differently between two evaluators: A 92-year-old male participant case in Figure 33

	A paper-and-pencil drawing	A tablet-and-stylus drawing
Evaluator 1	Total score = 12 | C9	Total score = 12 | C9
Evaluator 2	Total score = 12 | C9	Total score = 12 | C9

Table 10 summarizes how two different evaluators evaluated the drawings of a 92-year-old male participant. Two evaluators unanimously scored the four drawings as 12 points out of 13. Since the time should have been set at 11: 10 instead of 10: 50, criterion #9 was not satisfied. In this one example, there was perfect agreement between the two raters on the nature of the errors in the two types of drawings.

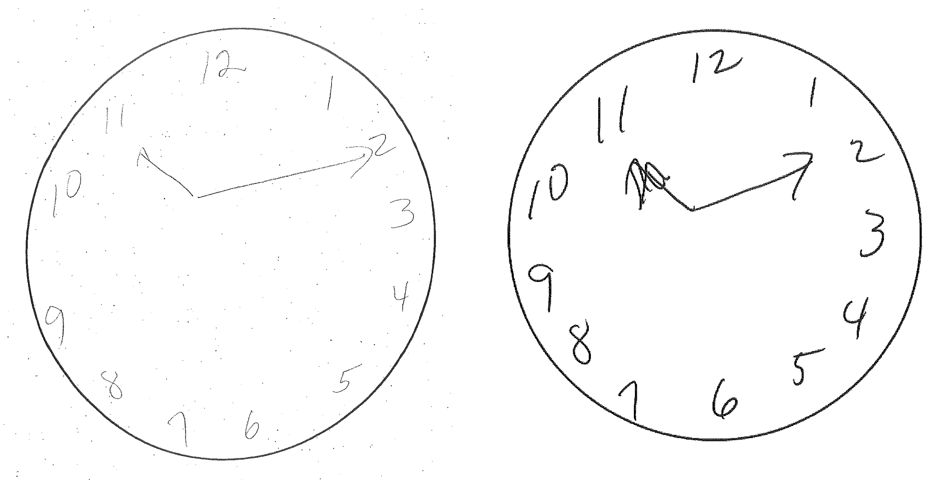


Figure 34. Two clock drawings from a 76-year-old female participant: A paper-and-pencil drawing (left) and a drawing using a stylus on the Tablet PC (right)

Figure 34 above shows two clock drawings from a 76-year-old female participant. The two drawings are almost identical. The only difference is in the superfluous lines of the clock hand indicating time in the drawing from the ClockReader application. In fact, the participant did not use an eraser, although we had instructed her about the Eraser button. Since this difference is due to the application's function usage, we consider this difference not to be critical in revealing anything about the usability of the ClockReader.

Table 11. Comparison of the total score and the criteria evaluated differently between two evaluators: A 76-year-old female participant case in Figure 34

	A paper-and-pencil drawing	A tablet-and-stylus drawing
Evaluator 1	Total score = 13	Total score = 12 | C11
Evaluator 2	Total score = 11 | C5, C13	Total score = 11 | C11, C13

Table 11 shows how two different evaluators analyzed the drawings of the 76-year-old female participant in Figure 34. With respect to the paper-and-pencil drawing, Evaluator 1 scored the drawing as 13-points out of 13, which means that the drawing did not show any cognitive impairment. However, Evaluator 2 scored the same drawing as 11-points out of 13. Criterion # 5 (the correct position of the numbers) and criterion #13 (the presence of a center) were not satisfied. The differences are the result of subjectivity of interpretation in terms of what an error is. For example, Evaluator 2 believed that the digit 9 was not drawn in the correct position, since it was drawn a little downward from where it normally is. However, Evaluator 1 thought that all of the numbers were located in the correct position. With respect to the tablet-and-stylus drawing, both evaluators commented on criterion 11, since the superfluous marking existed at the hand, indicating time. In addition, Evaluator 2 indicated that both drawings did not have a center. On the other hand, Evaluator 1 believed that both drawings did have a center, which eventually resulted in Evaluator 2's total score being higher than Evaluator 1's total score. This example clearly shows that differences exist due to human evaluator subjectivity. Automatic scoring will improve the consistency of scoring by minimizing this kind of inter-rater subjectivity.

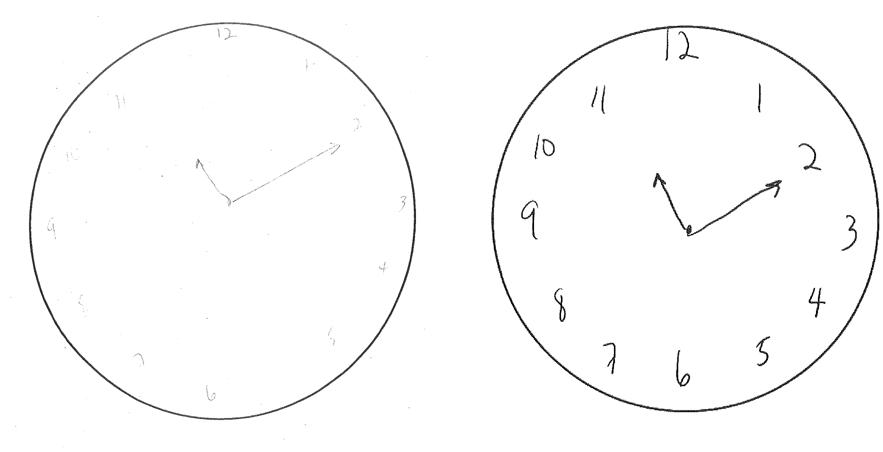


Figure 35. Two clock drawings from a 62-year-old female participant: A paper-and-pencil drawing (left) and a drawing using a stylus on the Tablet PC (right)

Figure 35 above shows two clock drawings from a 62-year-old female participant. The paper-and-pencil drawing on the left is exactly the same as the tablet-and-stylus drawing on the right.

Table 12. Comparison of the total score and the criteria evaluated differently between two evaluators: A 62-year-old female participant case in Figure 35

	A paper-and-pencil drawing	A tablet-and-stylus drawing
Evaluator 1	Total score = 13	Total score = 13
Evaluator 2	Total score = 13	Total score = 13

Table 12 summarizes how two different evaluators evaluated the drawings of a 62-year-old female participant. Two evaluators unanimously scored the four drawings as 13 points out of 13. This clearly shows that there is no difference between a paper-and-pencil-drawing and a drawing from the ClockReader application, as well as between different evaluators.

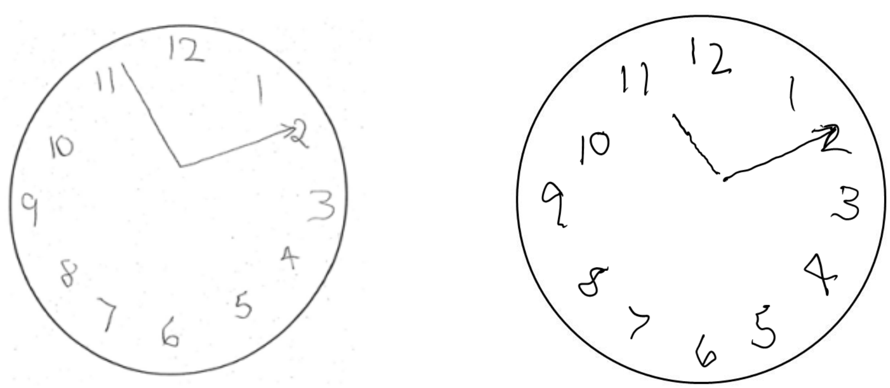


Figure 36. Two clock drawings from a 76-year-old female participant: A paper-and-pencil drawing (left) and a drawing using a stylus on the Tablet PC (right)

Lastly, Figure 36 above shows two clock drawings from a 76-year-old female participant. Both drawings are almost identical. The only difference shows the length of the hand indicating the hour. The paper-and-pencil drawing on the left shows that the hand indicating the hour is almost of similar proportion with the minute hand. Interestingly, the way the participant indicates only minutes with arrows is the same in both drawings.

Table 13. Comparison of the total score and the criteria evaluated differently between two evaluators: A 76-year-old female participant case in Figure 36

	A paper-and-pencil drawing	A tablet-and-stylus drawing
Evaluator 1	Total score = 12 | C10	Total score = 13
Evaluator 2	Total score = 11 | C10, C13	Total score = 12 | C13

Table 13 summarizes how two different evaluators evaluated the drawings of a 76-year-old female participant. With respect to the paper-and-pencil drawing, Evaluator 1 gave 12 points out of 13. On the other hand, Evaluator 2 assigned 11 points out of 13. The difference lies in how the evaluators considered whether or not a center existed. With respect to the tablet-and-stylus drawing, Evaluator 1 gave 13 points out of 13. On the other hand, Evaluator 2 gave 12 points out of 13 because criterion 13 was not satisfied. Criterion 13 showed the greatest number of inter-rater disagreement. The evaluation instructions advised evaluators to analyze a clock by observing whether a center is present (drawn or inferred) by the joining of the clock hands. However, if an evaluator infers the presence of the center, it is difficult to reach objectivity or minimization of subjectivity. We can also see the value of computerized automatic scoring to improve upon this ambiguous interpretation of a “center.”

Looking at the results of 20 participants and 40 clock drawings, a paired-sample t-test revealed significant differences in medical technicians’ assessments of patients’ drawings. Specifically, drawings were judged differently in terms of whether the hands were in correct proportions (criterion #10; $t(23) = 2.46, p < .05$) and whether a center was present at the joining of the clock hands (criterion #13; $t(23) = 3.10, p < .01$). Also, there was a trend for overall difference between technicians for Criterion #5, the numbers being in the correct position ($t(23) = 1.81, p < .09$). Furthermore, the final scores for each patient were also significantly different between the two technicians ($t(23) = 3.47, p < .01$), indicating that the same patient’s clock drawings can be differentially scored across technicians.

CHAPTER 6

EVALUATION

6.1. Overview

I deployed and evaluated the ClockMe System at the Emory Alzheimer's Disease Research Center (ADRC) over five months. The evaluation focused on whether the system was appropriately designed for patients, especially older adults for conducting a screening test; and whether it is appropriately designed to support dementia screening in clinical use.

The procedures of the study were threefold: to conduct the user study with patients, to conduct the user study with medical technicians and medical practitioners, and to analyze the computer-collected data. In each sub-section below, I specifically stated how the specific user group used the system.

6.2. The Evaluation of the ClockReader Application

The ClockReader is an application that provides patients with a Tablet PC-based format for a clock-drawing task. The application also records and saves each patient's drawing data for future clinical use. Now, I describe the user evaluations of the ClockReader Application by the three different user groups: patients, medical technicians, and medical practitioners.

6.2.1. The Use of the ClockReader by Patients

The evaluation focused on (1) how patients perceived the use of a computerized assessment in terms of ease-of-use; and (2) whether the drawings from a computer method could be different from the traditional paper-and-pencil method.

6.2.1.1. Participants

I recruited 45 volunteers enrolled in Emory ADRC's Honor Research Registry. Honor subjects include older adults without any cognitive impairment, individuals with MCI and individuals with dementia, including both Alzheimer's disease as well as other dementias, including Frontotemporal dementia, Parkinson's disease with dementia, and Lewy Body Dementia.

Among the 45 participants, there were 20 males and 24 females. Their average age was 71-years-old, and the range was diverse. The youngest was a 30-year-old male who was concerned about dementia due to his family history. The oldest was a healthy 96-year-old female. Sixty percent of the participants were older than 70 years old. The participants had somewhat higher educational levels, since 80% of the participants graduated college: only seven participants had a high school diploma; 15 participants were college graduates; 21 participants had a graduate-level education. Two participants could not complete their surveys due to a lack of time or other personal reasons. Therefore, there was some missing information. Table 14 shows the number of participants by gender, age, and education. The following Figures 37, 38, and 39 also represent the number of participants as bar graphs.

Table 14. The Number of Participants by Gender, Age, and Education

Gender	Number of Participants
Male	20
Female	24
Unknown	1
Total	**45**

Age	Number of Participants
30~ 39	1
40~49	0
50~59	6
60~69	10
70~79	13
80~89	13
90~99	1
Unknown	1
Total	**45**

Education	Number of Participants
High School	7
College	15
Graduate	21
Unknown	2
Total	**45**

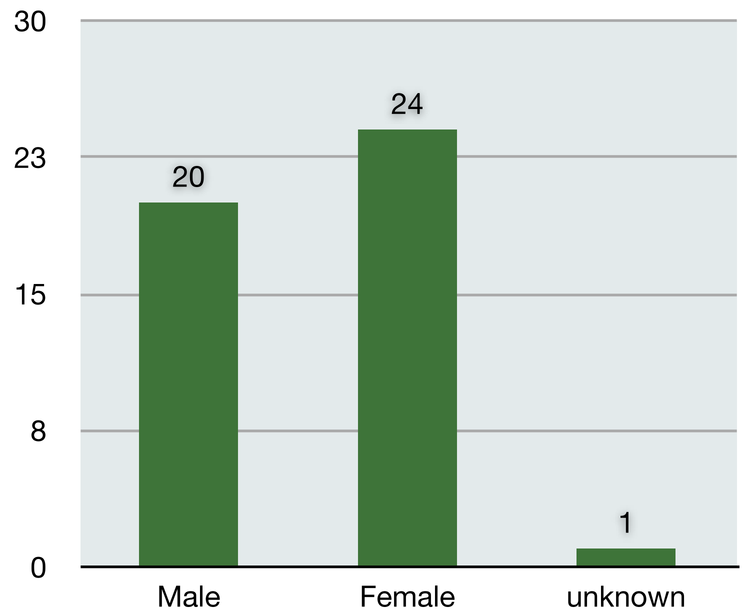


Figure 37. The Number of Participants by Gender

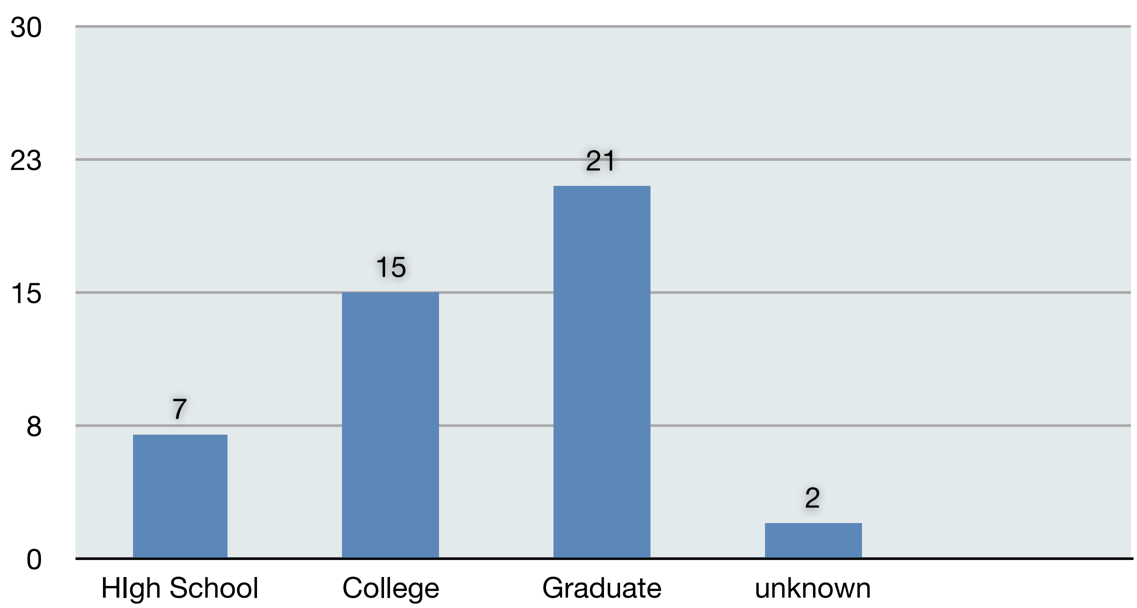


Figure 38. The Number of Participants by Education

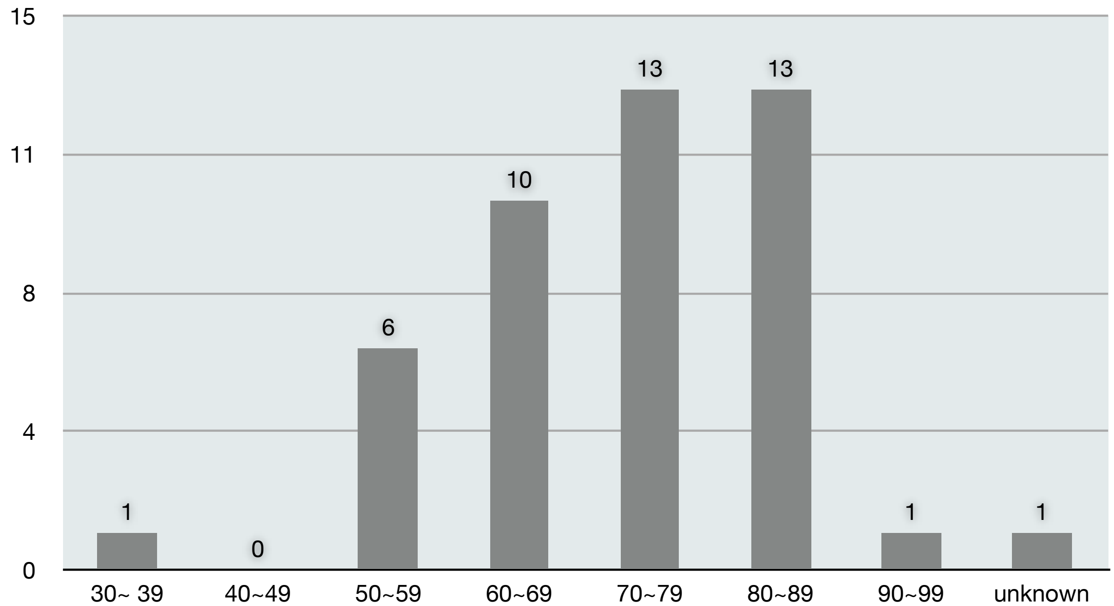


Figure 39. The Number of Participants by Age

6.2.1.2. Procedures

First, medical technicians asked research volunteers whether they were interested in joining the ClockMe study. If they agreed to participate, the technicians had them read and sign the IRB consent form. Then, the technicians asked the participants to complete the background survey form, which included questions about computer experience, as well as their demographic information. One group of participants completed a paper-based CDT first and then performed a computerized CDT. The other group of participants conducted a computerized CDT first and then completed a paper-based CDT. Once they completed the task, the technicians conducted a post-test interview.

6.2.1.3. Results

A. How did patients perceive the use of a computerized assessment in terms of ease-of-use? And did their computer experience influence the use of the ClockReader System?

All 45 people were successfully able to complete the Clock Drawing Test in a computerized format, as well as the traditional paper-and-pencil format. None of them had difficulties or complained about the use of a stylus with their limited- and/or no-computer experience.

Figure 40 shows the paper-clock drawing, as well as the computer-clock drawing from Patient ID # 71. She was the oldest participant, a 96-year-old female. She had never used a computer in her life and had no idea about Tablet Computing. However, she was happy to draw a clock on the computer, and there were no differences between the two different methods for her clock drawings.

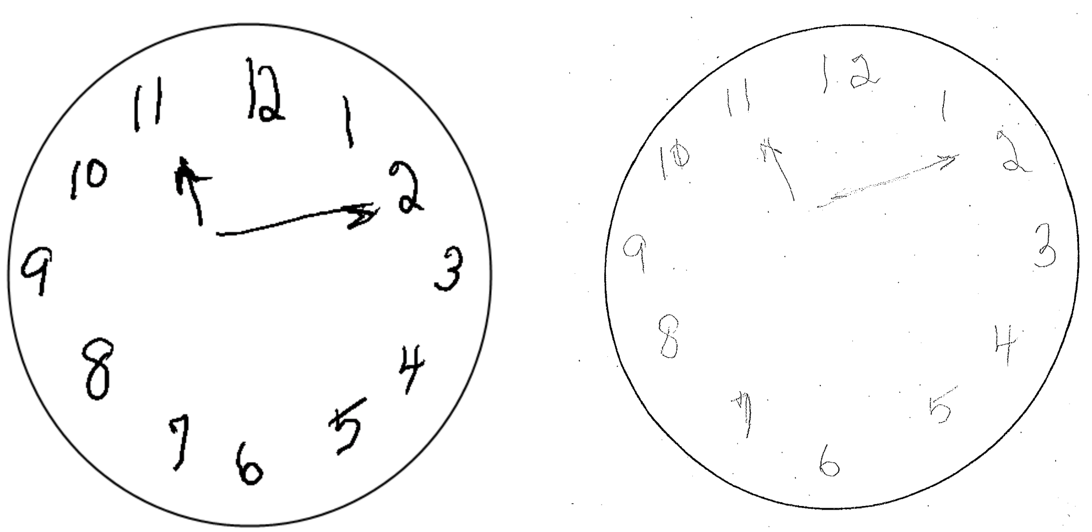


Figure 40. Two clock drawings from a 96-year-old female participant: A drawing using a stylus on the Tablet PC (left) and a paper-and-pencil drawing (right)

Recalling that the participants had an average age of 70 years old and had somewhat higher education levels, they had interestingly diverse computer usage experience. The following Table 15 reports the results of computer experience by the number of participants.

Table 15. Results of the Computer Experience Survey

6. Are you familiar with using a computer?	Number of Participants
Yes	39
No	4
Unknown	2
Total	45
7. Computer Literacy	**Number of Participants**
High	5
Medium	19
Low	18
Unknown	3
Total	45
8. What is the most uncomfortable thing about using a computer?	**Number of Participants**
Reading text on the screen	7
Using a mouse	8
Typing on a keyboard	23
Unknown	27
Total	45

Table 15 continued

9.	What would you prefer if you could use a computer differently? Describe it. (For example, I would like to use a computer with voice-activated control)
Touch screen	3
Voice-Activated Control	15
Unknown	21
Does not want to change	6
Total	45

10. What are your most common activities using your computer?	Number of Participants
Document/Word Processing	3
Internet	36
Games	1
No Comments	5
Total	45

In brief, a majority of the participants (N = 39) were familiar with using a computer. However, when I asked how they rated their computer literacy, 19 out of 45 participants answered medium. 18 out of 45 responded low. Only five participants expressed their computer literacy as high. The major barriers they perceived in using computers consisted of typing on a keyboard (N= 23), mouse use (N= 8), and reading text

(N=7) on the screen. Their most common activity using the computer was Internet use (N= 36).

Furthermore, some patients expressed their preference using the computerized Clock Drawing Test. One patient reported that it was easier for her to draw something on the computer because there was less friction compared to using a pencil on paper. Since older adults tend to have less manual strength, she expressed her liking and asked me where she could buy this "toy" for her personal use. She said, "I used to love drawing when I was young, and this Tablet computer and digital pen reminds me of my childhood. If it's not too expensive, I'd love to buy this and would want to draw with it every day." The other patient also stated that he preferred to draw a clock with the ClockReader because he could start all over. Unlike the paper-and-pencil Clock Drawing Test, the START OVER button provides patients with the option to easily erase everything. By simply clicking a button, they can have a clean slate of the drawing test.

The first design of the ClockReader System included the ERASER button, which enabled patients to use the stylus like a real eraser. Thus, a patient needed to rub the stylus in order to erase the drawing. Based on my observation of the preliminary study, most of them did not use the function, and even two people who had used the function reported that it was not easy to use. As a result, we decided to remove the ERASER button, and instead, we embedded the START OVER button.

B. With respect to an individual patient, were the drawings from the computer method different from the traditional paper-and-pencil method? In other words, did the patients draw the same clock from both methods?

To investigate the differences among the computer-clock-drawings, as well as the paper-clock-drawings, I asked a medical technician at Emory ADRC to evaluate both drawings per individual patient. Before I gave them the drawings, I removed all identifying information of the drawings so the technician did not know the origin of the drawing method, as well as the identification of an individual patient.

According to the data analysis, there was statistically no significant difference in terms of scoring results for all thirteen criteria when the same medical technician evaluated both computer-based and paper-based clock drawings completed by the same patients (ps> .05).

In brief, patients using the ClockReader created the same drawings as the paper-and-pencil clock drawings. Table 16 below shows the raw data I collected for the analysis. ID P1 to ID P45 indicated computer-clock-drawings. ID 46 to ID 89 indicated paper-clock-drawings. I collected two sets of clock drawings from 45 patients. A total of 89 drawings were collected instead of 90 drawings because one patient's paper-clock drawing data was missing.

Table 16. CDT score by Participants

ID	**1.** all 12 numbers (only 12)	**2.** only arabic numbers	**3.** correct order	**4.** drawn without rotating paper	**5.** close to quadrants	**6.** inside circle	**7.** two hands present	**8.** hour correct	**9.** minute correct	**10.** correct proportion	**11.** no superfulou s marks	**12.** hands joined	**13.** center	total
P1	1	1	1	1	1	1	1	1	0	0	1	1	0	10
P2	1	1	1	1	1	0	1	1	1	1	1	1	1	12
P3	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P4	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P5	0	1	1	1	0	1	1	0	0	0	1	1	1	8
P6	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P7	1	1	1	1	1	1	1	1	1	1	0	1	0	11
P8	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P9	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P10	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P11	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P12	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P13	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P14	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P15	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P16	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P17	1	1	1	1	1	1	1	1	0	0	1	1	1	11
P18	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P19	1	1	1	1	1	1	0	1	1	0	0	1	1	10
P20	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P21	1	1	1	1	1	1	1	0	1	1	1	1	1	12
P22	1	1	1	1	0	1	1	1	1	1	1	1	1	12
P23	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P24	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P25	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P26	0	0	0	0	0	0	0	0	0	0	0	0	1	1
P27	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P28	1	1	1	1	1	1	1	1	1	1	0	1	1	12
P29	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P30	1	1	1	1	1	0	1	1	0	1	1	1	1	11
P31	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P32	0	1	1	1	0	0	1	1	1	1	0	1	1	9
P33	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P34	0	1	1	1	0	0	1	1	1	1	1	1	1	10
P35	0	1	1	1	1	1	1	1	1	1	1	1	1	12
P36	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P37	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P38	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P39	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P40	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P41	1	1	1	0	1	1	1	1	1	0	1	1	1	11
P42	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P43	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P44	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P45	1	1	1	1	1	1	1	1	0	0	1	1	1	11
P46	1	1	1	1	0	1	1	1	0	0	1	1	0	9
P47	1	1	1	1	1	0	1	1	1	1	1	1	1	12
P48	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P49	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P50	0	1	0	1	0	1	1	1	1	0	0	1	1	8
P51	1	1	1	1	1	1	1	1	1	0	0	1	1	11
P52	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P53	1	1	1	1	1	1	1	1	1	0	0	1	1	11
P54	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P55	1	1	1	0	1	1	1	1	1	0	0	1	1	10
P56	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P57	1	1	1	1	1	1	1	1	1	0	0	1	1	11
P58	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P59	1	1	1	1	1	1	1	1	1	1	0	1	1	12
P60	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P61	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P62	1	1	1	1	1	1	1	1	0	0	1	1	1	11
P63	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P64	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P65	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P66	0	1	1	1	1	0	1	1	1	1	0	1	1	10
P67	1	1	1	1	0	1	1	1	1	1	1	1	0	11
P68	1	1	1	1	1	1	1	1	1	1	0	1	1	12
P69	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P70	1	1	1	1	1	1	1	1	1	1	1	1	0	12
P71	0	1	1	1	1	1	1	1	1	0	1	1	1	11
P72	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P73	1	1	1	1	1	1	1	1	1	1	0	1	1	12
P74	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P75	1	1	1	1	1	0	1	1	1	1	1	1	1	12
P76	1	1	1	1	1	1	1	1	1	1	1	0	0	11
P77	0	1	1	1	1	0	1	1	1	1	0	1	1	10
P78	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P79	1	1	1	1	1	0	1	1	1	0	1	1	1	11
P80	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P81	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P82	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P83	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P84	1	1	1	1	1	0	1	1	1	1	1	1	0	11
P85	1	1	1	1	1	1	1	1	1	1	1	1	1	13
P86	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P87	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P88	1	1	1	1	1	1	1	1	1	0	1	1	1	12
P89	1	1	1	1	1	1	1	1	0	0	1	1	1	11

6.2.2. The Use of the ClockReader by Medical Technicians and Medical Practitioners

Medical technicians and medical practitioners use the ClockReader System to administer the Clock Drawing Test. The System also allows them to electronically collect a patient's test data. The evaluation focused on how medical technicians and medical practitioners perceived the use of a computerized assessment. Through phone- and in-person interviews, I also investigated room for improvement.

6.2.2.1. Participants

I recruited three medical technicians who currently administer the paper-and-pencil Clock Drawing Test (CDT) at Emory ADRC. They evaluated whether the computerized CDT, the ClockReader application, was useful and designed for ease-of-use in their practice. To understand the characteristics of the users, I collected their demographic information in a pre-survey.

Two females and one male were recruited. All of them were under 25 years old and were trained as medical technicians. Their academic backgrounds were diverse. The male studied psychology in his undergraduate study, and one female earned a bachelor's degree in writing and minored biology. They both had only six-months of working experience in terms of the CDT administration. The other female, working as a lead research specialist at Emory ADRC, held a Master's degree in Neuroscience. She worked to administer the CDT for approximately one-and-a-half years. None of them had studied technological fields.

They all reported that they used computers multiple times per day in their work. Unlike their computer use, they stated that they were not accustomed to using Tablet PC or pen-based computing. Two of them stated that they had no experience in using Tablet PC. The other participant reported that he had some experience, but he never used it for practical purposes.

The below Table 17 summarizes the participants' demographic information, as well as their computer experience. Please see Appendix B for the survey form I used.

Table 17. Survey Results of Medical Technicians

	Technician A	Technician B	Technician C
AGE	Under 25	Under 25	Under 25
EDUCATION	BA in Writing	BA in Psychology	MA in Neuroscience
WORKING EXPERIENCE	0.5 year	0.5 year	1.5 year
COMPUTER USE	Multiple times a day	Multiple times a day	Multiple times a day
FAMILIARITY of PEN COMPUTING	No experience	Some experience, but no practical use	No experience

I also completed a user study of the ClockReader with Emory ADRC's medical practitioners. Please refer to their demographic information under Section 6.3.1. Participants.

6.2.2.2. Procedures

I had a one-hour demo and training session for the medical technicians in late February 2012 prior to the ClockReader deployment in their work. First, I demonstrated how to use the system and then how to store each patient's drawing data. After my demonstration, each technician individually used the system. Then, we had a Q-and-A session before wrapping up the session. They learned how to use the ClockReader, as well as to store each patient's drawing data. Please see Appendix C for the user manual I authored for their training.

At the end of the training session, all technicians completed the System Usability Scale (SUS) questionnaire developed by John Brooke in 1986(10, 99). In addition to these questionnaires, during the system deployment time, I conducted monthly in-person interviews and weekly phone interviews. I conducted weekly phone interviews only for the first month during the deployment period to ascertain their initial use.
Moreover, I also provided them with field notes, which they could freely use to report their opinions while using the system. These efforts helped me capture their system use experience at every moment. In particular, I liked to know whether they would change their opinion of using the system while they were longitudinally using it. In brief, the way I collected data was via a post-survey, weekly phone interviewing, in-person interviews, and field notes. Figure 41 shows a summary of the procedures of the user study with the technicians.

Figure 41. Summary of the Procedures

6.2.2.3. Results

The results of the System Usability Scale (SUS) Questionnaire showed that the ClockReader System was easy to use and was appropriately designed to use in the technicians' and medical practitioners' work. Tables 18 and 19 show the results of the SUS Questionnaire[2]per three individual technicians, as well as three individual medical practitioners. The three technicians scored it 100 points, 100 points, and 93 points, respectively. The three medical practitioners unanimously scored it 100 points. If the score is higher than 60 points in SUS score results, the system is considered as well designed. The SUS scores have a range from 0 to 100.

Table 18. System Usability Scale for the ClockReader Application Use by Medical Technicians

	Technician A	Technician B	Technician C
1. I think that I would like to use this system frequently.	5	5	5
2. I found this system unnecessarily complex.	1	1	1
3. I thought the system was easy to use.	5	5	5
4. I think I would need the support of a technical person to be able to use this system.	1	1	1
5. I found the various functions in this system were well integrated.	5	5	3
6. I thought this system was too inconsistent.	1	1	1
7. I would imagine that most people would learn to use this system very quickly.	5	5	4
8. I found the system very cumbersome to use.	1	1	1

[2]A Likert scale was used to indicate the degree of agreement or disagreement with each statement on a 5-point scale. For example, 5-points indicates "strongly agree," while 1-point indicates "strongly disagree." To calculate the SUS score, the score contribution is used; each item's score contribution ranges from 1 to 4. For items, 1, 3, 5, 7, and 9, the scale position minus 1 is the score contribution. For items 2, 4, 6, 9, and 10, the contribution is 5 minus the scale position. Then to obtain the overall SUS score, multiply the sum of the score by 2.5 (99).

Table 18 continued

9. I felt very confident using the system.	5	5	5
10. I needed to learn a lot of things before I could get going with this system.	1	1	2
Total Score	40	40	36
SUS Score	100	100	93

Table 19. System Usability Scale for the ClockReader Application Use by Medical Practitioners

	Medical Practitioner A	**Medical Practitioner B**	**Medical Practitioner C**
1. I think that I would like to use this system frequently.	5	5	5
2. I found this system unnecessarily complex.	1	1	1
3. I thought the system was easy to use.	5	5	5
4. I think I would need the support of a technical person to be able to use this system.	1	1	1
5. I found the various functions in this system were well integrated.	5	5	5
6. I thought this system was too inconsistent.	1	1	1
7. I would imagine that most people would learn to use this system very quickly.	5	5	5
8. I found the system very cumbersome to use.	1	1	1
9. I felt very confident using the system.	5	5	5
10. I needed to learn a lot of things before I could get going with this system.	1	1	1
Total Score	40	40	40
SUS Score	100	100	100

As for the interviewing results, none of them initially expressed any difficulties using the system and had not changed their opinion of the application use over the five-month period. They all said that since they were getting accustomed to it, it became very fun to use. One thing they would like to add is a “Print” function, so that they can easily replicate and include computer-generated patient drawings in their patient paper charts.

6.3. The Evaluation of the ClockAnalyzer Application

The main concern for the ClockAnalyzer System was whether it was well designed for them to use in their everyday practice. I had questions such as, “How easily is the application used?” and, “How useful are the functions designed in the application?” In this section, I first describe the participants and procedures. Then, the ClockAnalyzer is evaluated in three different perspectives: overall system use, usefulness of the functions, and novel behavioral data visualization.

6.3.1. Participants

I conducted an in-depth interview with three medical practitioners who screen dementia and its related diseases in their work. Medical Practitioner A obtained a Ph.D. in neuropsychology. Medical Practitioner B, as a neuroscientist and a psychiatrist, holds an MD and Ph.D. in neuroscience. Both are affiliated with the Emory Alzheimer’s Disease Research Center. The third participant, Medical Practitioner C, obtained an MD and an MPH in Public Health. As a board certified doctor in internal and geriatric medicine, he is a geriatrician who works at Emory School of Medicine's Geriatric Medicine and Gerontology department. He also works at the Emory Center for Health in Aging; in his position, his focus is on how to increase the early detection of dementia through screening. All of them are potential users who would play a pivotal role in deploying the ClockMe System in real clinical environments.

6.3.2. Procedures

The main concern for the ClockAnalyzer System was whether it is well designed for them to use in their everyday practice. I asked three medical practitioners to use the system by themselves, and I interviewed them as to how they used the system. I first gave them a demonstration on how to use the ClockReader and I explained how the data were stored and saved. Then I showed them how they could access an individual patient's data in the ClockAnalyzer Application. I then asked them to work with both applications. I explained how to use each function and asked them about their opinion of each function and investigated any room for improvement. In addition to interviewing, after each application use, all of the participants completed the SUS questionnaire.

6.3.3. Results

6.3.3.1. Overall use of the ClockAnalyzer Application

To understand the overall use of the application, I conducted the SUS questionnaire. Table 20 shows the SUS results of the ClockAnalyzer System. Unlike the unanimous agreement of the SUS results for the ClockReader System, medical technicians expressed slight difficulties using the ClockAnalyzer System. Especially, two of them indicated they would imagine that most people would learn to use this system with some support. However, overall scores (95, 93, 100 respectively) showed that the ClockAnalyzer was also well designed to use.

Table 20. System Usability Scale for the ClockAnalyzer Application Use

	Medical Practitioner A	Medical Practitioner B	Medical Practitioner C
1. I think that I would like to use this system frequently.	5	5	5
2. I found this system unnecessarily complex.	1	2	1
3. I thought the system was easy to use.	5	5	5
4. I think I would need the support of a technical person to be able to use this system.	4	1	1
5. I found the various functions in this system were well integrated.	5	4	5
6. I thought this system was too inconsistent.	1	1	1
7. I would imagine that most people would learn to use this system very quickly.	4	4	5
8. I found the system very cumbersome to use.	1	1	1
9. I felt very confident using the system.	5	5	5
10. I needed to learn a lot of things before I could get going with this system.	1	1	1
Total Score	32	37	40
SUS Score	95	93	100

6.3.3.2. Use (fullness) of the Functionality

The ClockAnalyzer reports the analysis of an individual patient's clock drawing. Here, I reported how medical practitioners liked to use each function in their practice through in-depth interviews. Overall, they stated that all of the functions were useful and they liked to collect this particular information for their patients for future clinical research. Figure 42 shows the screen capture of the ClockAnalyzer application.

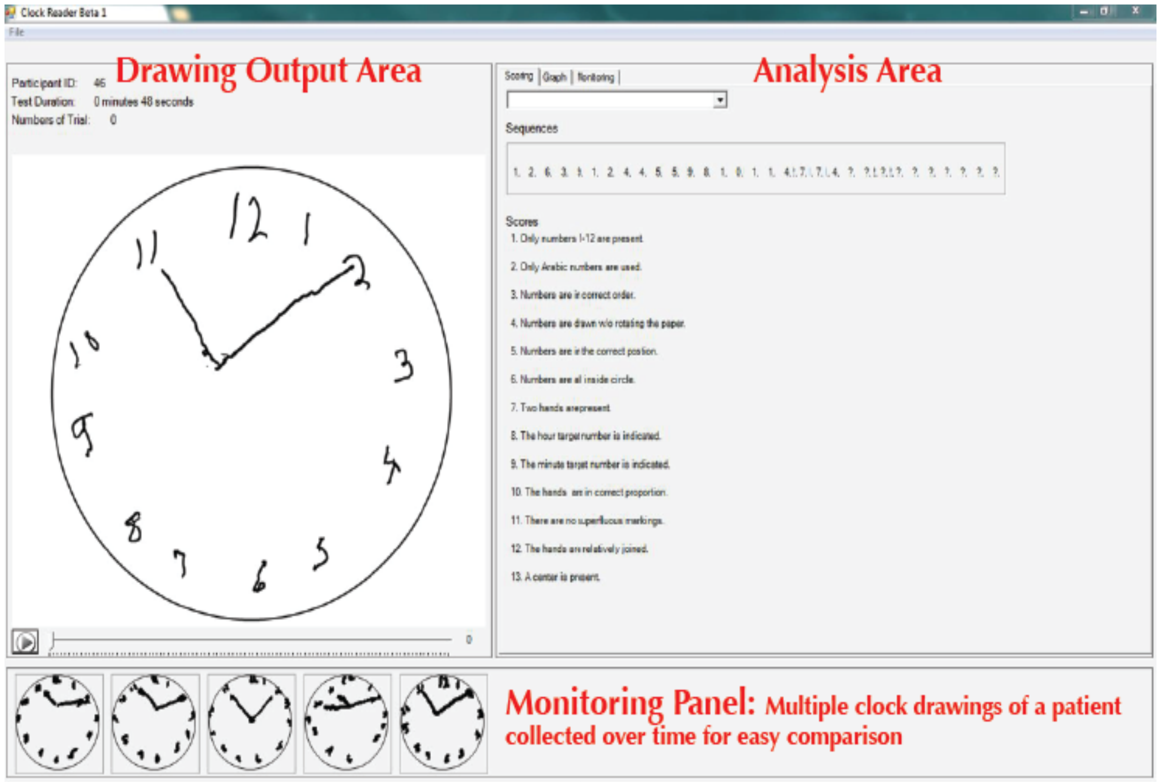


Figure 42. Screen Capture of the ClockAnalyzer Application

A. Time to Completion

All three medical practitioners unanimously agreed that the completion time would be very helpful, and they loved to collect this for their patients. The time to completion per each drawing could them track the CDT data longitudinally together with the scores. Two practitioners especially hypothesized that time to completion could help identify early signs of cognitive impairment. For example, some people may draw a clock correctly; however, they may take a significantly longer time to complete the task in the following year. The average completion time was 42 seconds. Figure 43 specifically shows the number of participants per each completion time range.

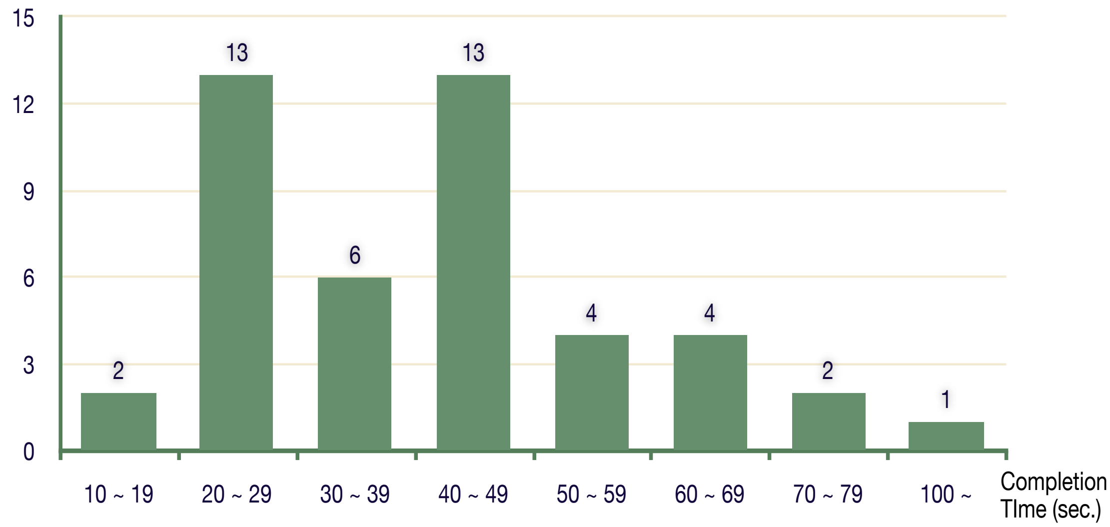


Figure 43. The number of participants per completion time

I also explored how the values correlated with other factors of the participants, such as age, education, CDT score, MMSE [3]score, depression measurement [4]score and gender. Statistically, the time to completion had a positive relationship with age and depression level ($r = .40$, $p < .01$ for age, and $r = .39$, $p < .05$ for depression, respectively). These results indicate that the older the patient was, the longer time it took for him/her to complete the test. Also, patients with higher depression levels tend to spend more time

---

[3]MMSE stands for the Mini-Mental State Examination, which is popularly used to screen for cognitive impairment. The MMSE specifically consists of a 30-point questionnaire with questions including orientation to time and place, registration, attention, calculation, recall, language, repetition, and complex commands. As for the interpretation of MMSE scores, any score greater than or equal to 25 (out of 30) points indicates normal cognition. However, a score below 25can indicate severe (less than or equal to 9 points), moderate (between 10 to 20 points), or mild cognitive impairment (between 21 to 24 points).

[4]All of our study participants were required to take the Geriatric Depression Scale (GDS) in order to measure depression level. Scoring higher than 5 may indicate the presence of depression. I was able to collect MMSE and Depression level scores because all of the participants in my study had taken both as a requirement of being in the Honor registry.

completing the test. MMSE score, CDT score, and Education were statistically negatively correlated with the time to completion (r = -.45, $p < .01$ for MMSE; r = -.31, $p < .05$ for CDT; and r = -.30, $p < .05$ for Education, respectively). If a patient had a higher MMSE score and CDT score, the completion time was shorter. The more education a patient had, the less time he or she took for the test. In addition, there was no relationship with respect to gender and the time to completion.

B. Numbers of Trials

Three practitioners all liked to capture the number of trials per each drawing session. The number of trials recorded if a patient pressed the START OVER button and liked to draw restart to draw a clock. They all said that they thought the number of trials would be useful. However, they did not explicitly explain how to use the data for future use. For example, Medical Practitioner A said, "I usually asked a patient to draw a clock one time; however, if a patient wants to restart to draw a clock, it would be good to track it." Medical Practitioner C also commented, " I also want to record the clock a patient attempted to draw and erase. So I like to see the clock deleted together with the one they submitted." In our current system, the one they submit by the COMPLETE button is the only one recorded. In the next iteration of the ClockAnalyzer user interface, I intend to add a button that shows the history of multiple trials of clock drawings per individual patient.

Among the 45 patients' data, six participants only tried to draw a clock one more time. Table 21 shows the number of trials by participants who joined the deployment study.

Table 21. The Number of Trials by Participants

Number of Trials	Number of Participants
1	39
2	6

Statistically, there was no relationship among the number of trials and other factors such as gender, age, time to completion, education, MMSE score, CDT score, depression level, and computer usage experience.

C. Animation

All three medical practitioners reported that the animation function was very useful and they liked to use it. Medical Practitioner A said, " I definitely want to use it. This is wonderful! Because I can get qualitative data of the clock draw, I can understand how patients organize and (I hope) it may help me to understand frontal disorders."

Both Medical Practitioners B and C commented that they liked to have a button used for a video player as a fast-forward. Medical Practitioner B said, "I think it would be great if there were a button which could provide animation speed at either two- or four-times a faster view. Then, I can save time and review all of the patient's data at once." Medical Practitioner C also said, "The reason I'd like to add a fast-viewing function is that there are some patients who really draw a clock slowly. I saw a patient who drew a clock in four minutes."

D. Sequences/Process

All three medical practitioners stated that the function capturing the sequences and drawing process can be useful. They thought such information would be considered as the patients' planning strategy, so they liked knowing each patient's sequence of clock drawing. Medical Practitioner C said, "It is similar to the animation function. But I like the animation better. What I really want to know is whether a patient had a space planning strategy or not. For example, some patients drew a clock with the pattern of 12, 3, 6, and 9, while others drew a clock without any pattern such as 12, 1, 4, 7… " Among 45 participants' drawing, 22 participants drew a clock with the pattern of either 12, 3, 6, and 9 or 12, 6, 9, 3. Table 22 shows the number of participants using a space planning strategy.

Table 22. The Number of Participant Who Used a Space Planning Strategy

Use of Space Planning Strategy	Number of Participants
Yes	21
No	24

E. Monitoring

The three medical practitioners all agreed that the last function, the monitoring tab, which shows multiple clocks per patient, would give them more insight into understanding the patient's symptoms. Medical Practitioner A said, " Tracking the clinical condition is important, since I can understand whether patients are getting worse or better. Or I can also see how patients respond to the medication."

6.3.3.3. Use of Novel Behavioral Data Visualization

In this section, I describe the interview results of current visualization in the ClockAnalyzer. Specifically, I provide two (for airtime) and three (for pressure) different visual representations of the behavioral data, and I asked the practitioners which ones they liked to use and how they wanted to use those data for their potential research.

A. Airtime

I created two different graphs for airtime visualization. As shown in Figure 44, in the first graph, Graph a's x-axis indicates a stroke order written when a patient drew a clock. The y-axis of the Graph a indicates the airtime of each stroke (numbers of the hand arrows) drawn. The unit of airtime is captured at the millisecond level.

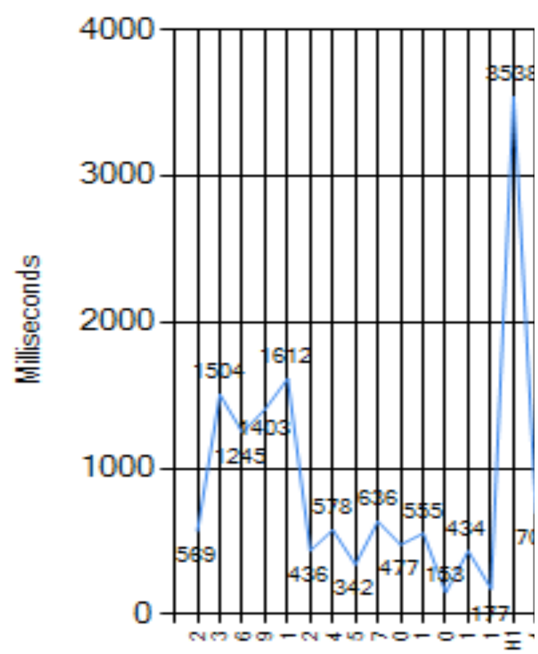


Figure 44. Graph a: Airtime by Recognized Character

Figure 45 shows the second graph, Graph b, which is the modified version of the first one. Graph b represents the lengths of the x-axis as airtime duration. The more airtime a user spends before drawing each digit, the longer the line will be. For example, a user spent 1504 milliseconds before she wrote the digit “3” and spent 1245 milliseconds before she wrote digit “6.”

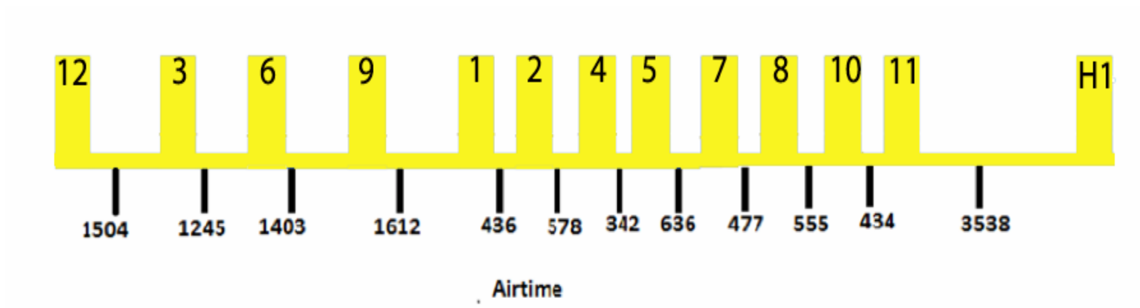


Figure 45. Graph b: Airtime by Recognized Character

A.1. Which graph do you prefer to use?

All three medical practitioners said that they liked “the second graph (graph b)” better than the first one (graph a). They unanimously told me that the second one was easier to understand in terms of the airtime. However, Medical Practitioner B said that he liked to have the first graph (graph a). He said, “ I think it would be good to have both graphs…because I like to see fluctuations or patterns of movement. For example, if we click a button, the second graph (graph b) can be switched to the first graph (graph a), or vice-versa. Anyway, it is really good that we can capture those data.”

A.2. Do you think that airtime can be used to differentiate different types of dementia such as Parkinson’s, Alzheimer’s, and Lewy-Body?

Two medical practitioners agreed that they could use the data to understand different types of dementia. Medical Practitioner B said, “ If we have a bunch of data, I want to know about airtime differences among different types of dementia. I can see potentials of using them. I’d also like to add another measurement such as the jaggedness of lines. For example, smoothness of individual pen-strokes or line-segments, especially, I believe that those data can be used to understand Parkinson’s patients.” However, Medical Practitioner C commented that he did not know how to use them to differentiate sub-types of dementia.

A.3. Some researchers argue that airtime can be related to depression, which may help them make a differential diagnosis between depression and dementia. What do you think about this argument?

Medical Practitioners A and B agreed that airtime can be used to make a differential diagnosis between depression and dementia. Medical Practitioner A said, “ I do not read the article you mentioned. However, I think airtime can be very interesting biomarkers for the identification of MCI or other types of dementia. As for depression and dementia, I would be very interested in using it to test several research questions I have. For example, I am assuming that depressed people will be slower, but they will be able to draw a clock. People with dementia may be faster, but they will be incorrect.” In addition, Medical Practitioner C also agreed and said, “Even though I have no idea about using airtime for depressed patients, I think delayed initiation can be related to depression.”

B. Pressure

I created three different graphs for pressure visualization. Figure 46 shows the first graph (graph c) with raw pressure data[5]. The x-axis indicates a stroke order written, and the y-axis shows the raw pressure value.

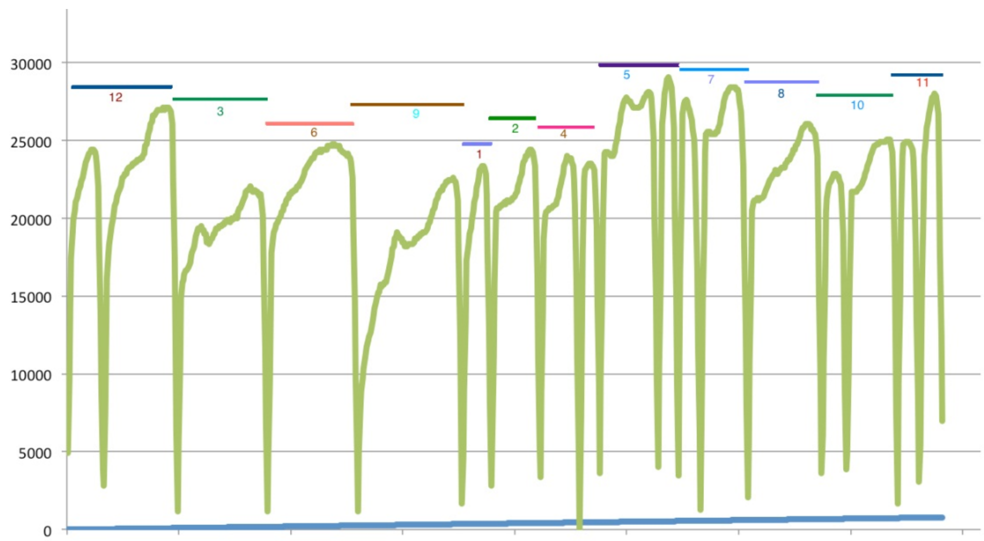


Figure 46. Graph c: Pressure (Raw Value) by Recognized Character

Figure 47 shows the second graph (graph d) with the average pressure data. While a patient writes an individual digit, the pressure value is not represented as a single digit. For example, if a patient writes the digit "12," the pressure value can be different per each digit. Therefore, the raw pressure value can be too much information for medical practitioners. Thus, I averaged the pressure value per each recognized digit to simplify the first graph (graph c).

[5]I used the Fujitsu Lifebook T900 Tablet PC for the application development. The ranges of pressure value captured by the T900 are a minimum value of 0 to a maximum value of 32767.

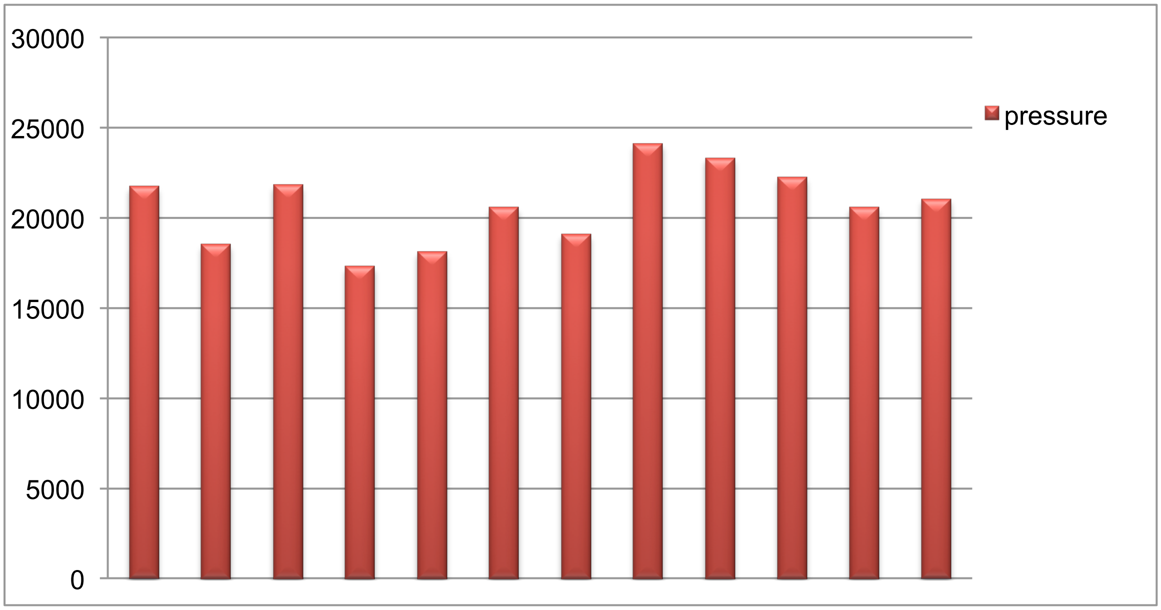


Figure 47. Graph d: Average Pressure by Recognized Character

The third graph, graph e, shown in Figure 48, includes three basic data such as the lowest, highest, and average values of pressure by recognized characters.

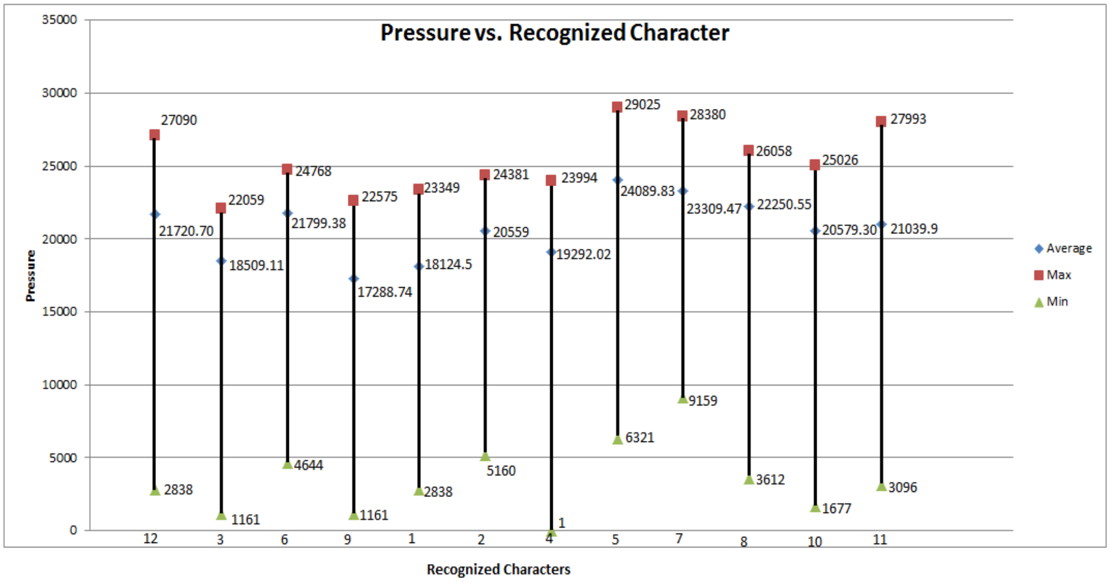


Figure 48. Graph e: Pressure (Min, Max, and Average Value) by Recognized Character

B.1. Among the three visualization of pressure data, which one do you like the best? Why? Do you have any ideas to create a better visualization?

Interestingly, the three medical practitioners had different preferences of pressure visualization. Medical Practitioner A said that she liked the second one best because it was the simplest one, which made it easily understandable. Medical Practitioner B liked the first one best. He said, “I like the first graph because the raw values enable me to explore the overall patterns of pressure by looking at the smoothness of the line. But it would be great if you can combine the first and the second graphs. For example, put the average mark as a hash mark into the first graph.” Medical Practitioner C said, “I like the third one best. But, I want to change something. First, the unit of the Y-Axis can be simplified as 5, 10 rather than 5,000 and 10,000. Second, for the numbers in the line graph, I only want to show the median value and remove the maximum and minimum values. For those values, if I want to see them, I can always look at the Y-axis to find out the exact values."

B.2. Do you think that there is some potential value to using these pressure data for clinical purposes? If so, what value would that be?

All three medical practitioners reported that they saw potentials in using pressure for Parkinson's patients. They all hypothesized that a patient with Parkinson's had a different lower pressure value, compared to others. Medical Practitioner A said that she thought pressure could be useful and have a great potential to understand patients with Parkinson's, since tremor and pressure can be related to each other. Thus, she was interested in exploring pressure data with Parkinson's patients. Medical Practitioner B commented that pressure could be an important factor in understanding different movement disorders such as Parkinson's disease. Moreover, Medical Practitioner C mentioned, " I think that pressure can also be related to some level of certainty. So it would be interesting to explore correlations with other factors."

C. Combination of Airtime and Pressure

Combined data sometimes can provide more insights or provide a better understanding of the information. Thus, I created two graphs visualizing airtime and pressure together. Figure 49 shows the first graph, graph f, which combines the raw pressure value together with airtime. The X-axis of graph f indicates airtime and the Y-axis represents the value of pressure. Further, the recognized characters are shown at the middle of the individual pressure value graph. The second graph, graph g, shown in

Figure 50, is a combination of the average pressure value together with airtime. Each recognized character is shown at the top of the bar graph.

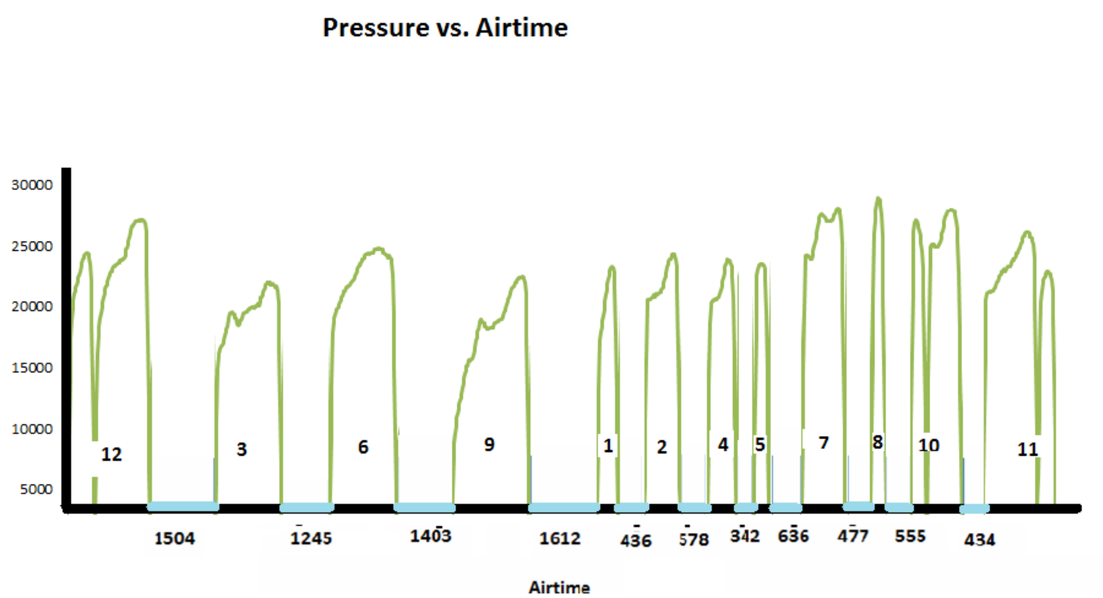


Figure 49. Graph f: Pressure (Raw Value) and Airtime by Recognized Character

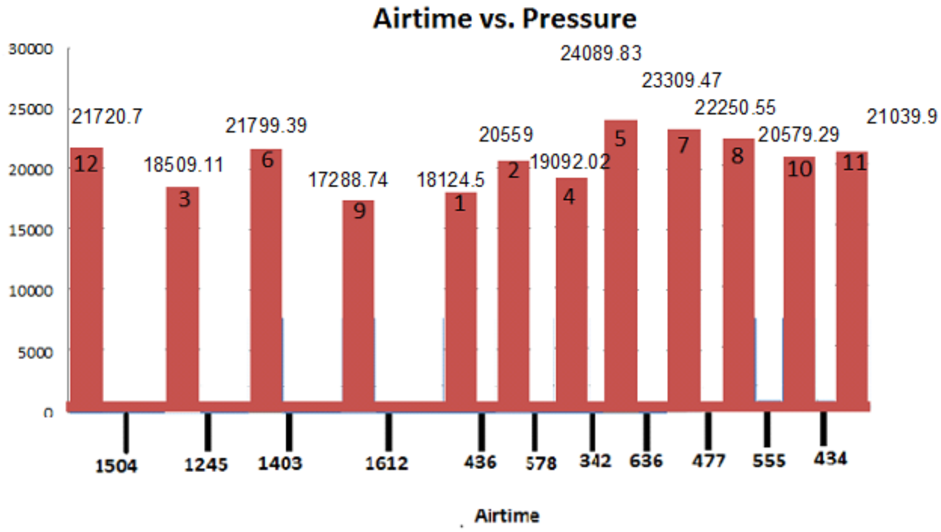


Figure 50. Graph g: Average pressure and airtime by recognized character

C.1. What do you think about those combined data visualizations? Are they better than separate data visualization? If yes, why do you think so? Do you have any other ideas to make it better?

Again, interestingly, all three medical practitioners had mixed preferences of the combined data visualization. One preferred the first graph (graph f), and the other liked the second graph (graph g). However, the last one did not like the idea of representing two combined data. Especially, Medical Practitioner A said, “I do not see the advantages of combining two data. Airtime and pressure cannot be related to each other. It doesn’t seem important to be together. So I don’t like the idea of showing both together.”

6.4. Summary and Discussion

In the following section, first I discuss the relationship among different factors through the statistical analysis. Then, I synthesize the results of the user study from all of the participants, such as the medical technicians, medical practitioners, and patients.

6.4.1. Investigating the correlation among different factors

First, through the statistical analysis of the collected data, I was able to investigate whether different factors correlated with each other. This helped me answer my third research question, " Can a computerized tool perform better?" Specifically, the behavioral data, the Time to Completion, shows many co-relationships among different factors. Therefore, medical practitioners are able to use these data for their future studies or to test their hypotheses.

Prior to reporting the statistical analysis, I would like to review the way I collected the data. The data consisted of (1) standard cognitive assessment data such as MMSE scores, CDT scores, Boston Naming scores, and depression scores; (2) each patient's demographic information such as age, gender, and education; and (3) computerized CDT information, such as the time to completion, sequences, and the number of trials. I used the Emory ADRC NECO test results data to collect standard cognitive assessment data. Since all participants were enrolled in the Honor registry, Emory ADRC conducted standard cognitive assessment, called the NECO test, every year. From the pre-survey results, I collected individual patients' demographic information. From the computerized Clock Drawing Test, through using the ClockReader, I was able to collect the time to completion, sequences, and the number of trials per each drawing. Here are the summaries of the data analysis.

- **The Time to Completion positively correlates with age and depression levels**($r = .40$, $p < .01$ for age, and $r = .39$, $p < .05$ for depression, respectively)**.**

If patients were older, they took more time to complete the CDT. Also if patients had a higher depression score, they spent more time to take the CDT test.

- **The Time to Completion negatively correlates with the CDT score** (r = -.31, p< .05 for CDT)**.**

  If patients had a higher score in CDT, they spent less time to complete the CDT.

- **The Time to Completion negatively correlates with MMSE score and Education** (r = -.45, p< .01 for MMSE and r = -.30, p< .05 for Education, respectively)**.**

  If patients had a higher score in MMSE and were educated more, they spent less time to complete the CDT.

- **Education is a factor that has a positive relationship between the MMSE score and the Boston Name score test** (r = .39, p< .01 for MMSE, and r = .45, p< .01 for Boston Name score, respectively)**.**

  If patients had a higher education, they had a better score on the MMSE and on the Boston Naming score test. This indicates that education can highly affect the results of standard assessment tests.

Second, as shown in Section 6.2.1.3. Results, I statistically proved my second research question, "Can a computerized tool replicate the exact same results as the paper-and-pencil test?" I triangulated three different data to answer this question: patients' interviews and their drawings, and an analysis of both their drawings (computer and paper-and-pencil) from an evaluator. There were statistically no significant differences in the two different methods. Moreover, none of the patients, despite their age, education, gender, and computer experience, reported any difference in drawing a clock using either method. In brief, what the participants drew and reported in using the computerized system showed consistency. In addition, their two drawings were evaluated and were shown to be similar in terms of their test results.

Third, to explore my third research question in terms of accuracy, "Is this tool able to increase the accuracy of cognitive impairment assessment?" I investigated whether the different evaluators, the medical technicians, would score the drawings differently; in other words, whether scoring inconsistency would exist among different evaluators.

I asked two medical technicians to evaluate the same clock drawings and found out that there were differences that existed among the different evaluators. Please refer to Section 5.4. for detailed examples. However, statistically the results from the two deployment studies were mixed. The first deployment study shows that there were statistical differences existing between two criteria, while the second deployment study did not show any statistical differences. Despite the absence of statistical differences, incidents of discrepancy occurred 12 times. These discrepancies resulted in 8 cases of different total scores between the two technicians. These results indicate that human subjectivity or levels of experience always influence scoring results. Therefore, computers can increase consistency. Appendix D contains the table showing discrepancy of each criterion.

6.4.2. User Study: Ease of Use, Usefulness, and Extension

First, the interface of the ClockMe System was designed for ease-of-use. All six participants (three medical technicians and three medical practitioners) reported that the ClockMe System was easy to use. They really liked the simplicity of the user interface. None of them had difficulties in administering the test, as well as drawing a clock. After testing, they all successfully saved the data. They also stated that they believed there would be no problem for a patient to draw a clock compared to the paper-and-pencil-based CDT. In short, all 45 participants agreed that the system was easy to use and that there were no problems performing the CDT on the computer.

Second, all of the participants reported that the ClockMe System was useful. Patients could take the test anywhere and anytime if they have access to the system. This would encourage them to have fewer trips to the clinic; consequently, they could save their money by minimizing the need to see a doctor. Medical technicians expressed the system's usefulness in terms of data saving and automatic scoring. This would save their time in scanning analog paper data into electronic data. Last, medical practitioners really appreciated the usefulness of the ClockMe System in terms of the data-capturing methods. For example, the completion time and sequences of drawing were easy to collect by the computerized Clock Drawing Test, and the completion time showed several correlations with other factors. Medical practitioners also indicated that they would like to use novel behavioral data to test their hypotheses and to understand their patients in-depth. Thus, the ClockMe System would be able to empower medical practitioners with strong evidence, such as computer-collected information.

Third, medical practitioners expressed their hope to enhance the current system. For example, Medical Practitioner B wanted to include a printing function in the ClockReader System. In fact, the ClockAnalyzer already includes this function. However, he said, “Even though the function exists in the ClockAnalyzer, clinicians definitely wants to have this function in the ClockReader too… Because this would provide clinicians with handy paper-version-clocks.” This indicates that despite the fact that the medical practitioners all liked and agreed with the advantages of computerized assessments, they also wanted to have an option of the traditional way to keep patients’ data. Furthermore, three of them stated that if the ClockMe System could run on a smaller and lighter machine like the iPad, it would be ideal for them. Rather than carry around a heavy, bulky machine, a lighter machine would be more convenient. Medical practitioner A specifically said, “As long as the machine has a clear contrast, the size of the machine does not matter. I am only concerned about the screen contrast. Since many older adults have low vision, it is important to provide a strong contrast on the screen.”

CHAPTER 7

CONCLUSION AND FUTURE WORK

Among many other approaches intended to cure dementia-related diseases, prevention through delay is currently proposed as the best way to tackle dementia (86). To prevent dementia, early detection is indispensable because it could delay the onset and progression of the disease. Computing technologies can contribute to empowering medical practitioners to conduct timely and effective screening for identifying people at risk for or with asymptomatic dementia. In this dissertation, I investigated the effectiveness of computing technologies to enhance the dementia-screening process by presenting the design, implementation, and evaluation of the ClockMe System. The following section briefly reviews the answers to my three research questions. I then conclude by stating the contributions of my study, as well as future research directions.

7.1. Summary of the Research Question Answers

I had three research questions answered through the iterative design, development, and evaluation process of the ClockMe System. Figure 51 provides a brief method overview of how I answered those questions.

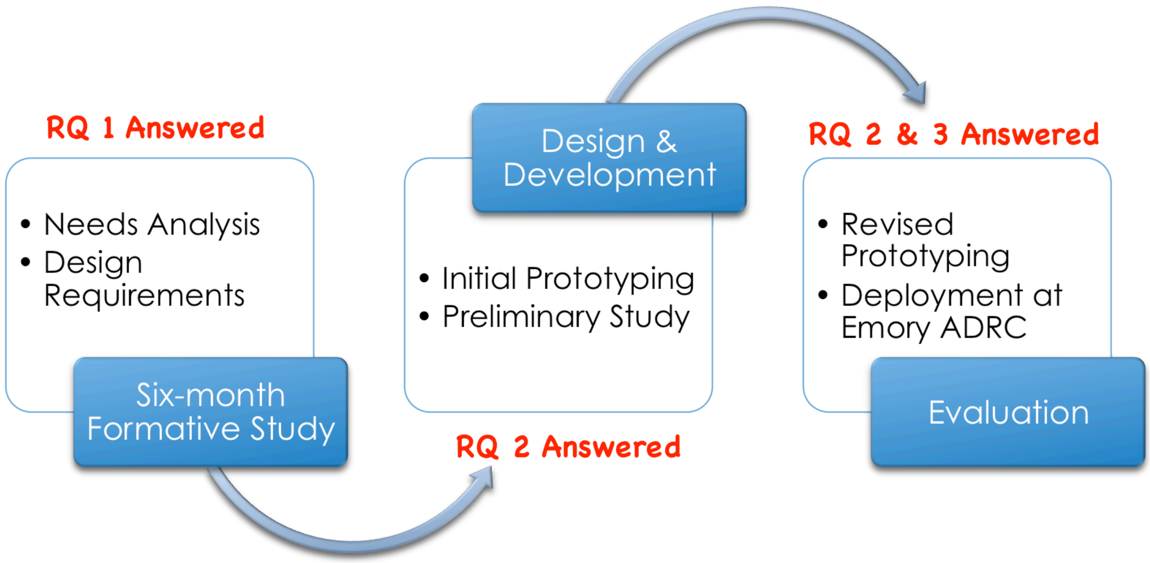


Figure 51. Overview of Research Methods

Research Question 1 (RQ1)

What do doctors need? What are the design requirements of a computerized tool for neuropsychologists and neurologists to do an analysis, screening, and diagnosis for dementia?

In my six-month formative study at Emory ADRC, I discovered the barriers of screening dementia and explored the possible technological opportunities to support the screening process. Through observations of paper-and-pencil memory testing sessions, I specifically identify three issues that can be addressed by technologies.

- The drawings are predominantly analyzed at the end of the tests as a final product.
- Clinicians are required to do many things at once, such as measure the time and observe the patient. As a result, some critical moments can be easily missed.

- The way data are collected and stored is not appropriate for monitoring any changes of cognitive condition.

Those issues informed me to design several components of the ClockMe System. Specifically, to fully utilize the patient's drawings, I propose that the user interface should integrate all of the data that will support the clinician's decision. For example, clinicians cannot examine or analyze the process of how the drawings are constructed because of the difficulty in the data collection. Providing an automated capturing process, such as an animation, would be helpful in that such information contributes to creating an automatic replay of the drawing process, time, and pressure. In addition to these automated data collection features, I have gained an in-depth understanding of contextual knowledge, such as interactions between clinicians and patients in the screening process, and interactions with the materials. This knowledge also led me to appropriately design a user interface with respect to the characteristics of potential users: older adults who are concerned about their cognitive condition.

Research Question 2 (RQ2)

Can the computerized tool yield the same results as the paper-and-pencil Clock Drawing Test?

- **Do patients feel equally comfortable using both methods?**
- **Do patients perform the same?**
- **Do technicians score both drawings the same way?**

To answer this question, first, I conducted a usability test with older adults, a potential user group for dementia screening. The activities included a demographic and background survey of older adults (prior computer experience and Internet use), use of the computerized screening tool, and post-interviews regarding the tool use. Then, I compared the testing results, including the quality of the drawing and the scoring, between the traditional and computerized methods. I investigated whether the differences were simply due to the use of computerization, or some other factors. According to the results from both the preliminary and deployment studies, there were no statistically significant differences between the two different methods. To sum, patients felt no differences when they used the computer tool versus when they used the paper-and-pencil tool. Patients also produced the same quality of drawings for their cognitive assessment between both methods. Further, those drawings were evaluated similarly by the evaluator who had administered and scored the tests. Therefore, I argue that people using the computerized tool replicate the exact same results as the paper-and-pencil test.

Research Question 3 (RQ3)

Can a computerized tool perform better? Would a computerized tool provide more information? Can this tool increase the consistency of cognitive impairment assessment? Can this tool provide new diagnostic insights for medical practitioners about their patients by providing concrete data beyond that offered by the paper-and-pencil test?

I argue that a computerized tool can perform better in dementia screening. First, computers avoid influencing human subjectivity during the scoring process. Sometimes it is difficult to unanimously agree upon interpretations of all criteria. Such lack of unanimity can be due to the inherent ambiguity of drawings or a lack of training on the part of an individual evaluator. My preliminary study indicates that two evaluators showed discrepancies in their scoring. Second, a computerized screening tool can provide more and richer data than the current pencil-and-paper-based tool. Therefore, this tool can help medical practitioners gain better diagnostic insights. For example, with a computerized tool, we can record the time to completion per each CDT session. Also, the data – time to completion – collected from my deployment study showed that the completion time positively correlates with age and depression level; on the other hand, it negatively correlates with MMSE score, CDT score, and Education.

7.2. Contributions

Now, I outline four major contributions I have made throughout the research activities for my dissertation work.

- The creation of the ClockMe System, a computerized dementia screening tool, which can help clinicians identify cognitive impairment through a more accessible and quick-and-easy screening process
- The delivery of computer-collected novel behavioral data, which may offer new insights and a new understanding of a patient's cognition, enabling clinical researchers to ask new questions about the relationship between the CDT and cognitive abilities

- An in-depth understanding of different stakeholders such as patients, medical technicians, and medical practitioners, and the identification of their common user needs and desires within a complicated healthcare workflow system
- The triangulation of multiple data collection methods such as ethnographical observations, in-person interviews, focus group meetings, weekly phone interviews, and quantitative data from a user survey, in a real-world deployment study

Ultimately, the utmost contribution of this dissertation shows the promise of truly interdisciplinary work. This study not only contributes to the research in the Human Centered Computing field, but it also contributes to the field of medicine, especially dementia screening as well as neuropsychological assessments, with the potential of using computer-collected behavioral data. Patients' depression assessment data positively correlate with computer-collected behavioral data such as patients' CDT completion time. This finding indicates that people with depression tend to take more time to complete the CDT. Furthermore, people with a lower CDT score also tend to take more time to complete the CDT. In other words, time to completion negatively correlates with correctness of the CDT. In brief, people usually take more time to complete the CDT if she or he has depression or dementia. Thus, medical practitioners can use time-to-completion data from the CDT in order to differentiate dementia from depression. Some people may take more time to complete the CDT due to depression; if they have an appropriate higher score on the CDT, they may have depression. Others may take more time to complete the CDT, together with a lower score on the CDT; this, on the other hand, may indicate a diagnosis of dementia. Figure 52 shows the feature comparison summary of the paper-and-pencil Clock Drawing Test and the ClockMe System.

	Paper-and-Pencil Clock Drawing Test	The ClockMe System
Score	✔	✔
Time to Completion		✔
Number of Trials		✔
Animation		✔
Sequences/Process		✔
Monitoring		✔
Air Time		✔
Pressure		✔

Figure 52. Feature Comparison of the paper-and-pencil Clock Drawing Test and the ClockMe System

7.3. Future Work

This dissertation work opens the door to enhancing current screening practices by developing a computerized Clock Drawing Test. I would like to take the initiative to conduct future research in three directions: enhancement of the current system, extension of other drawing applications, and the deployment of the system for clinical trials.

First, I plan to enhance the current system by creating other scoring modules. The current ClockMe System includes only Freedman's 13-point scoring algorithm. By adding these new scoring algorithm modules, medical practitioners will be able to choose and compare results calculated from different scoring criteria. For example, some neurologists follow Freedman's 15-point evaluation criteria, while others use Mendez' 20-item qualitative criteria (61, 96). Therefore, they may derive different diagnoses and treatments if the user interface incorporates different evaluation algorithms for medical practitioners to choose and compare results calculated from the different scoring criteria. Providing an analysis with multiple criteria will enable medical practitioners to critically analyze the current situation with diverse perspectives and will advance knowledge of early cognitive impairment detection never before possible.

Second, in order to increase sketch recognition in drawings, the online recognition algorithm we implemented should be integrated with an offline sketch recognition algorithm. First, I plan to use a large number of patients' scanned data at the Emory ADRC to develop machine-learning algorithms for offline sketch recognition. Then, to verify a new algorithm, I will gather a large collection of clock-drawing data, both from patients and a normal aging population, through system deployment at diverse places such as community centers or clinics.

Third, the system can also be enhanced through modification to meet the needs of different stakeholders and their environments. The tool has many different types of stakeholders in terms of medical practitioners, caregivers, and consumers. For example, primary care doctors who do not have expertise in dementia screening may want to use a tool situated in a waiting room area. Senior housing nurses or caregivers may want to put the tool in their library or common areas so that anyone who wishes to test their cognitive functioning can try it. If the system is available on the Internet, it could also be used as a home-based cognitive assessment tool for dementia screening. This would eventually increase the opportunity to screen people at risk for dementia. Furthermore, specialists, such as neurologists or neuropsychologists, may want to explore additional types of behavioral data collected with such a tool. It would be ideal if this tool could serve these different needs and purposes.

Fourth, from the formative and deployment studies I conducted, I learned that there are many other drawing assessments available. The doctor who attended the user study, in particular, expressed his/her desire to develop a digital version of complex figure assessment. By simply using the current technologies, we can extend its utilization. The need to make a digital version of the paper-and-pencil-based assessment comes from the fact that a pencil drawing is usually considered to be a final output analysis. However, with digital pen-based computing, doctors can collect two different types of drawings: a static drawing and a dynamic one. This dual drawing process enables them to conduct a drawing analysis from two different perspectives.

Last, conducting clinical trials using the system may advance an understanding of the different types of cognitive impairment. These clinical trials can especially help utilize novel behavioral data, such as pressure or airtime. For example, clinicians may find interesting patterns of airtime from aggregated data through a large number of patients with Parkinson's disease, compared to control groups. Therefore, it is necessary to conduct clinical trials with specific disease groups for subsequent study. I also plan to improve the current visualizations of behavioral data by conducting in-depth interviews and focus group meetings with an appropriate large number of doctors. Additionally, when the system is used for clinical trials, issues of data encryption and security, along with the integration of patients' electronic medical records, need to be thoroughly studied.

APPENDIX A:

CLOCKREADER PROJECT: USABILITY COMPARISON TEST

PRE-QUESTIONNAIRE

Please answer the following questions. All of your answers will be strictly confidential. Any published reports will not identify individuals with their answers. If there is a question you do not wish to answer, please just leave it blank and go on to the next one.

Basic information about participant		
1. Gender:	□Male	□ Female
2. Age:	( )	
3. Education:	□High school graduate □College □Graduate school:	Major ( )
4. Previous (or current) Job:	( )	

Computer Experience	
5. Are you familiar with using computers?	
□Yes	□No
6. (If yes), How do you rate your computer literacy? □High □Medium □Low	

7. What is the most uncomfortable thing about using a computer?

□Reading texts on the screen
□using a mouse
□typing a keyboard

---

8. What would you prefer if you could use a computer differently? Describe it. (For example, I would like to use computer with voice-activated control.)

---

9. What are your most common activities using your computer?

□Document/ word processing	□Internet (if you choose this option, please answer the following questions: Internet use)	□Games

Internet use and search activity

10. How long (on average) do you use the Internet per day?

□0 to 2 hours	□6 to 8 hours	□more than 12 hours
□3 to 5 hours	□9 to 11 hours	

---

11. What do you think isthe best way to find information on the Internet?

□search engines (Google, yahoo)	□subject directories
□listservs and newsgroups	□other ( )

---

12. What is your main goal in spending time on the Internet?

□emails	□news/shopping	□games
□research	□communication (chatting)	

---

13. How do you organize your favorite places (sites) for frequent Internet visits?

□Tagging □others (   )
□Favorite or Bookmark

Your information is very valuable to us. Thank you for your time and participation in our study.

APPENDIX B:

PRE-SURVEY FOR MEDICAL SPECIALISTS AND SYSTEM USABILITY SCALE SURVEY

Pre-Survey for medical specialists

1. How old are you?

   25 or under   26-35   36-45   46-55   56 and over

2. Gender:

   Male   Female

3. How long have you been administering the Clock Drawing Test?

4. Degrees obtained (e.g. B.S. in Education):

5. How often do you use a computer?
   a. Multiple times per day
   b. Once a day
   c. Few times a week
   d. Once a week
   e. Few times a month
   f. Once a Month
   g. Less than once a month

6. Familiarity with Tablet PC or Pen-based Computing :

   a. Little or no experience
   b. Some experience but no practical experience
   c. Extensive experience

System Usability Scale survey
©Digital Equipment Corporation, 1986

	Strongly disagree 1	2	3	4	Strongly agree 5
1. I think that I would like to use this system frquently.					
2. I found the system unnecessarily complex.					
3. I thought the system was easy to use.					
4. I think I would need the support of a technical person to be able to use this system.					
5. I found the various functions in this system were well integrated.					
6. I thought this system was too inconsistent.					
7. I would imagine that most people would learn to use this system very quickly.					
8. I found the system very cumbersome to use.					
9. I felt very confident using the system.					
10. I needed to learn a lot of things before I could get going with this system.					

Thank you!

APPENDIX C:

THE CLOCKREADER USER MANUAL

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

User Manual 1.0

Hyungsin Kim | hyungsin@gatech.edu

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

1. Make sure you have the correct laptop for the Clock Drawing Test. It should look like this.

2. Open the Laptop.

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

3. Turn-on the Computer by simply sliding the power button, which is near the bottom right corner of the monitor, to the right. Wait about one minute.

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

4. Once you turn-on the computer, just sit back and wait till you get to the login window.

5. To log in to the computer, please click on icon called ClockReader.

6. Please type **emorygatech** in the Password field. Then press Enter, or else find the password at the top left corner of the screen edge.

2

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

7 Once you are logged in, you should be able to see the desktop screen like this.

8 Now, please hold the screen with your hand, as shown in the picture on the right.

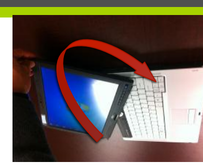
9. While holding the screen, push the screen back till it becomes horizontal (flat).

ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

10. Now hold the top right corner of the screen as shown in the picture, and while keeping it flat, rotate it 180°.

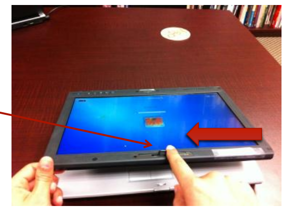
11. Once you get to the position as shown in the picture, stop rotating.



ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

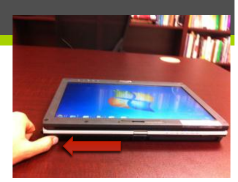
12. While holding the left side of the screen with your left hand, use your right index finger to slide the black latch to the left in order to lock the screen in place.



ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

13. In the left bottom corner of the computer there is a special, pressure sensitive stylus. In order to take it out, slide the stylus to the left.

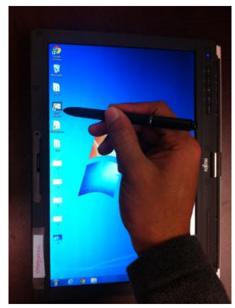


ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

14. Rotate the laptop one quarter turn to the right. (You will see icons on the left.)

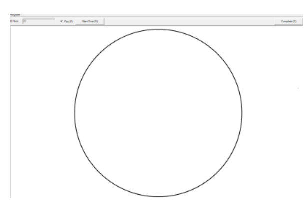
15. With the stylus, double click on the "CDT Shortcut" icon on the monitor.



ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

15. A screen with a circle will open.



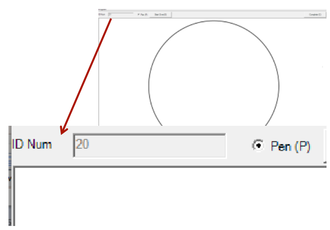
ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

16. On the top left hand of the screen, find the ID Num. This is the patient's ID number.

17. Write down the ID Num on the patient's chart.

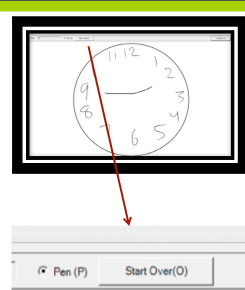
18. Now, the laptop is ready for a patient to draw a clock. Give the instructions to the patient.



ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

19. If the patient wants to re-draw the the clock, she can simply just press the start over(O) button as shown. This will erase all of the drawing.

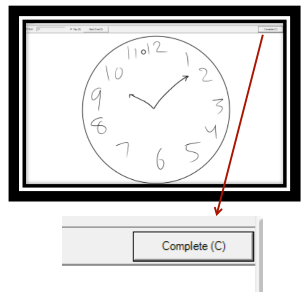


ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

20. Once the patient is done drawing, she can simply click the button **Complete(C),** located in the top, right-hand corner of the screen.

21. Now, you should take the laptop from the patient. Make sure that the stylus is returned as well.

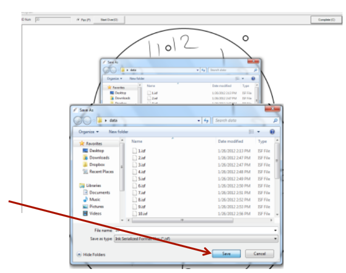


ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

21. Once the Complete(C) button is pressed, you will see a pop-up window as shown on the right.

22. Then simply press the save button**.**

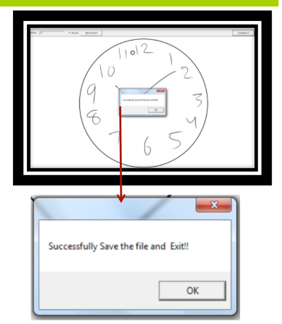


ClockReader: Computerized Clock Drawing Test

Georgia Tech and Emory School of Medicine

23. Finally Click on OK to exit the program.
    This may take a few minutes.

24. Please make sure to return the stylus to its dock.



APPENDIX D:

NUMBER OF OCCURRENCE OF DISCREPANCY BETWEEN TWO EVALUATORS

	# of occurrence of discrepancy between two evaluators
1.all 12 numbers (only 12)	2
2.onlyArabic numbers	0
3.correct order	1
4.drawn without rotating paper	0
5.close to quadrants	0
6.inside circle	1
7.two hands present	2
8.hour correct	0
9.minute correct	0
10.correct proportion	1
11.no superfluous marks	2
12.hands joined	2
13.center	1
Total	8

REFERENCES

1. **Agrell, B., and O. Dehun.** 1998. The clock-drawing test. Age and ageing **27:**399-403.

2. **Aharonson, V., I. Halperin, and A. D. Korczyn.** 2007. Computerized diagnosis of mild cognitive impairment. Alzheimer's & dementia : the journal of the Alzheimer's Association **3:**23-27.

3. **Alzheimer's Association.** 2011. 2011 Alzheimer's disease facts and figures. Alzheimer's and Dementia: The Journal of the Alzheimer's Association **7:**208-244.

4. **Arias-Torres, D.** 2006. The Design and Evaluation of a Pen-Based Computer Interface for Novice Older Users. ENC '06. Seventh Mexican International Conference on Computer Science**:**142-150.

5. **Ashford, J. W.** 2008. Screening for memory disorders, dementia and Alzheimer's disease. Aging Health **4:**399-432.

6. **Backman, L., S. Jones, A. K. Berger, E. J. Laukka, and B. J. Small.** 2005. Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. Neuropsychology **19:**520-531.

7. **Berg, R., M. D. Franzen, and D. Wedding.** 1994. Screening for brain impairment : a manual for mental health practice, 2nd ed. Springer Pub. Co., New York.

8. **Beyer, H., and K. Holtzblatt.** 1999. Contextual design. interactions **6:**32-42.

9. **Brodaty, H., and C. Moore.** 1997. 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 **12:**619-627.

10. **Brooke, J.** 1996. SUS: A quick and dirty usability scale. *In* P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. K. McClelland (ed.), Usability evaluation in industry. Taylor & Francis, London.

11. **Brown, J., G. Pengas, K. Dawson, L. A. Brown, and P. Clatworthy.** 2009. Self administered cognitive screening test (TYM) for detection of Alzheimer's disease: cross sectional study. BMJ **338:**b2030.

12. **Burns, J. M., and J. C. Morris.** 2008. Mild cognitive impairment and early Alzheimer's disease : detections and diagnosis. John Wiley & Sons, Chichester, England ; Hoboken, NJ.

13. **Butcher, J. N.** 2003. Computerized Psychological Assessment, Handbook of Psychology. John Wiley & Sons, Inc.

14. **Butcher, J. N., J. N. Perry, and M. M. Atlis.** 2000. Validity and utility of computer-based test interpretation. Psychological assessment **12:**6-18.

15. **Cho, Y. S., H. Kim, and E. Y.-L. Do.** 2010. Curved-straight lines-analysis (CSLA) algorithm for handwritten digit recognition enhancement. Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia**:**150-154.

16. **De Pandis, M. F., M. Galli, S. Vimercati, V. Cimolin, M. V. De Angelis, and G. Albertini.** 2010. A new approach for the quantitative evaluation of the clock drawing test: preliminary results on subjects with Parkinson's disease. Neurology research international **2010:**283890.

17. **Dennis, M., T. J. Boll, B. K. Bryant, and American Psychological Association.** 1988. Clinical neuropsychology and brain function : research, measurement, and practice. American Psychological Association, Washington, DC, Hyattsville, MD.

18. **Desai, A. K.** 2010. Healthy brain aging : evidence based methods to preserve brain function and prevent dementia. Saunders, Philadelphia, Penn.

19. **Dixon, R. A., D. Kurzman, and I. C. Friesen.** 1993. Handwriting performance in younger and older adults: age, familiarity, and practice effects. Psychology and aging **8:**360-370.

20. **Dougherty, J. H., Jr., R. L. Cannon, C. R. Nicholas, L. Hall, F. Hare, E. Carr, A. Dougherty, J. Janowitz, and J. Arunthamakun.** 2010. The computerized self test (CST): an interactive, internet accessible cognitive screening test for dementia. Journal of Alzheimer's disease : JAD **20:**185-195.

21. **Dwolatzky, T., V. Whitehead, G. M. Doniger, E. S. Simon, A. Schweiger, D. Jaffe, and H. Chertkow.** 2003. Validity of a novel computerized cognitive battery for mild cognitive impairment. BMC geriatrics **3:**4.

22. **Eyde, L. D.** 1988. Computerized Psychological-Assessment - a Practitioners Guide - Butcher,Jn. Contemp Psychol **33:**886-887.

23. **Feldman, H. H., C. Jacova, A. Robillard, A. Garcia, T. Chow, M. Borrie, H. M. Schipper, M. Blair, A. Kertesz, and H. Chertkow.** 2008. Diagnosis and treatment of dementia: 2. Diagnosis. Canadian Medical Association Journal **178:**825-836.

24. **Finkel , S.** 2003. Cognitive screening in the primary care setting. The role of physicians at the first point of entry. Geriatrics **58:**43-44.

25. **Freedman, M., L. Leach, K. E., W. E., K. Shulman, and D. Delis.** 1994. Clock Drawing: A Neuropsychological Analysis. Oxford University Press, Oxford.

26. **Freedman, M., L. Leach, E. Kaplan, G. Winocur, K. Shulman, and D. C. Delis.** 1994. Clock Drawing: A Neuropsychological Analysis, 1 ed. Oxford University Press, New York.

27. **Gangadhar, G., D. Joseph, and V. S. Chakravarthy.** 2008. Understanding Parkinsonian handwriting through a computational model of basal ganglia. Neural computation **20:**2491-2525.

28. **Gangadhar, G., D. Joseph, A. V. Srinivasan, D. Subramanian, R. G. Shivakeshavan, N. Shobana, and V. S. Chakravarthy.** 2009. A computational model of Parkinsonian handwriting that highlights the role of the indirect pathway in the basal ganglia. Human movement science **28:**602-618.

29. **Giovannetti, T., B. M. Bettcher, L. Brennan, D. J. Libron, R. K. Kessler, and K. Duey.** 2008. Coffee with jelly or unbuttered toast: commissions and omissions are dissociable aspects of everyday action impairment in Alzheimer's disease. Neuropsychology **22:**235-245.

30. **Goldstein, G., and T. Incagnoli.** 1997. Contemporary approaches to neuropsychological assessment. New York Plenum Press.

31. **Grober, E., C. Hall, R. B. Lipton, and J. A. Teresi.** 2008. Primary care screen for early dementia. Journal of the American Geriatrics Society **56:**206-213.

32. **Gualtieri, C. T.** 2004. Computerized Neurocognitive Testing and its Potential for Modern Psychiatry. Psychiatry **1:**29-36.

33. **Gualtieri, C. T.** 2004. Dementia screening in light of the diversity of the condition. Journal of Insurance Medicine **36:**298-309.

34. **Gualtieri, C. T.** 2004. Dementia Screening Using Computerized Tests. Journal of Insurance Medicine **36:**213–227.

35. **Gualtieri, C. T., and L. G. Johnsonb.** 2006. Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. Arch Clinical Neuropsychology **21:**623-643.

36. **Guha, A., H. Kim, and E. Do.** 2010. Automated Clock Drawing Test through Machine Learning and Geometric Analysis. International Workshop on Visual Languages and Computing. Workshop in DMS 2010**:**311-314.

37. **Guha, A., H. Kim, and E. Y.-L. Do.** 2010. Re-Examining the Mental Imagery Debate with Neuropsychological Data from the Clock Drawing Test,. Visual Representations and Reasoning A workshop of the 24th AAAI Conference on Artificial Intelligence (AAAI-10) 26-32.

38. **Harvan, J. R., and V. Cotter.** 2006. An evaluation of dementia screening in the primary care setting. Journal of the American Academy of Nurse Practitioners **18:**351-360.

39. **Hebert, L., P. Scherr, J. Bienias, D. Bennett, and D. Evans.** 2003. Alzheimer's disease in the U.S. population: Prevalence estimates using the 2000 Census. Archives of neurology **60:**1119-1122.

40. **Heinik, J., P. Werner, T. Dekel, I. Gurevitz, and S. Rosenblum.** 2010. Computerized kinematic analysis of the clock drawing task in elderly people with mild major depressive disorder: an exploratory study. International psychogeriatrics / IPA **22:**479-488.

41. **Hesse, B. W., C. Hanna, H. A. Massett, and N. K. Hesse.** 2010. Outside the box: will information technology be a viable intervention to improve the quality of cancer care? Journal of the National Cancer Institute. Monographs **2010:**81-89.

42. **Hodges, J. R. (ed.).** 2007. Cognitive Assessment for Clinicians, 2 ed. Oxford University Press, USA.

43. **Iliffe, S., L. Robinson, C. Brayne, C. Goodman, G. Rait, J. Manthorpe, and P. Ashley.** 2009. Primary care and dementia: 1. diagnosis, screening and disclosure. International journal of geriatric psychiatry **24:**895-901.

44. **Iracleous, P., J. X. Nie, C. S. Tracy, R. Moineddin, Z. Ismail, K. I. Shulman, and R. E. Upshur.** 2010. Primary care physicians' attitudes towards cognitive screening: findings from a national postal survey. International journal of geriatric psychiatry **25:**23-29.

45. **Ismail, Z., T. K. Rajji, and K. I. Shulman.** 2010. Brief cognitive screening instruments: an update. Journal of Geriatric Psychiatry **25:**111-120.

46. **Jacobs, D. M.** 1998. The role of neuropsychological testing in neurological disease. Medical Update for Psychiatrists **3:**139-143.

47. **Kay, G. G., and V. N. Starbuck.** 1997. Computerized Neuropsychological Assessment. *In* M. E. Maruish and J. A. Moses (ed.), Clinical Neuropsychology: Theoretical Foundations for Practitioners, 1 ed. Psychology Press.

48. **Kim, H.** 2011. Exploring technological opportunities for cognitive impairment screening. Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI2011)**:**887-892.

49. **Kim, H., Y. Cho, and E. Do.** 2011. Using Pen-Based Computing in Technology for Health Human-Computer Interaction. Users and Applications, p. 192-201. *In* J. Jacko (ed.), vol. 6764. Springer Berlin / Heidelberg.

50. **Kim, H., Y. S. Cho, and E. Y.-L. Do.** 2011. Computational clock drawing analysis for cognitive impairment screening. Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction**:**297-300.

51. **Kim, H., Y. S. Cho, and E. Y.-L. Do.** 2010. Context-Bounded Refinement Filter Algorithm: Improving Recognizer Accuracy of Handwriting in Clock Drawing Test. Visual Representations and Reasoning A workshop of the 24th AAAI Conference on Artificial Intelligence (AAAI-10) 53-60.

52. **Kim, H., A. Guha, Y. S. Cho, and E. Y.-L. Do.** 2010. ClockReader: Investigating Senior Computer Interaction through Pen-based Computing, p. 30-33, in CHI Workshop on Senior-Friendly Technologies: Interaction Design for the Elderly, Atlanta, GA, USA.

53. **Kim, H., A. Guha, Y. S. Cho, and E. Y.-L. Do.** 2010. Sketch-Based Screening for Cognitive Impairment Detection: A Human Centered Approach, p. 46-53, in SKCHI - ACM CHI 2010 Workshop of Designing Sketch Recognition Interfaces: The Past, Current, and Future, Atlanta, GA, USA.

54. **Kim, H., C.-P. Hsiao, and E. Y.-L. Do.** 2012. Home-based Computerized Cognitive Assessment Tool for Dementia Screening. Journal of Ambient Intelligence and Smart Environments **4:**429-422.

55. **Kluger, B. M., L. V. Saunders, W. Hou, C. W. Garvan, S. Kirli, D. B. Efros, Q. A. Chau, G. P. Crucian, G. R. Finney, K. J. Meador, and K. M. Heilman.** 2009. A brief computerized self-screen for dementia. J Clin Exp Neuropsychol **31:**234-244.

56. **Korczyn, A. D., and V. Aharonson.** 2007. Computerized methods in the assessment and prediction of dementia. Current Alzheimer research **4:**364-369.

57. **Lee, A. Y., J. S. Kim, B. H. Choi, and E. H. Sohn.** 2009. Characteristics of clock drawing test (CDT) errors by the dementia type: quantitative and qualitative analyses. Archives of gerontology and geriatrics **48:**58-60.

58. **Leposavic, I., L. Leposavic, and B. Saula-Marojevic.** 2010. Neuropsychological assessment: computerized batteries or standard tests. Psychiatria Danubina **22:**149-152.

59. **Letz, R.** 1991. Use of computerized test batteries for quantifying neurobehavioral outcomes. Environmental health perspectives **90:**195-198.

60. **Levey, A., J. Lah, F. Goldstein, K. Steenland, and D. Bliwise.** 2006. Mild cognitive impairment: an opportunity to identify patients at high risk for progression to Alzheimer's disease. Clinical therapeutics **28:**991-1001.

61. **Lezak, M. D.** 1995. Neuropsychological assessment, 3rd ed. Oxford University Press, New York.

62. **Libon, D., R. Swenson, E. Barnoski, and L. Sands.** 1993. Clock drawing as an assessment tool for dementia. Archives of Clinical Neuropsychology **8:**405-415.

63. **Luciana, M.** 2003. Practitioner review: computerized assessment of neuropsychological function in children: clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). Journal of child psychology and psychiatry, and allied disciplines **44:**649-663.

64. **Lumen, N., V. Fonteyne, G. De Meerleer, P. De Visschere, P. Ost, W. Oosterlinck, and G. Villeirs.** 2012. Screening and early diagnosis of prostate cancer: an update. Acta clinica Belgica **67:**270-275.

65. **Middleton, L. E., and K. Yaffe.** 2009. Promising strategies for the prevention of dementia. Archives of neurology **66:**1210-1215.

66. **Milne, A., A. Culverwell, R. Guss, J. Tuppen, and R. Whelton.** 2008. Screening for dementia in primary care: a review of the use, efficacy and quality of measures. International psychogeriatrics / IPA **20:**911-926.

67. **Mittal, C., S. Gorthi, and S. Rohatgi.** 2010. Early cognitive impairment: Role of clock drawing test. Medical Journal Armed Forces India **66:**25-28.

68. **Morris, J. C.** 2005. Mild cognitive impairment and preclinical Alzheimer's disease. Geriatrics **Suppl:**9-14.

69. **Peat, J. K., C. Mellis, K. Williams, and W. Xuan.** 2002. Health science research : a handbook of quantitative methods. SAGE, London ; Thousand Oaks.

70. **Petersen, R. C.** 2003. Mild cognitive impairment : aging to Alzheimer's disease. Oxford University Press, Oxford ; New York.

71. **Petersen, R. C., J. C. Stevens, M. Ganguli, E. G. Tangalos, J. L. Cummings, and S. T. DeKosky.** 2001. Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology **56:**1133-1142.

72. **Pinto, E., and R. Peters.** 2009. Literature review of the Clock Drawing Test as a tool for cognitive screening. Dement Geriatr Cogn Disord **27:**201-213.

73. **Pittman, J. A.** 2007. Handwriting Recognition: Tablet PC Text Input. Computer **40:**49-54.

74. **Poreh, A. M.** 2006. The quantified process approach to neuropsychological assessment, 1st ed. Taylor & Francis, New York.

75. **Poreh, A. M.** 2000. The quantified process approach: an emerging methodology to neuropsychological assessment. The Clinical neuropsychologist **14:**212-222.

76. **Powers, R. E., J. W. Ashford, and S. Peschin.** 2008. Memory Matters. Alzheimer's Foundation of America.

77. **PreventedAD.com**, Progression of Cognitive Impairment.

78. **Rosenberg, P. B., D. Johnston, and C. G. Lyketsos.** 2006. A clinical approach to mild cognitive impairment. The American journal of psychiatry **163:**1884-1890.

79. **Rosenblum, S., T. Dekel, I. Gurevitz, P. Werner, and J. Heinik** 2010. Handwriting process variables among elderly people with mild Major Depressive Disorder: A preliminary study. Aging: Clinical and Experimental Research **22:**141-147.

80. **Rosenblum, S., P. Weiss, and S. Parush.** 2003. Product and Process Evaluation of Handwriting Difficulties. Educational Psychology Review **15:**41-81.

81. **Rosenblum, S., and P. Werner.** 2006. Assessing the handwriting process in healthy elderly persons using a computerized system. Aging clinical and experimental research **18:**433-439.

82. **Royall, D. R., J. A. Cordes, and M. Polk.** 1998. CLOX: an executive clock drawing task. Journal of neurology, neurosurgery, and psychiatry **64:**588-594.

83. **Ryan, J. J., L. A. Glass, J. M. Bartels, and A. M. Paolo.** 2009. Base rates of "10 to 11" clocks in Alzheimer's and Parkinson's disease. International journal of neuroscience **119:**1261-1266.

84. **Schatz, P., and J. Browndyke.** 2002. Applications of computer-based neuropsychological assessment. The Journal of head trauma rehabilitation **17:**395-410.

85. **Schomaker, L. S.** 1998. From Handwriting Analysis to Pen-Computer Applications. IEE Electronics Communication Engineering Journal **10:**93--102.

86. **Shankle, W. R., and D. G. Amen.** 2005. Preventing Alzheimer's: Ways to Help Prevent, Delay, Detect, and Even Halt Alzheimer's Disease and Other Forms of Memory Loss. Perigee Trade

87. **Shulman, K.** 2000. Clock drawing: Is it the ideal cognitive screening test? . International journal of geriatric psychiatry **15:**548-561.

88. **Shulman, K. I., and A. Feinstein.** 2003. Quick cognitive screening for clinicians : mini mental, clock drawing and other brief tests. Martin Dunitz, London ; New York.

89. **Siek, K. A., S. E. Ross, D. U. Khan, L. M. Haverhals, S. R. Cali, and J. Meyers.** 2010. Colorado Care Tablet: the design of an interoperable Personal Health Application to help older adults with multimorbidity manage their medications. J Biomed Inform **43:**S22-26.

90. **Silagy, C., and A. Haines (ed.).** 2001. Evidence-based practice in primary care, 2 ed. BMJ Books.

91. **Smith, A.** 2009. On the use of drawing tasks in neuropsychological assessment. Neuropsychology **Mar;23:**231-239.

92. **Smith, R. A., S. W. Duffy, and L. Tabar.** 2012. Breast cancer screening: the evolving evidence. Oncology (Williston Park) **26:**471-475.

93. **Snellgrove, C. A.** 2005. Cognitive Screening for the Safe Driving Competence of Older People with Mild Cognitive Impairment or Early Dementia.

94. **Solomon, P., and C. Murphy.** 2005. Should we screen for Alzheimer's disease? A review of the evidence for and against screening Alzheimer's disease in primary care practice. Geriatrics **60:**26-31.

95. **Souder, E., P. O'Sullivan, and G. Pechenik.** 1999. Comparison of Scoring Criteria for Clock Drawing Test. Journal of Clinical Geropsychology **5:**139-145.

96. **Strauss, E., E. M. S. Sherman, and O. Spreen.** 2006. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary, 3 ed. Oxford University Press, USA.

97. **Sunderland, T., and National Institutes of Health (U.S.).** 1989. Clock drawing in Alzheimer's disease : a novel measure of dementia severity. National Institutes of Health, Bethesda, Md.

98. **Trenkle, D. L., W. R. Shankle, and S. P. Azen.** 2007. Detecting cognitive impairment in primary care: performance assessment of three screening instruments. Journal of Alzheimer's disease : JAD **11:**323-335.

99. **Tullis, T., and B. Albert.** 2008. Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Morgan Kaufmann.

100. **Tuokko, H., T. Hadjistavropoulos, J. A. Miller, and B. L. Beattie.** 1992. The Clock Test: a sensitive measure to differentiate normal elderly from those with Alzheimer disease. Journal of the American Geriatrics Society **40:**579-584.

101. **Unlu, A., R. Brause, and K. Krakow.** 2006. Handwriting analysis for diagnosis and prognosis of parkinson's disease, p. 441-450, Proceedings of the 7th international conference on Biological and Medical Data Analysis. Springer-Verlag, Thessaloniki, Greece.

102. **Wadley, V. G., O. Okonkwo, M. Crowe, D. E. Vance, J. M. Elgin, K. K. Ball, and C. Owsley.** 2009. Mild cognitive impairment and everyday function: an investigation of driving performance. Journal of geriatric psychiatry and neurology **22:**87-94.

103. **Werner, P., S. Rosenblum, G. Bar-On, J. Heinik, and A. Korczyn.** 2006. Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. The journals of gerontology. Series B, Psychological sciences and social sciences **61:**P228-236.

104. **Wilson, J. M. G., and G. Jungner.** 1968. Principles and practice of screening for disease. World Health Organization, Geneva.

105. **Woo, E.** 2008. Computerized Neuropsychological Assessments. CNS Spectr **13:**14-17.