

Designing the ClockSketch User Experience

by

Katherine Anna Hobbs

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

Master of Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

©Massachusetts Institute of Technology 2015. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document
in whole or in part in any medium now known or hereafter created.

Author
Katherine Anna Hobbs
Department of Electrical Engineering and Computer Science
September 16, 2015

Certified by.....
Randall Davis
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by.....
Dr. Christopher J. Terman
Chairman, Masters of Engineering Thesis Committee

Designing the ClockSketch User Experience

by

Katherine Anna Hobbs

Submitted to the Department of Electrical Engineering and Computer Science
on September 16, 2015, in partial fulfillment of the
requirements for the degree of
Master of Engineering

Abstract

In this thesis, we designed and implemented a user interface for doctors to view patient results from the Clock Drawing Test as computed by the ClockSketch system. We identified four primary goals for the interface: glanceability, familiarity, minimalism, and trustworthiness. Following a spiral-based design process, we systematically built and iterated over interface prototypes with these goals in mind.

Thesis Supervisor: Randall Davis

Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Professor Randall Davis for the continuous support of my MEng, for his patience, motivation, and insightful comments. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my MEng.

I would also like to thank my parents for their endless love, support, and dogged encouragement to work on my thesis even when I didn't want to.

Lastly, I would like to thank Isaac, who married me despite my thesis often coming between us. His continued and unfailing love, support, and understanding underpins my persistence in the graduate career and makes the completion of this thesis possible.

Contents

1	Introduction	11
1.1	Description of ClockSketch	11
1.2	Motivation for a ClockSketch User Interface	12
1.3	Structure of a ClockSketch User Interface	12
2	Related Work	13
2.1	Cognitive Ability Tests	13
2.1.1	Clock Drawing Test	13
2.1.2	MMSE	14
2.1.3	SLUMS	15
2.2	Medical Interface Design	15
2.2.1	Glanceable	16
2.2.2	Familiar	16
2.2.3	Simple/Minimalistic	17
2.2.4	Trustworthy	17
3	Design	19
3.1	Design Goals	19
3.1.1	Glanceable	19
3.1.2	Familiar	20
3.1.3	Simple/Minimalistic	21
3.1.4	Trustworthy	21
3.2	Design Process	22

3.2.1	Notional Dashboard	24
3.2.2	Paper Prototypes	26
3.2.3	Computer Mock-ups	27
3.2.4	Software Implementation	29
3.3	Design Decisions	30
3.3.1	Information Selection	31
3.3.2	Information Visualization	34
3.3.3	Data Modification	36
4	Implementation	39
4.1	System Flow	39
4.2	User Interface	40
4.3	Client-Server Communication	45
5	Evaluation	47
5.1	Procedure	47
5.1.1	Early Evaluation Procedure	47
5.1.2	Later Evaluation Procedure	48
6	Future Work and Conclusion	49
6.1	Future Work	49
6.2	Conclusion	50

List of Figures

1-1	Information Tree	12
3-1	Spiral-Based Design Process	23
3-2	Notional Dashboard	25
3-4	Computer Mock-ups	28
3-5	Final Computer Mock-ups	29
3-6	Initial Software Implementation	30
3-7	Information Tree	35
4-1	Patient Selector and Editor	41
4-2	Patient Editor	42
4-3	Patient Report	44
4-4	Scoring System Selector	44

Chapter 1

Introduction

The Clock Drawing Test [2, 11, 22] is a pen-and-paper test commonly administered by health professionals to screen for a variety of cognitive impairment, including Alzheimer’s disease and other dementias. The test first asks a patient to draw an analog clock displaying the time “10 minutes after 11” (called the Command clock) and then asks them to copy a pre-drawn clock showing that same time (the Copy clock). Using one of many possible scoring systems [7], doctors score the test, looking for telling errors in the drawings.

1.1 Description of ClockSketch

ClockSketch is a digital version of the Clock Drawing Test: patients perform the same test with a digitizing ballpoint pen, and their drawing is analyzed and scored by a computer based on work by Souillard-Mandar et al. [23]. ClockSketch can score tests according to seven commonly used manual scoring systems: Manos [13], Royall [19], Shulman [22], Rouleau [18], Mendez [16], MiniCog [5], and Libon [12]. The goal of ClockSketch is to provide a consistent and objective evaluation of a subject’s cognitive abilities using the Clock Drawing Test. Ideally, a doctor could rely on ClockSketch to provide a thorough and accurate analysis of and score for the subject’s cognitive abilities.

1.2 Motivation for a ClockSketch User Interface

Before this thesis, ClockSketch contained all the infrastructure for analyzing and scoring tests, but no simple means for displaying test scores and analytics. There already existed an interface for administering the Clock Drawing Test, but ClockSketch lacked a doctor-facing interface for viewing test results. Without a graphical means to review the test results, ClockSketch is of little use to doctors.

1.3 Structure of a ClockSketch User Interface

The information the ClockSketch UI needs to present falls naturally into a hierarchical, tree-like structure, which we visualized in Figure 1-1. We will explain the components of this hierarchy in Chapter 3.

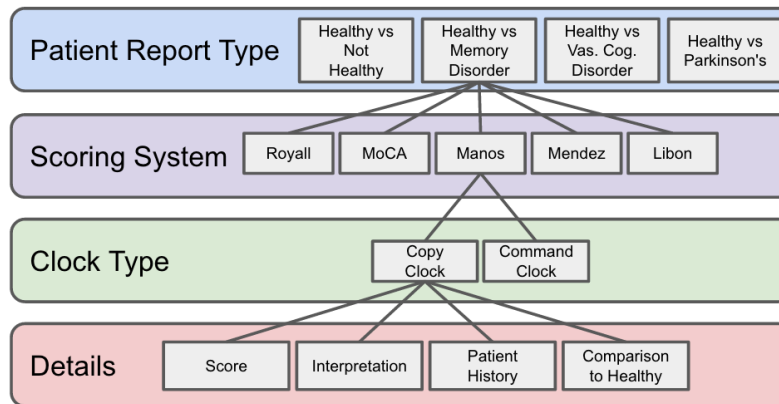


Figure 1-1: Information Tree

This thesis outlines the process of designing, building, and evaluating a doctor-facing graphical user interface to be used in conjunction with the ClockSketch program.

Chapter 2

Related Work

To the best of our knowledge, the ClockSketch interface is the first UI developed to present a doctor-oriented view of the output of automated Clock Drawing Test analysis. We examine two areas of related work: cognitive ability tests and medical-specific interface design.

2.1 Cognitive Ability Tests

There are a number of tests other than the Clock Drawing Test that doctors use to assess a patient's cognitive abilities. Two commonly used tests are the mini-mental state exam (MMSE) and the Saint Louis University Mental Status Exam (SLUMS). Both of these tests produce information useful for determining a patient's cognitive state, and doctors may use multiple tests to screen a patient. Understanding the results from the Clock Drawing Test and these other tests helped us determine what information doctors need to assess a patient's cognitive ability, which aided the design process of the ClockSketch user interface.

2.1.1 Clock Drawing Test

As the ClockSketch doctor-facing interface is intended to display the results of the Clock Drawing Test, it was necessary to understand the test itself and how it is

currently used by medical professionals. Via a remote consultation with the Alliance Medical Center in North Carolina, we observed a doctor manually administering the Clock Drawing Test in an effort to determine which parts of the test and test results doctors need in order to make a diagnosis. Typically doctors watched a patient perform the test, reviewed the clock drawings, assigned a score to each of them using a manual scoring system, and then interpreted the assigned scores using predefined cutoffs.

Since ClockSketch can score the clock drawings for doctors, we wanted to find out if doctors still needed to see the clock drawings in order to make a diagnosis. During user evaluations, when asked to make a diagnosis given only a patient’s numerical test scores, we found that doctors were reluctant to say anything before first reviewing the actual clock drawings. They also expressed their desire to be able to actually watch the patient perform the test. This taught us that doctors not only need to see a patient’s clock drawings on the interface, but they also need to be able to replay them and watch while they are “drawn.” This was useful information while designing the ClockSketch doctor-facing interface, and we made sure to feature the patient clock drawings in a prominent location.

2.1.2 MMSE

Commonly used in place of the Clock Drawing Test, the MMSE [8] is a 30-question survey intended to assess registration, attention and calculation, recall, language, and ability to follow simple commands and orientation. The test is administered verbally and an overall score is computed from the sum of the scores for each category. Doctors can use individual subscores to determine which type of dementia a patient tends toward. For instance, Alzheimer’s patients tend to score lower on orientation and recall. It is notable that doctors use the individual subscores—such as orientation and recall scores—of the MMSE test. Similarly to the MMSE category scores, some of the scoring systems used for the Clock Drawing Test have subscores, which may be of use to doctors in addition to their overall scores. Although the display of subscores was not included in this version of the ClockSketch doctor-facing interface, it is a

planned future feature in the next version. Examining the MMSE helped us identify test subscores as potentially useful information to display on the interface.

2.1.3 SLUMS

The SLUMS [21] is an 11-question test that includes the Clock Drawing Test. It seeks to identify patients with very early Alzheimer’s symptoms by testing orientation, short-term memory, calculations, naming of animals, clock drawing, and recognition of geometric figures. Overall scores for the SLUMS are simply the sum of the points received for each question. Doctors use a patient’s education level and overall score to make a diagnosis of normal, mild cognitive disorder, or dementia. We noted the use of semi-static patient information such as education level, ethnicity, handedness, etc. for diagnosis and discussed its utility with medical professionals. During user evaluations, doctors highly encouraged incorporating semi-static patient information in the ClockSketch doctor-facing interface. The SLUMS helped us identify semi-static patient information as a key feature to include in the ClockSketch doctor-facing interface.

2.2 Medical Interface Design

There are a multitude of digital programs used in the medical and health industries, each of which have their own unique interfaces. And each of these interfaces have different goals for their user experience. We interviewed doctors and other medical professionals to identify the design goals for the ClockSketch doctor-facing user interface. Then we examined a number of medical interfaces with shared design goals in an effort to determine the optimal way to achieve our design goals.

Our design goals (further described in Chapter 3) are for the interface to be glanceable, familiar, simple/minimalistic, and trustworthy. Below we describe what we learned by investigating other medical interfaces with similar design goals.

2.2.1 Glanceable

Glanceability is a common design goal for program interfaces that are intended to provide rapid and immediate feedback while the user completes some sort of action. For example, TripleBeat [6] is a program that provides real-time feedback about a runner’s physiological responses through a glanceable mobile interface. This is a very different use case than what is intended for ClockSketch, but since glanceability is one of our design goals, it was useful to study interfaces for real-time feedback programs to learn how best to make the ClockSketch interface glanceable.

Color can be a quick and easy way to alert users and make them pay attention to that part of the interface [14]. TripleBeat uses color to tell runners whether or not they are running fast enough to meet their goals. This is where we got the idea to use colors to distinguish failing scores—since failing scores are by-and-large the most important thing for a doctor to know about, we want them to “stick out” when a doctor glances at the interface. Healthy scores are not as important as unhealthy scores—if a patient scored in the healthy range for all of the scoring systems except one, we want that one failing score to be the first thing a doctor sees. TripleBeat helped us identify color as a tool for implementing glanceability.

2.2.2 Familiar

In their 1995 paper “User interface design and evaluation for electronic medical record system” [3], Aisaka et al. discuss the design for a user interface for an electronic medical record system (EMRS). One of their main design goals is familiarity, which they achieve through a windows layout that simulates conventional paper forms, so that doctors can quickly grow accustomed to the EMRS. Since the ClockSketch interface does not involve a lot of user input, it was difficult to incorporate this idea. However, since doctors are most familiar with pen and paper clock drawings, we made sure to include images of the drawings in the interface.

2.2.3 Simple/Minimalistic

Sage Research & Design built an interface for a medical application [1] used to program a medical device surgically implanted in patients. It is a very complex tool used by several different types of users (surgeons, clinicians, and customer representatives), so one of their primary goals was for it not to require any training to use. Sage Research & Design performed user research and found that a task-based flow through the interface was the optimal way to achieve their goal; users are asked what they would like to do and then moved through the appropriate sequence of screens. It is difficult to apply this principle to the ClockSketch interface since it is not intended to guide the user through a process, but we did note the simplification of the Sage Research & Design interface through the use of tabs that hide extra information. We also identified graphics as a simplification solution, and we chose to implement tabs and graphics in the ClockSketch interface.

2.2.4 Trustworthy

According to “Trustworthy Medical Device Software” by Kevin Fu [9], trustworthiness is “a system property measuring how well a software system meets requirements such that stakeholders will trust in the operation of the system.” This is exactly what we want for the ClockSketch interface—we do not want to give doctors any reason to doubt the results. Fu goes on to say that clearly specifying roles and responsibilities can make a program more trustworthy. We applied this thinking to the ClockSketch interface in the following way: the ClockSketch interface will never make a diagnosis—it will only every provide numerical scores and scoring system-defined interpretations, because it is the doctor’s responsibility to make a diagnosis. This prevents ClockSketch from being blamed for making an incorrect diagnosis. The doctor’s responsibility is clearly to make the diagnosis while ClockSketch’s role is clearly to be a tool, which makes it more trustworthy.

Chapter 3

Design

In this chapter we present the design and the principles that motivated the ClockSketch doctor-facing interface. We spent extensive time learning about the environment the Clock Drawing Test is presented in before we ever put pen to paper to design the first paper prototype. Through an iterative, spiral-based design process, we identified concrete design goals for the interface and made key design decisions.

3.1 Design Goals

We first identified concrete design goals through interviews with doctors and other professionals with a stake in the ClockSketch project. Although they remained largely the same, the design goals evolved somewhat due to feedback during user evaluation. After interviewing a range of medical professionals from dementia specialists who are well-acquainted with the Clock Drawing Test to family practice doctors who are typically colloquially familiar with the Clock Drawing Test, we identified four main goals for the interface:

3.1.1 Glanceable

During the interviews, doctors emphasized their desire for a glanceable interface—they wanted the important parts of the interface to “pop out” at them and be easily

spotted. In other words, users should not need to read everything on the screen in order to get the gist of a patient’s results. Our goal was to include all the details and explanations for any ClockSketch-generated score or interpretation but highlight only the important numbers and figures.

Glanceability proved to be a difficult design goal because doctors also wanted every possible detail included in the interface. Initially, doctors insisted that the interface should have explanations for every number or graph, explaining where it came from, what it meant, and how to interpret it. They also requested a feature rich interface with lots of options and customization. But this proved to be too much—even with different fonts, bright coloring, and other textual emphasis methods, the interface simply had an overwhelming amount of visual information that left it feeling extremely cluttered. And after preliminary user evaluations of basic designs, doctors and developers alike requested a more minimalistic version without all of the explanations and details.

Working with doctors, we found a happy medium; we identified the absolute most important parts of the interface, increased their textual emphasis, and hid the explanations and text in such a way that they were available but only on the screen if the user wanted to see them. We struck a balance, maintaining our initial goal of including all the details and explanations while ensuring that the most important parts are findable at a glance.

3.1.2 Familiar

Our second design goal was for the interface to be familiar, i.e. similar to existing manual Clock Drawing Test results and other digitized test results that doctors already use. This goal was identified while demoing an early prototype for the interface—doctors did not specifically request that the interface be similar in appearance to what they currently use, but they disliked when the interface did not have images of a patient’s clock drawings, which doctors are used to seeing during diagnosis.

Another reason for our familiarity design goal is that users are able to learn a new interface much faster if it resembles an interface they have used in the past [17]. If

the interface resembles other interfaces that the doctor is familiar with, it will take less time for the doctor to come up to speed and start using it. For these reasons, one of our design goals is for the interface to maintain visual similarities with manual Clock Drawing Test results.

Through studying interfaces with familiarity as a design goal and observing doctors interact with the interface, we determined how best to make our interface familiar. We achieved familiarity in our interface by including video replays of a patient’s clock drawings and scores calculated using the most popular scoring systems, with which doctors are likely to be familiar.

3.1.3 Simple/Minimalistic

Another goal, which overlaps with the glanceable goal, is for the interface to be simple. Doctors insisted that the interface must not require a lot of training to use. Doctors do not have time to be trained to use a complicated new interface. Additionally, doctors do not have a lot of time to spend with each patient and need to be able to review a patient’s results and diagnose them quickly and efficiently. Our goal is for the interface to be simple enough that it requires little to no training to understand and use.

We achieved this goal through much the same means as we achieved the glanceable goal: we added menu bars, tabs, and tooltips to reduce the amount of information on the screen at any one time. Additionally, through our study of related work performed by Sage Research & Design (see Section 2.2.3), we learned that graphics can aid interface simplification, so we also included graphics in our interface.

3.1.4 Trustworthy

Our final goal is for the interface to be trustworthy; it must give doctors no reason to doubt the results. Doctors may not trust a new system with “magic numbers”—they will want to know where the numbers are coming from and the reasoning behind them (at least until they are comfortable with it and have used it for a time and

grown to trust it). This is much of the motivation behind doctors requesting that the interface include every detail and explanation possible (see Section 3.1.1). When a doctor manually scores a patient’s Clock Drawing Test, it is easy for them to point to a specific part of the drawing and say “this is why they lost these points.” It is much more difficult for a computer program to do this, which makes it difficult for someone who did not develop the program to trust it—they do not know where the numbers are coming from.

The area under the curve for the healthy/unhealthy classifier for ClockSketch is 0.91 while the standard algorithms used by the doctors have an AUC of 0.72 on this task. ClockSketch has a higher accuracy rate than doctors do for scoring the Clock Drawing Test [23]. But simply presenting doctors with statistics proving the efficacy of ClockSketch is not enough to convince them that it is reliable and trustworthy; the user interface design itself must increase trustworthiness [20].

As discussed in section 2.2.4, clearing defining roles and responsibilities is also a good way to make the interface trustworthy. We made sure that there is no place in the interface that will label a patient as “healthy” or “unhealthy,” because it is the doctor’s responsibility to assign that label. This way ClockSketch is never responsible for an incorrect diagnosis.

We achieved interface trustworthiness by including explanations for every score, interpretation, and graph displayed on the interface. We also included videos of a patient’s clock drawings so that doctors can manually verify the scores a patient received if they do not trust the scores computed by ClockSketch. Implementing these features made our interface trustworthy.

3.2 Design Process

We approached the design process iteratively, electing to use a spiral model (see Figure 3-1) as opposed to a waterfall model. The spiral model [4] is an iterative process model that builds on itself, refining a project over and over again. Within every iteration we underwent four main phases:

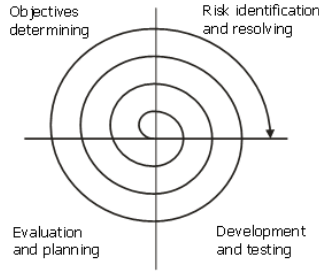


Figure 3-1: Spiral-Based Design Process

1. **Determine Objectives.** During this phase, we interviewed medical professionals who commonly administer the Clock Drawing Test and the developers of ClockSketch. We identified both party’s goals for the system and made plans to reconcile them into one cohesive set of system goals (see Section 3.1 for a list of goals we identified). During this phase we also performed cost-benefit analyses for the addition, removal, or modification of specific aspects of the user interface based on feedback and observations during user evaluations.
2. **Identify and Resolve Risks.** During this phase, we identified and resolved risks and evaluated alternatives relative to the objectives. This phase sometimes involved additional interviews and discussion with doctors and developers in order to reach a solution acceptable by all parties.
3. **Development and Testing.** During this phase, we designed, implemented, and evaluated the interface. In the first few iterations we focused on designing and redesigning a low-fidelity prototype to meet the system objectives and mitigate the identified risks. In later iterations we implemented a refined design as a computer prototype and iterated on the computer prototype design until we reached an optimal design. During our final iterations, we implemented our computer prototype in software.
Throughout the design process, we tested our interface through evaluations with ClockSketch developers and doctors. For more information on the evaluation process, see Chapter 4.
4. **Plan the Next Iteration.** During this phase, we reviewed our notes from the

previous three phases and determined what we intended to accomplish over the next iteration. We refined our objectives, risk mitigation, and test plan based on the findings from this iteration and all prior iterations.

During our numerous design iterations we built everything from low-fidelity paper prototypes to high-fidelity software implementations. We progressed through four distinct prototypes: a notional dashboard jointly-designed by doctors and ClockSketch developers, paper prototypes, computer mock-ups, and finally a software implementation. We spent many iterations on each prototype before progressing to the next prototype. We now discuss in detail how user testers reacted to prototypes and how they evolved due to feedback.

3.2.1 Notional Dashboard

The notional dashboard (Figure 3-2) was an initial draft created by doctors and developers before we began work on the ClockSketch doctor-facing interface. It provided a jumping off point to springboard the design process. We studied the notional dashboard and interviewed the doctors and developers who designed it to better understand what they wanted.

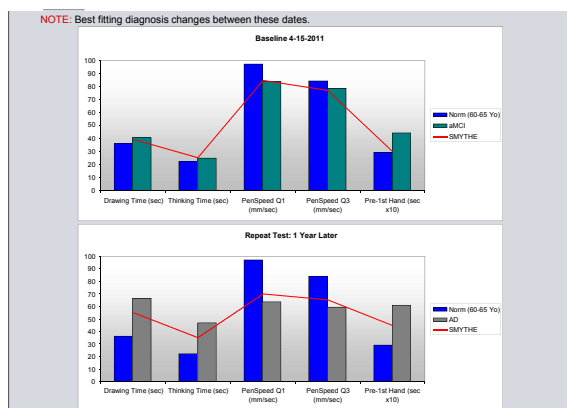
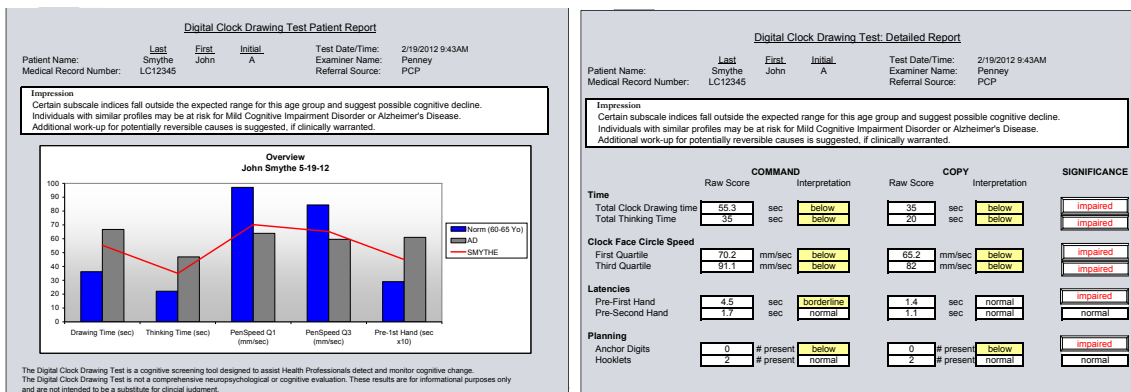


Figure 3-2: Notional Dashboard

As evidenced by the sheer amount of data present, the doctors and developers that designed the notional dashboard wanted to have access to every number and figure ClockSketch could provide. But through discussion we learned that doctors wanted the interface to be more compact—they wanted everything to be on viewable on one screen instead of three. Doctors conceded that the amount of information displayed in the notional dashboard would not all fit on one screen, but insisted it be present in some form on the interface.

While reviewing the notional dashboard we considered how well it achieved our design goals: The dashboard is not very glanceable—users cannot quickly identify whether a patient received mostly failing scores or not. However, the use of color to identify below average interpretations and significances helps somewhat, but the colors are not consistent—both yellow and red are used. The dashboard also fails

the familiarity goal; there are no clock drawings or even scores from common scoring systems for doctors to reference. Although the graphics do help simplify it somewhat, the dashboard is not very simple because it has so much information on it. Lastly, the dashboard is not trustworthy because it has no way to gain doctors’ trust—it only presents “magic” numbers without any explanations.

Following our review of the notional dashboard, we identified aspects of the dashboard we planned to keep, change, or remove during paper prototyping: We decided to keep the comparison of the patient to healthy patients in the graphics. We decided to condense the dashboard into a single screen during paper prototyping, add images of clock drawings, and add explanations for scores and interpretations. We did not like the fact that graphs were placed side-by-side, because it made it too easy for users to compare scores that should not be compared, so we decided to rearrange non-comparable graphs. We also decided to make the color scheme consistent across the interface, making it obvious which scores and interpretations are below average. Lastly, we decided not to display all of the numbers for all of the scoring systems at one time because it is overwhelming for the user to be presented with so much data at one time.

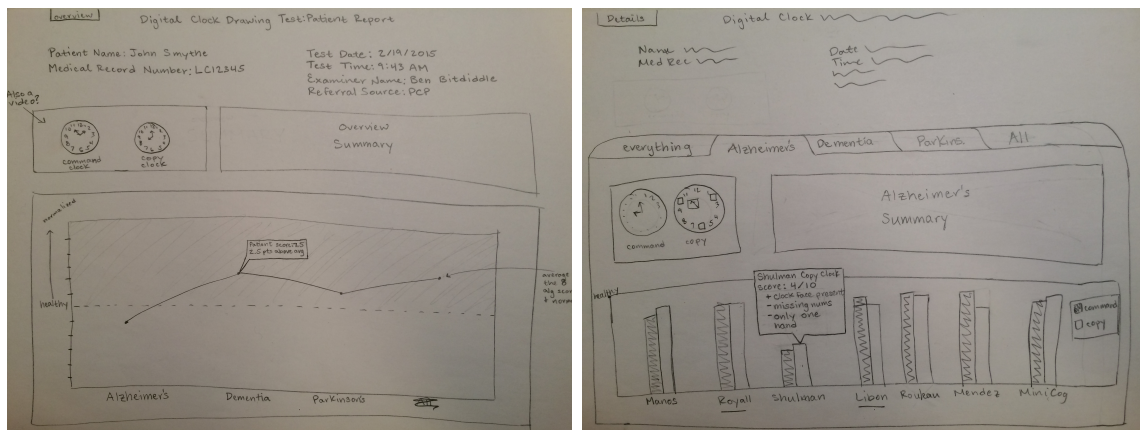
3.2.2 Paper Prototypes

After extensive review of the notional dashboard, we began designing paper prototypes based loosely on the notional dashboard. Through iterations, our paper prototype evolved quite a bit.

Our first paper prototype (Figure 3-3a) incorporated many of the positive aspects of the notional dashboard and added a few additional features too. For instance, it included a comparison between the patient and healthy patients, images of the clock drawings, a purely graphical representation of patient scores, and it was condensed into one screen. We showed the prototype to doctors and the ClockSketch developers and got a lot of useful feedback. Doctors appreciated the images of the clock drawings and found them to be a useful reference. Doctors and developers alike did not like how the graph displayed an averaged score—they wanted to see the individual scores

from every scoring system. Doctors also wanted to see explanations for the scores. We took all of these comments into consideration during the next few iterations of the paper prototype and eventually ended up with our final paper prototype (Figure 3-3b).

Our final paper prototype displayed the individual scores for all of the scoring systems just as doctors had requested. It also included tooltips with explanations for every score. Overall, doctors and developers were pleased with the prototype, but wanted to see if they still liked it after it was mocked up on a computer.



(a) Early Paper Prototype

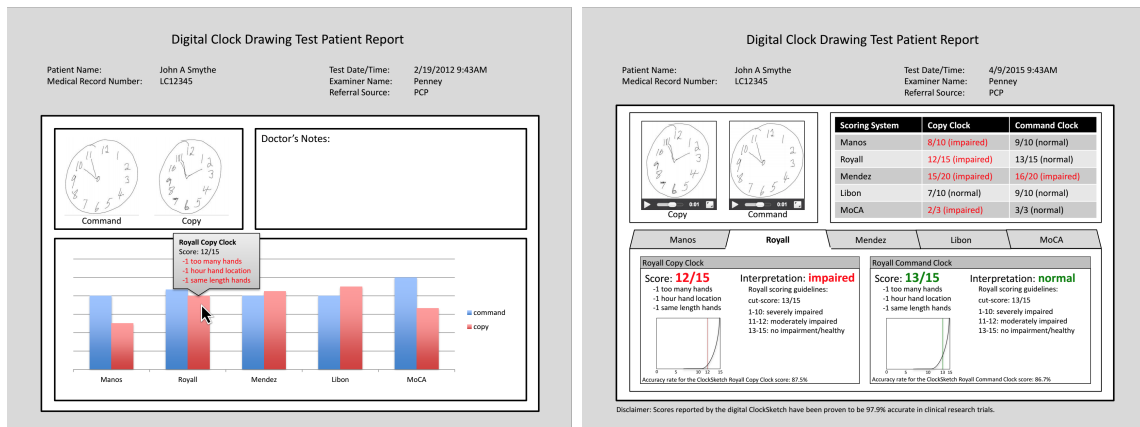
(b) Later Paper Prototype

3.2.3 Computer Mock-ups

We spent the longest amount of time iterating over the computer mock-ups. Since they were rather high-fidelity prototypes, it was easier for doctors to evaluate them and give us meaningful feedback.

Our first computer mock-up (Figure 3-4a) was almost an exact replica of the final paper prototype. We met with doctors and performed informal user evaluations to learn how to improve the interface. Doctors did not like the comment box in the upper right-hand corner of the interface because they did not have time to write anything extra. Doctors also said they would prefer to have scores represented numerically rather than graphically. We also met with ClockSketch designers who requested a video of the clock drawings in place of images. We took all of this into consideration while designing the second computer mock-up.

The second computer mock-up (Figure 3-4b) removed the score graph and replaced it with a tabbed panel displaying information for only one scoring system at a time. We also removed the comment box and replaced with a summary table containing scores and interpretations for every scoring system. In response to the designers' request, we replaced the clock drawing images with clock drawing videos. Lastly, we added a graph comparing the patient to healthy patients. We met with doctors and developers again and learned that doctors want to be able to see a patient's history while making a diagnosis. Also, some doctors did not want to see results from the scoring systems they did not regularly use. Lastly, developers wanted a way to switch between patient report types. We notes all of the comments and implemented them in our final computer mock-up



(a) computer mock-up 1

(b) computer mock-up 2

Figure 3-4: Computer Mock-ups

Our final computer mock-up (Figure 3-5) included patient history graphs, the ability to add or remove scoring systems through a buttons on the summary table or a popup window, access to different report types via the menubar, and tooltip score and interpretation explanations. Doctors and developers were both pleased with the final computer mock-up and wanted to see it implemented.

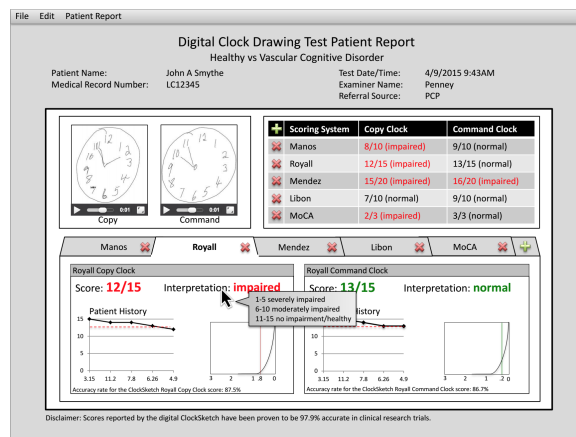
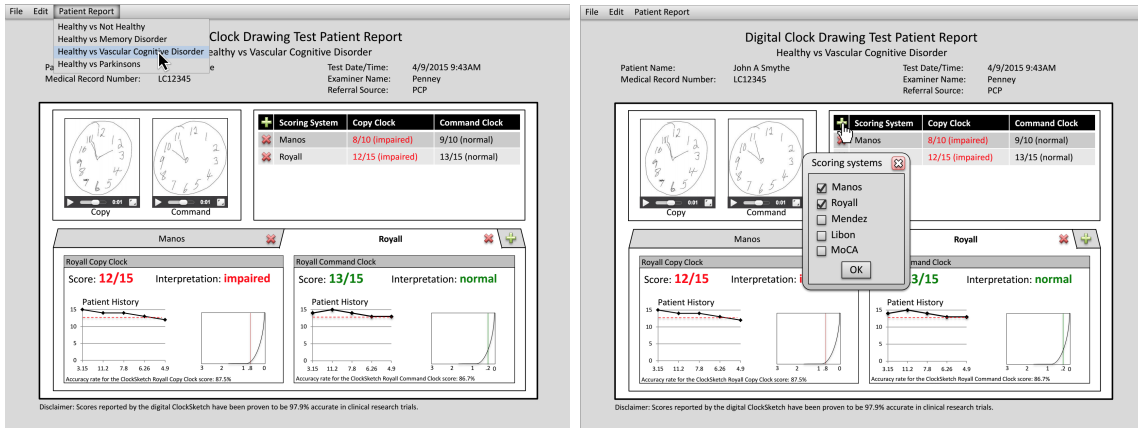


Figure 3-5: Final Computer Mock-ups

3.2.4 Software Implementation

The final few iterations focused on the software implementation, which, thanks to a multitude of iterations on the paper prototypes and computer mock-ups, took a relatively short time. The most difficult aspect of the software implementation stage was simply determining how to implement some of the features we designed for the computer mock-ups.

Our initial software implementation (Figure 3-6) closely resembles the final software implementation. The main differences are sizing and coloration due to operating system configurations. We met with doctors and developers several times between the initial and final software implementations and received a lot of feedback. Doctors

wanted the patient history graph to have background shading to make it easier to interpret how well a patient had performed in the past. Developers requested identical axes on all graphs of the same type that were viewable on one screen. Lastly, doctors and developers wanted a landing page that would appear when the program was started, asking the user to select a patient and verify that their information is correct before being able to view their patient report. We implemented all of these features in our final software implementation which we will discuss in Chapter 4.

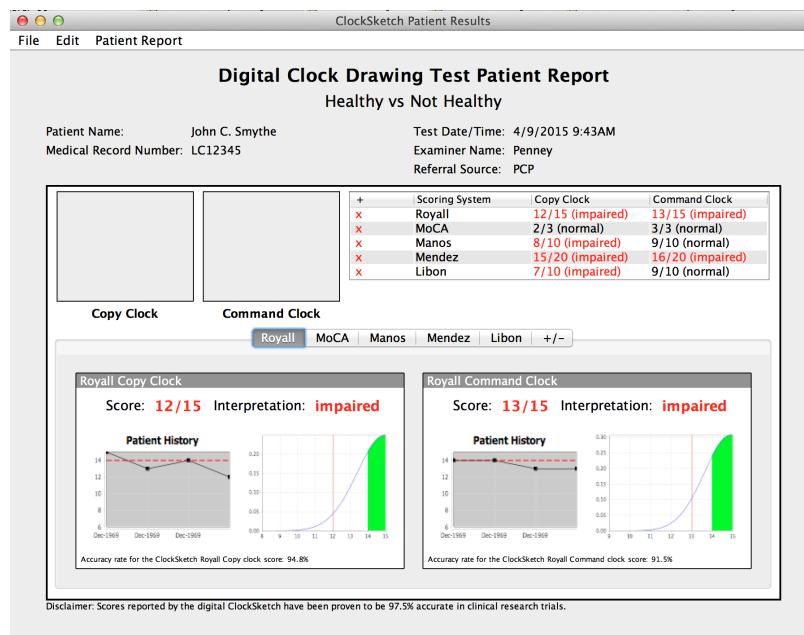


Figure 3-6: Initial Software Implementation

3.3 Design Decisions

Through the design process we came across several contending design options which required us to make key design decisions. We identified three main problem areas which required us to make major design decisions: information selection, information visualization, and data modification.

3.3.1 Information Selection

We identified information selection as a problem area right from the start. As mentioned in the Design Goals section, doctors requested that the interface include every detail and explanation while also maintaining simplicity. Reconciling these two requests meant making a design decision as to what and how much information to include in the interface.

This brings up the question of how well-versed the user is in all things related to the Clock Drawing Test. If we assume that the user is a specialist and understands the scoring mechanisms for all of the provided scoring systems, then there is no need to provide details and explanations for the different features. If, however, the user is a general practitioner who may be only vaguely familiar with the Clock Drawing Test, they will rely on the interface to guide them in understanding the unfamiliar scoring systems and meaning of the scores themselves. Through discussion with the ClockSketch developers and doctor testers, we elected to gear the doctor-facing interface toward medical professionals who are familiar with the Clock Drawing Test, but not necessarily familiar with all the scoring systems because most doctors exclusively use only one or two scoring systems.

In order to determine what information to include in the interface, we made a list of essential features. This list was based on the values, figures, and other information doctors currently review when assessing a patient's Clock Drawing Test. We also asked doctors to identify the vital parts of the interface and included their answers in the list. Together we identified the following essential components for the interface:

1. scores for the copy and command clocks for each scoring system
2. interpretations for the clocks for each scoring system
3. photos of the patient's clock drawings

After we identified the minimum viable features, we began trying to determine what additional, non-required information that doctors do not currently have might be useful. This was somewhat difficult because doctors do not always know what

will help them or are not aware of the possibilities. We realized that it was better to look at what additional information we are now capable of providing doctors with that they could not or did not have access to previously (when they were manually-administering the Clock Drawing Test). We made a list of possible features and then narrowed it down through user evaluations and doctor interviews. Below, we enumerate the additional features we ultimately implemented in the interface and those that did not make the cut.

Included Features:

1. **Clock Drawing Videos.** Doctors loved the idea of being able to replay a patient's clock drawing. This feature is especially useful for doctors who were not present while a patient performed the Clock Drawing Test.
2. **Summary of Scores and Interpretations.** This is an unnecessary feature as all of the information within it is available through other parts of the UI, but it greatly aids glanceability and efficiency of the interface. Doctors can quickly glance over the ever-present summary table for quick score look-up.
3. **Patient History.** Doctors liked the idea of all the patient's history being aggregated in one place. Of course, for this to be a useful feature, ClockSketch must be in use for a significant amount of time and patients must have completed the Clock Drawing Test on multiple occasions. Doctors liked the idea of being able to review a patient's score progression or regression over time.
4. **Comparison to Healthy Patients.** ClockSketch has access to thousands of scores for the Clock Drawing Test which allow us to create graphics showing how a specific patient's score compares to the scores of thousands of healthy patients. Doctors can use this to assess indication of a problem based on the number of healthy patients who received the same score.
5. **Tooltip Explanations.** Doctors wanted to have access to explanations for all of the scores and interpretations, but they did not want the explanations to be visible at all times because they cluttered the interface.

6. **Accuracy Statistics.** The ClockSketch developers requested that statistics stating the accuracy rates for ClockSketch’s scores be included somewhere on the interface to increase trustworthiness.

Excluded Features:

1. **Less Popular Scoring Systems.** There are a lot of possible scoring systems and we had to draw the line somewhere. ClockSketch is capable of scoring the Clock Drawing Test according to seven different mainstream scoring systems. We asked doctors to rank the scoring systems from most to least commonly used in practice and then selected the four most popular scoring systems to include. We also included one new scoring system designed by the developers of ClockSketch. If there is significant support for another scoring system, it can easily be added to the interface. The interface is flexible and allows for easy addition and removal of scoring systems.
2. **Timing Information.** There is a significant amount of temporal information captured by the digital pen used by patients in the digital Clock Drawing Test, which is not currently visualized in the ClockSketch interface. The overall time to complete the sketch, as well as the ”thinking” and ”inking” times between each component, can be calculated from the digital pen data. Displaying this timing information is a feature the developers of ClockSketch would like to introduce in the future, but we want to wait until doctors are comfortable using ClockSketch with familiar scoring systems before it is introduced.
3. **Score Comparison.** Separate scoring systems should *not* be compared because they are on different scales and mean different things. Similarly, a patient’s score for their command clock drawing should not be compared to their score for their copy clock drawing. We purposely placed graphical displays for scores on separate viewing tabs for different scoring systems so they can not be easily compared. We also tried to adequately separate graphical score displays for the copy and command clocks for the same scoring system.

4. **Doctor Notes.** Doctors will likely be jotting down notes elsewhere and will not need to write notes on the interface itself. There should not be two separate places for notes. ClockSketch is simply a tool doctors can use to diagnose patients—if doctors think the ClockSketch-generated scores are incorrect, there is no need for doctors to make a note of it in ClockSketch because it will not affect the recorded score or scoring system.
5. **Overall Score Summary.** ClockSketch is not capable of amalgamating all the data from all the scoring systems and providing a cohesive summary. Also, this would be too much like a diagnosis which doctors should do on their own.

3.3.2 Information Visualization

Design decisions concerning information visualization naturally followed design decisions for information selection. Once we identified the information we wanted to display, we had to figure out how to display it. Below, we explain how we decided to visualize each piece of information we selected to include in the interface:

1. **Clock Drawing Videos.** The main concern with visualizing the clock drawing videos was resizing them to fit within the interface. Patients typically perform the Clock Drawing Test on standard 8.5 by 11 inch paper, so some of the clock drawings can be quite large. We chose to shrink the videos to approximately 1.5 by 1.5 inches, which is still large enough for users to identify the important features of most clock drawings.
2. **Summary of Scores and Interpretations.** We were deciding between a summary table and a summary paragraph like in the notional dashboard (Figure 3-2). We chose to display the score summary in a table as opposed to a paragraph because it is easier for users to scan a table for numbers than a paragraph for words [15].
3. **Patient History.** We visualized patient history as a line chart. Each point on the chart represents a score from a previous Clock Drawing Test in chronological

order. This way doctors can easily identify trends in a patient’s Clock Drawing Test performance over time.

4. **Comparison to Healthy Patients.** We compared the patient to healthy patients through a graph plotting the percentage of healthy patients that received each score. We added a line to the graph representing the patient’s score, which made it easy to see how the patient’s score compared to the scores of healthy patients.
5. **Tooltip Explanations.** In an effort to provide doctors with as many details and explanations as possible without cluttering the interface, we hid score breakdown information and interpretation details in readily accessible tooltips that are hidden by default.
6. **Accuracy Statistics.** We included accuracy statistics for every scoring system and ClockSketch as a whole as footnotes in the interface.

The information we wanted to display fit naturally in a tree-like structure (Figure 3-7), so we began experimenting with the design to determine the most effective way to display a tree of information.

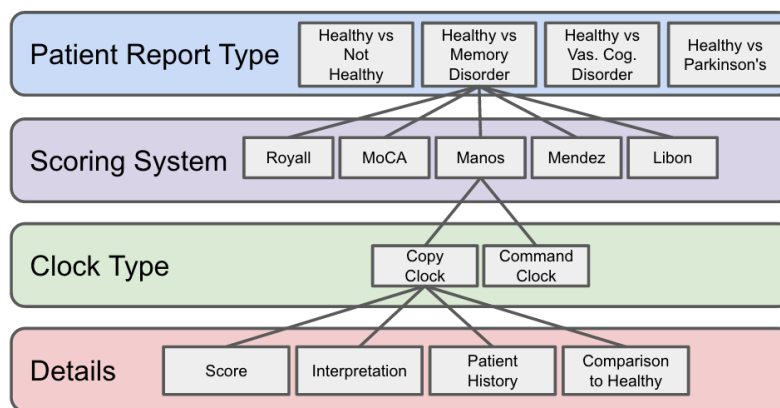


Figure 3-7: Information Tree

We identified branches at the same level within the tree as natural divisions in the interface. And, in an effort to maintain simplicity and reduce clutter throughout the

interface, we chose to only display information from one Patient Report branch and one Scoring System branch of the tree at a time. We determined that doctors were most likely to need to view information at the Scoring System branch level, so we hid information from other Scoring System branches that were not the current primary focus. This forces doctors to focus on the results from one scoring system at a time, without distractors present.

We separated the Patient Report branches through a menu bar drop down that allows users to select the patient report type they would like to view. Within a patient report, the Scoring System branches are separated using tabs for each scoring system. Once a user selects a patient report type, they can then select a scoring system. Only one scoring system can ever be selected at a time—all other scoring systems are hidden. Within a scoring system, all sub-branches are visible: all details pertaining to both the copy and command clock are visible.

To give an example, a doctor might open the Healthy vs. Not Healthy patient report and then click on the Royall tab. At this point the doctor will only see information pertaining to the Healthy vs. Not Healthy patient report and details for the Royall scoring system. The doctor will always be able to see a summary table that includes scores and interpretations for other scoring systems, but they will not see all of the details for other scoring systems. Lastly, the doctor will also always be able to see the the clock drawing videos.

3.3.3 Data Modification

During interviews, doctors wanted to know what they should do if they disagreed with a score computed by ClockSketch. This brought up an interesting design question—should the interface allow data modification? In other words, should doctors be able to “correct” a ClockSketch generated score?

Allowing doctors to modify the data could be useful for the developers of ClockSketch if the modifications were recorded and accessible to developers. Developers could review the modifications and look for trends in order to identify problems and refine ClockSketch’s scoring algorithms. On the other hand, doctors do not need to

be able to change scores because ClockSketch is only a guide for diagnoses—not the final word. The doctor can and should interpret the results however they choose, only relying on ClockSketch as a tool. Also, doctors are not as accurate as ClockSketch at scoring the Clock Drawing Test and they may not be (correctly) taking into account all of the metrics which ClockSketch uses to compute scores.

Developers thought it was unlikely that they would review modification logs and also pointed out the fact that ClockSketch has already proven to be more accurate than doctors at scoring the Clock Drawing Test. So, after weighing the pros and cons and discussing it with developers, we decided not to allow data modification.

Chapter 4

Implementation

In this chapter we present the finished software implementation of the ClockSketch doctor-facing interface. We will provide a walkthrough of the anticipated system flow and we will explain all components of the user interface. Lastly, we will describe our modifications to the underlying ClockSketch system that were made in order to accommodate client-server communication.

4.1 System Flow

We wanted to maintain the look and feel of the existing ClockSketch interface that our doctor-facing interface would plug into, so we chose to implement it using Java Swing just like the rest of the ClockSketch program. We sought to create a seamless transition from entering data for a test to viewing the results of a test.

We envision two different people interacting with ClockSketch. One is a nurse or medical assistant that administers and uploads the test and the other is a doctor who will only ever look at the results interface (which is what we built). Of course, in some cases the same person might perform both jobs. In which case it is important for there to be a seamless transition between the existing ClockSketch interface and our results interface.

To give a better idea of how our interface fits into the system flow, we present the imagined ClockSketch usage scenario from start to finish (as described by the

developers and doctors who first imagined ClockSketch):

1. **Perform Tests.** A nurse or medical assistant administers the Clock Drawing Test in the usual fashion using a digital pen.
2. **Enter Data.** The test administrator starts ClockSketch, opens the patient's file (or creates a new file if one does not already exist), and then uploads the patient's test data from the digital pen. They review the uploaded data and verify that ClockSketch identified features correctly in the patient's clock drawings.
3. **Doctor Review and Diagnosis.** A doctor opens the patient report, peruses it a bit, and decides what to do next (e.g., referral, additional testing, etc.).

4.2 User Interface

Now we will walk through the final software implementation for the results interface.

The results interface is composed of four distinct windows:

1. **Patient Selector and Editor.** The Patient Selector and Editor window (Figure 4-1) is the first thing a user is presented with upon opening the interface. It consists of a browseable textbox for looking up a patient's file and an embedded Patient Editor window (see the next window). Currently the interface is only set up to accept only locally saved files, but it will not be this way when it is actually hooked up to a hospital system. The Patient Selector and Editor is also accessible through the Patient Report window by opening the File menu dropdown and selecting "Open..."

Patient data file:

/Dropbox/mit/meng/mugs/code/patient2.json [Browse...](#)

Patient Information

Patient Name: Alyssa P. Hacker

Birthdate: December 13 1995

Gender: Female Male

Education: High School

Ethnicity: Unk

Race: White/Cauc

Handedness: Left Right

[Open Patient File](#)

Figure 4-1: Patient Selector and Editor

- Patient Editor.** The Patient Editor window (Figure 4-2) allows users to modify patient information. It includes fields for the patient's name, birthdate, gender, education level, ethnicity, race, and handedness. The patient's name and birthdate are uneditable fields. Users can access the Patient Editor window through the Patient Report window by opening the Edit menu dropdown and selecting "Edit Patient Information".

Patient Information

Patient Name: Alyssa P. Hacker

Birthdate: December 13 1995

Gender: Female Male

Education: High School

Ethnicity: Unk

Race: White/Cauc

Handedness: Left Right

Save Patient Info

Figure 4-2: Patient Editor

3. **Patient Report.** The Patient Report window (Figure 4-3) is the primary window in the interface. It is accessible through the Patient Selector and Editor window by selecting a patient file and clicking “Open Patient File”. The Patient Report window houses all data pertaining to Clock Drawing Tests and is intended to be the window users spend the majority of their time looking at. The Patient Report window is composed of five main components:

- (a) **Menubar.** The Menubar component consists of four dropdown menus: File, Edit, Viewing Preferences, and Patient Reports. The File dropdown contains the “Open...” menu item which launches a Patient Selector and Editor window. The Edit dropdown contains the “Edit Patient Report” menu item which launches a Patient Editor window. The Viewing Preferences dropdown has three menu items: “General Practitioner”, “Neurologist”, and “Neuropsychologist”. Selecting a Viewing Preference menu item will dynamically simplify or complexify the Patient Report window by adding/removing parts of the interface. Lastly, the Patient Reports dropdown contains four menu items: “Healthy vs Not Healthy”,

“Healthy vs Memory Disorder”, “Healthy vs Vascular Cognitive Disorder”, and “Healthy vs Parkinson’s”. Selecting a Patient Report menu item will remove all data for the previously selected Patient Report and replaces it with data for the newly selected Patient Report.

- (b) **Patient Information.** The Patient Information component consists of five uneditable textfields displaying the patient’s name, medical record number, test date and time, examiner name, and referral source.
- (c) **Clock Drawing Videos.** The Clock Drawing Videos component contains scaled versions of the patient’s copy and command clock drawings.
- (d) **Summary Table.** The Summary Table component contains a row for every selected scoring system. Each row contains the name of the scoring system, the score and interpretation for the patient’s copy and command clock drawings, and a button for removing the scoring system from the Patient Report window.
- (e) **Scoring System Details.** The Scoring System Details component is a tabbed panel with tabs for every selected scoring system. Each scoring system tab contains two sub-components: a Copy Clock box and a Command Clock box, which both contain their respective Clock Drawing Test score and interpretation. The Clock Boxes also contain two graphs: a patient history graph and a patient comparison graph. The patient history graph is a line chart with each point on the chart representing a score for the selected scoring system from a previous test in chronological order. The patient comparison graph plots the percentage of healthy patients that received each score for the selected scoring system. There is also a line on the patient comparison graph representing the patient’s score.

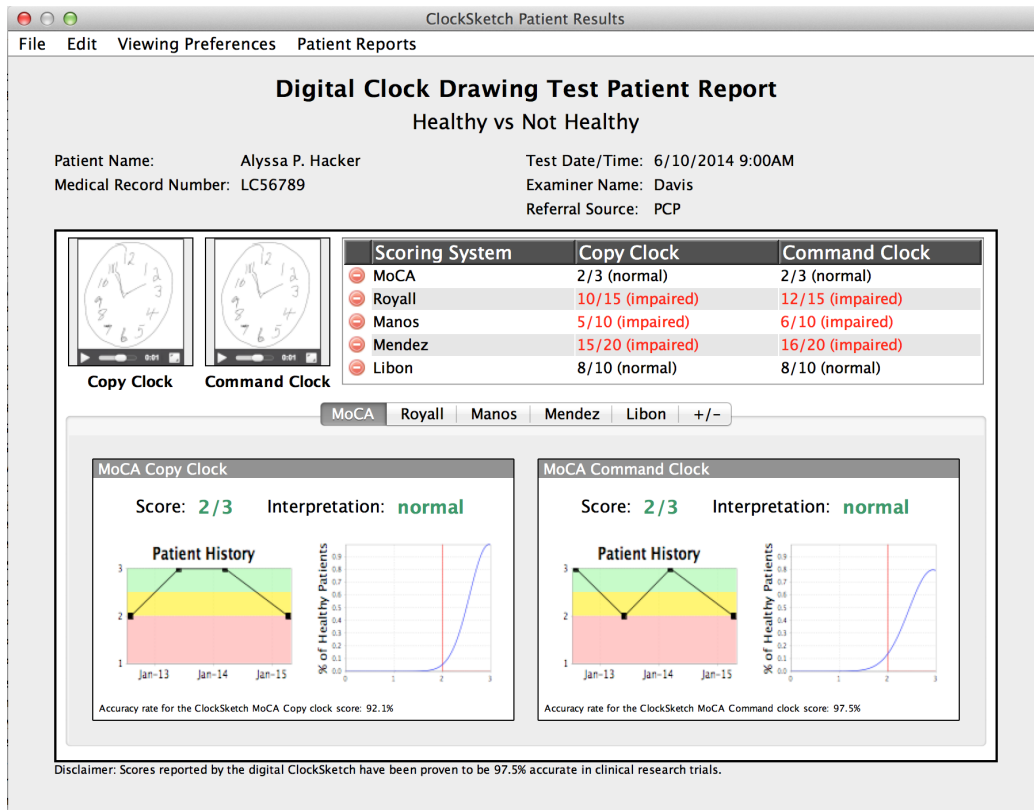


Figure 4-3: Patient Report

4. **Scoring System Selector.** The Scoring System Selector window (Figure 4-4) allows users to dynamically add and remove scoring systems from the Patient Report window. It is accessible through the +/- tab on the Scoring System Details component in the Patient Report window. It consists of a list of checkboxes for each of the available scoring systems. Only the checked scoring systems will be displayed in the Patient Report window.

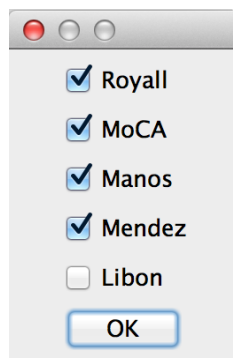


Figure 4-4: Scoring System Selector

4.3 Client-Server Communication

When integrated into an electronic medical records (EMR) system, the ClockSketch UI will eventually allow a doctor to query a centralized server containing patient records. The doctor will query the database with a patient name or other identifier, select the correct patient from the results listing, confirm the patient's information via a preview window, and then load the patient's ClockSketch medical records from the centralized server.

Since we did not have a functional EMR system to integrate with, we built an interface that simulated one by storing patient data information and testing records in files. In the current interface, when a doctor opens a new patient's records they are prompted to select the file showing that record. The file open window is augmented to show a preview of basic patient demographic information so the author can confirm the patient selected matches the one they want. Some of the demographic information—ethnicity, race, gender, and handedness—can be edited directly and is immediately persisted to the patient file.

For our purposes, we use a JSON [24] structure to represent the on-disk structures. There are many JSON serialization libraries available, which makes it easy for an interface on the centralize server to import and export the ClockSketch UI models. In our particular implementation, we use the popular, full-featured Google GSON [10] library which uses reflection in the Java API to serialize our ClockSketch data structures. Thanks to the abstractions we built in the ClockSketch UI, integration to a centralized EMR server should be trivial.

Chapter 5

Evaluation

In this chapter we discuss how we evaluated our interface prototypes during the development and testing phase of the iterative process. We will present our procedure for evaluating each prototype.

5.1 Procedure

We worked closely with three medical professionals and the developers of ClockSketch to evaluate our prototypes. During the development and testing phase for the notional dashboard and paper prototypes, tests were more loosely defined; we presented prototypes to doctors and the developers of ClockSketch and discussed modifications. In order to evaluate our computer prototypes and software implementation, we performed more in-depth user evaluations with doctors following every design iteration.

5.1.1 Early Evaluation Procedure

While iterating over the notional dashboard and paper prototypes, our evaluation process focused on interviews and discussions with doctors and developers to ascertain their needs and goals for the interface. During our interviews with doctors we described the intended purpose of ClockSketch and asked them to evaluate the interface for usability; we asked them if they would be willing to routinely use the interface

to screen patients and what things they liked, disliked, or would like to change about the interface. Our discussions with developers mainly circulated around determining what data and information would be available for the interface to display and how to display them.

5.1.2 Later Evaluation Procedure

Our procedure for evaluating the computer prototypes and software implementations was more thorough than it was for the notional dashboard and paper prototypes. In addition to the interviews and discussions mentioned above, we walked through the interface piece by piece and discussed every part with doctors and developers. We asked doctors and developers about the usability of every feature and what they would like us to add, change, and remove from the interface.

Chapter 6

Future Work and Conclusion

In this chapter we discuss anticipated future work for the interface and conclude with a summary of our current progress.

6.1 Future Work

While the current interface is usable as is, there are a few additional features it may benefit from in the future. First, as mentioned in Section 2.1.2, we would like to add subscores for the individual scoring systems to the user interface. Since doctors use this information and it is already collected, it is simply a matter of adding the appropriate UI elements to display it—perhaps via tooltips on the detailed scoring system chart.

As discussed in Section 4.3, we leave to future work the task of integrating ClockSketch with a hospital’s EMR systems and having it act as a client interacting with records on a remote server.

We also envision adding more detailed panes to the neurologist view showing timing information—this information is captured by the digital pen, but not currently visualized in UI elements. A timeline-style view could be used to present the relative progression of the clock sketch, which a tracking bar that moves across the UI elements as the interface viewer plays one of the clock or command sketch recordings.

6.2 Conclusion

We have presented a doctor-facing ClockSketch UI, a visual user interface for the results of Clock Drawing Test. This interface is intended to be used in conjunction with a human expert interpreting the results. Through vigorous design iteration and evaluation, we incrementally improved and applied feedback to the interface, resulting in the final software implementation presented in Section 4.2.

Bibliography

- [1] Complex medical user interface. Technical report, Sage Research & Design, <http://www.sage-research.com/complex-medical-gui/>.
- [2] Berit Agrell and Ove Dehlin. The clock-drawing test. *Age and Ageing*, 27(3):399–404, 1998.
- [3] K. Aisaka, K. Tsutsui, Y. Murakami, H. Ban, A. Hashizume, Y. Oka, and S. Ishikawa. User interface design and evaluation for electronic medical record system. *Medinfo*, 8 Pt 1:781–784, 1995.
- [4] B Boehm. A spiral model of software development and enhancement. *SIGSOFT Softw. Eng. Notes*, 11(4):14–24, August 1986.
- [5] Soo Borson, James Scanlan, Michael Brush, Peter Vitaliano, and Ahmed Dokmak. The mini-cog: a cognitive vital signs measure for dementia screening in multi-lingual elderly. *International journal of geriatric psychiatry*, 15(11):1021–1027, 2000.
- [6] Rodrigo de Oliveira and Nuria Oliver. Triplebeat: Enhancing exercise performance with persuasion. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '08, pages 255–264, New York, NY, USA, 2008. ACM.
- [7] Hayley E Richardson and John N Glass. A comparison of scoring protocols on the clock drawing test in relation to ease of use, diagnostic group, and correlations with mini-mental state examination. *Journal of the American Geriatrics Society*, 50(1):169–173, 2002.
- [8] Marshal F. Folstein, Susan E. Folstein, and Paul R. McHugh. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1974.
- [9] Kevin Fu. Trustworthy medical device software. *Public Health Effectiveness of the FDA*, 510, 2011.
- [10] Google. google-gson, aug 2015.

- [11] Michael Kirby, Aisling Denihan, Irene Bruce, David Coakley, and Brian A Lawlor. The clock drawing test in primary care: sensitivity in dementia detection and specificity against normal and depressed elderly. *International Journal of Geriatric Psychiatry*, 16(10):935–940, 2001.
- [12] David J Libon, Rodney A Swenson, Edward J Barnoski, and Laura Prouty Sands. Clock drawing as an assessment tool for dementia. *Archives of Clinical Neuropsychology*, 8(5):405–415, 1993.
- [13] Peter J Manos and Rae Wu. The ten point clock test: a quick screen and grading method for cognitive impairment in medical and surgical patients. *The International Journal of Psychiatry in Medicine*, 24(3):229–244, 1994.
- [14] A. Marcus, W. B. Cowan, and W. Smith. Color in user interface design: functionally and aesthetics. *Conference on Human Factors in Computing Systems Proceedings*, page 25, 1989.
- [15] Fourcan Karim Mazumder and Utpal Kanti Das. Usability guidelines for usable user interface. *International Journal of Research in Engineering and Technology*, 3(9):79–82, sep 2014.
- [16] Mario F Mendez, Thomas Ala, and Kara L Underwood. Development of scoring criteria for the clock drawing task in alzheimer’s disease. *Journal of the American Geriatrics Society*, 40(11):1095–1099, 1992.
- [17] Jakob Nielsen. Fresh vs. familiar: How aggressively to redesign. Technical report, Nielsen Norman Group, sep 2009.
- [18] Isabelle Rouleau, David P Salmon, Nelson Butters, Colleen Kennedy, and Kathryn McGuire. Quantitative and qualitative analyses of clock drawings in alzheimer’s and huntington’s disease. *Brain and cognition*, 18(1):70–87, 1992.
- [19] Donald R Royall, Jeffrey A Cordes, and Marsha Polk. Clox: an executive clock drawing task. *Journal of Neurology, Neurosurgery & Psychiatry*, 64(5):588–594, 1998.
- [20] Content Science. Ux strategy + design: Usable interface design boosts a products user adoption. Case study, Content Science, Atlanta, GA, sep 2010.
- [21] Tariq SH1, Tumosa N, Chibnall JT, Perry MH 3rd, and Morley JE. Comparison of the saint louis university mental status examination and the mini-mental state examination for detecting dementia and mild neurocognitive disorder—a pilot study. *Am J Geriatr Psychiatry*, 14(11):900–10, nov 2006.
- [22] Kenneth I Shulman, Dolores Pushkar Gold, Carole A Cohen, and Carla A Zuccherro. Clock-drawing and dementia in the community: A longitudinal study. *International Journal of Geriatric Psychiatry*, 8(6):487–496, 1993.

- [23] William Souillard-Mandar, Randall Davis, Cynthia Rudin, Rhoda Au, David J Libon, Rodney Swenson, Catherine C Price, Melissa Lamar, and Dana L Penney. Learning classification models of cognitive conditions from subtle behaviors in the digital clock drawing test. *Machine Learning. Accepted with Minor Revision*, 2015.
- [24] Ed. T. Bray. The JavaScript Object Notation (JSON) Data Interchange Format. RFC 7159, RFC Editor, March 2014.