2012

Evaluation Of Information Visualization For Decision Making Support In An Emergency Department Information System.

Quaneisha L. Jenkins-Penha
North Carolina Agricultural and Technical State University

EVALUATION OF INFORMATION VISUALIZATION FOR DECISION MAKING SUPPORT IN AN EMERGENCY DEPARTMENT INFORMATION SYSTEM

by

Quaneisha L. Jenkins-Penha

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Department: Industrial and System Engineering
Major: Industrial and System Engineering
Major Professor: Dr. Steven Jiang

North Carolina A&T State University
Greensboro, North Carolina
2012

ABSTRACT

Jenkins-Penha, Quaneisha L. EVALUATION OF INFORMATION VISUALIZATION FOR DECISION MAKING SUPPORT IN AN EMERGENCY DEPARTMENT INFORMATION SYSTEM. (**Major Advisor: Dr. Steven Jiang**), North Carolina Agricultural and Technical State University.

Emergency departments collect large amounts of data to make decisions regarding patient care. New technologies are used to support the decision-making process. Various information visualization (IV) techniques can be used to support healthcare professionals in visualizing patterns that can be helpful in the decision-making process. Implementation of decision support tools with IV techniques in EDIS is difficult to achieve and thus an evaluation of the techniques is needed. The purpose of this dissertation is to propose an evaluation framework to assess various IV techniques in EDIS and provide recommendations for developers.

A comprehensive assessment framework was developed based on performance measures, user opinion, mental workload, and eye tracking metrics to evaluate IV techniques for EDIS. A heuristic evaluation, an empirical study, focus groups, and a usability test with domain experts were conducted to demonstrate the potential of utilizing this methodology. A significant difference in performance, usability, mental workload, and eye tracking metrics was found for the visualization techniques as applied to EDIS. The findings of these studies suggest that when applied to an EDIS the density chart, tree map, and network diagrams have lower performance times, better accuracy, higher usability opinion, and lower mental workload than the 3D scatter plot, scatter plot

matrix, and parallel coordinates. From these results, a set of guidelines is recommended for designers of EDIS that employ the use of visualization techniques. Future work includes further use of this assessment framework to develop a model for IV effectiveness and its application to other complex systems.

School of Graduate Studies
North Carolina Agricultural and Technical State University

This is to certify that the Doctoral Dissertation of

Quaneisha L. Jenkins-Penha

has met the dissertation requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2012

Approved by:

Dr. Steven Jiang Major Professor	Dr. Daniel Mountjoy Committee Member
Dr. Xuili Qu Committee Member	Dr. Forrest Toms Committee Member
Dr. Paul Stanfield Department Chairperson	Dr. Sanjiv Sarin Associate Vice-Chancellor for Research and Graduate Dean

DEDICATION

To those who came before me, I understand and appreciate your foresight and journeys.

To those that will come after me, remember what you want.

To those who have come with me, this dissertation is as much mine as it is yours.

BIOGRAPHICAL SKETCH

Quaneisha Lynnette Jenkins-Penha was born on June 8, 1984 in Freehold, NJ. She graduated with honors from Lakewood High School, Lakewood NJ in June 2002. She received her Bachelor of Science from Stanford University in June 2007 in Science, Technology, and Society with a focus in Management Science & Engineering and double minors in African & Afro-American Studies as well as Classics. In May 2009, she received her Master of Science degree in Industrial & Systems Engineering (ISE) with a focus in Human-Machine Systems from North Carolina A & T (NC A&T) State University. In fall 2009, she began the doctorate program in ISE keeping the same focus at NC A&T. She is a member of Alpha Pi Mi Industrial Engineering Honor Society, Golden Key International Honour Society, Phi Kappa Phi Honor Society, Human Factors and Ergonomics Society, Institute for Industrial Engineers (IIE), and the National Society of Black Engineers (NSBE). She has received technical writing awards from the Symposium on Human Interaction with Complex Systems and the NSBE National Convention. She has also presented her research at the IIE Research Conference and the Applied Ergonomics Conference. She worked as a coordinator for Stanford Womens' Community Center (2003-2007); research assistant at Kent State University (2004); engineering intern for Parsons in (2005-2006); research associate at SA Technologies (2008), General Electric Transportation (2009), and interned for IBM Software Group (2010). Upon completion of the doctorate program, she will take a usability engineering position with Boeing in the field of cyber security information systems.

ACKNOWLEDGMENTS

First, I would like to thank my committee members, Dr. Daniel Mountjoy, Dr. Xuili Qu for their assistance during this research effort providing me their expertise and constructive feedback while shaping this research. I am grateful to my committee member, Dr. Forrest Toms, for his research guidance, advice, and valuable discussions that helped not only sort out the details of my research but of my life and personal outlook. I would like to thank my advisor, Dr. Steven Jiang for his guidance, patience, encouragement and giving me the opportunity to build a unique idea from the ground. I know I have been very fortunate to have an advisor that has given me the freedom to explore ideas on my own while still guiding me through my faltered steps. In the last stages, his support was instrumental to me overcoming many challenging situations and finishing this dissertation.

The faculty of the Industrial and Systems Engineering, Leadership Studies, and Computer Science, have provided me with a graduate education that has helped me refine my research and writing skills and approach my work as a research professional.

I need to express my gratitude to the nursing students from Guilford Technical Community College, nursing staff from Guilford Health Care Center, the emergency department staffs of Wesley Long Hospital and Duke Hospital that helped make this work possible with their feedback and participation.

I am indebted to my colleagues at NCA&T who worked with me on this project Khaliah Hughes, Katherine Bagley and Antonio Lee for their comments, participation, and supportive ideas; I am honored to be a member of the Inaugural Class of Brass.

This work was definitely not doable without the technical assistance of Ritson Delpish, Khary Popplewell, Leonard Washington, and Dr. Kenneth Williams.

Finally, I would like to acknowledge with much gratitude, my family for their unwavering support and friends who answered the phone when it was the darkest hour.

Without the aforementioned individuals, this research would have been most difficult and without meaning for me. I also acknowledge the institutional support that I have received while completing my dissertation. This work was funded through the Title III program of North Carolina A&T State University.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE **PAGE**

LIST OF TABLES

KEY TO ABBREVIATIONS

SME	Subject Matter Expert
IV	Information Visualization
EDIS	Emergency Department Information System
HIS	Healthcare Information System
ED	Emergency Department

CHAPTER 1

INTRODUCTION

Since the mid 1980s, the emergency department in hospitals has virtually replaced regular physician visits for many persons in the United States as the cost of medical care has increased. In 2010, there were 123.8 million visits to the emergency departments throughout the country with 13% resulting in hospital admission (CDC, 2010). As a result, professionals in the emergency department need to make their decisions quicker and better than before. One of the most important tools to support decision-making is through the technology of information visualization. This dissertation explores utilizing information visualization (IV) techniques to display patient derived data from a national hospital emergency department survey conducted by the Center for Disease Control and Prevention (CDC) and to assist decision-making. A framework to evaluate IV in an emergency department information system (EDIS) is presented. An understanding of the contributing variables that affect decision-making in healthcare such as time, accuracy, and mental workload can improve the performance of healthcare personnel.

1.1 Background

Electronic medical equipment is common place in today's healthcare system. A healthcare system is a complex, dynamic, unpredictable, integrated set of dependent components that must work together in order to achieve success in patient care. Adding to the complexity are diverse teams that include but not limited to attending physicians, residents, pharmacists, nurses, and technicians. Particularly, intensive care units and emergency departments are the most challenging subspaces within a healthcare system, bringing them to the attention of many researchers in the decision sciences.

There is an enormous amount of data collected for each patient from their arrival time to the end of their follow up visits to the emergency department (ED). The patient record is a composite of all the data acquired and created during a patient's path through the healthcare system (Tang & McDonald, 2001). The focus of patient care is not just on the period of illness, but also includes the cycle of wellness to illness to recovery, and back to wellness again. Therefore, a patient's record must integrate data from multiple sources across multiple providers and for various times. Two significant categories of decisions in healthcare are diagnosis and the treatment plan. The decision makers, usually physicians, must use the raw data acquired and the views of other staff members involved to make decisions.

There are many persons involved in healthcare decision-making. For the purposes of this research, 'healthcare' system or setting will be limited to just hospitals. In hospitals, there are two main types of medical physicians; general practitioners or those trained in a specialized area. Within these groups of staff members, there may be different levels of seniority (interns, residents, attending, etc.). These are the persons most likely with whom key decisions will end with in regards to diagnosis and treatment. Nurses are another type of staff that support physicians and usually have the most patient contact. They may also have a specialty and have different levels of seniority (aides, licensed, registered, and professional). Technicians are members of staff with a limited range of duties and expertise. They usually are responsible for obtaining patient samples to be delivered to a lab for analysis. Pharmacists are staff members that supply medications for patient treatment. A physician making a diagnosis decision may look at the vitals collected by a nurse, the lab analysis from a technician, and prescribe a treatment plan that the pharmacist will supply to the patient.

An emergency department (ED) for this research will be defined as a facility that specializes in acute care for patients without prior appointments, typically attached to a hospital or other primary care unit. An ED is a complex system with a high degree of complexity, fatality risks, uncertainty, and dynamic environment. In addition, there is a tremendous amount of patient data collected in various formats. An information system designed for emergency department use needs to be able to display collected patient information and support decision-making for healthcare practitioners. The system is a complex web of information in which analysis can prove to be difficult, affecting patient care outcomes. Complex data analysis occurs in various areas of patient care including the following tasks (Mirel, 2003):

- Emergency responders prioritizing logistics
- Nurses dispensing medicine and giving other treatments
- Communication among various health care professionals
- Corporate and government project managers allocating resources

There are differences between the tasks associated with complex systems versus well-structured tasks (Mirel, 2003). For tasks associated with complex systems, healthcare professionals must search and sift through much more information and determine how to allocate their attention among the abundance of information. Due to this, data analysis and decision-making are already burdensome for the user, when an unusable display is added to the situation there are even less cognitive resources of the user available. In a complex healthcare environment, such as an emergency department, time is a key component. Good decisions made too late turn into bad decisions leading up to tragic consequences. Often there may not be a clear indication of positive or negative results when a person makes a decision.

In healthcare decision-making is multi-dimensional, dense, and often can result in death. The factors affecting the decision-making process are numerous and multifaceted. The process itself has been measured by objective (user performance) and subjective (user preference) variables. These measured variables can be influenced by the technological support a user receives throughout their decision-making process, such as a decision support tool that employs information visualization.

Information visualization is not a novel technology and has been researched extensively over the decades. This technology can simultaneously provide a "big picture" and a "small picture" for clarification in an on-demand format for users. There are many types of information display designs, but information visualization can improve decision-making by influencing user cognitive resources. This relationship provides an avenue of human factors research into information visualization. Visualizations can be a significant display method for complex information systems, such as an emergency department information system (EDIS). There are various information visualization techniques in existence, and not all are appropriate for any decision-making situation in an emergency department. Therefore, there exists the need to analyze visualization techniques for their effectiveness and support in healthcare decision-making.

1.2 Problem Statement

Although assessment methods have been developed to evaluate information visualization techniques, many of them are not comprehensive and are just one dimensional. The existing methods although successfully applied in some domains, including medical, are not necessarily an exact match for the domain of emergency department patient information. Due to this mismatch, there are no set of guidelines for evaluating visualization techniques for EDIS.

As stated earlier, there is a significant amount of data, types of data, and data sources collected on patients during emergency department visits. Usually, there is also the accompaniment of time constraints coupled with the overwhelming data that causes problems for healthcare professionals to perform their jobs. Physicians and other healthcare practitioners need to be able to filter pertinent information when necessary for diagnostic and treatment processes.

Much of the research on analyzing visualizations focuses on the development of the techniques and somewhat less focus on the evaluation of the techniques; in either case, none focuses on EDIS evaluation (Au, Carey, Sewraz, Guo, & Rüger, 2000). As a result, empirical studies are needed that collect objective and subjective data to assess the effectiveness of the techniques (Mirel, 1998). The need for the assessment is necessary in optimizing the user experience, improving efficiency, and productivity.

Decisions in an ED that could be supported by the employment of visualization techniques include searching for a specific type of patients for clinical studies, analyzing drug treatments (what are the most commonly prescribed drugs), or analyzing diagnoses (most commonly paired diagnoses). Yet, difficult decisions are made daily in complex hospital systems without the support of visualization techniques (Falkman, 2000). A particular analysis of visualization techniques that are properly introduced into an EDIS would potentially aid in the decision-making processes.

1.3 Motivation

Presently, there are no existing information visualization guidelines for comprehensive emergency medical datasets. This particular set of guidelines would be beneficial because it answers the questions of what technique to use and when to use a particular technique. Several assessment techniques exist for information visualizations. Some studies assess multiple techniques for the business domain, while other studies assess a single visualization for the medical domain (Ziegler, 2002; Rester, 2007, Pohl, Rester & Wiltner, 2007). Eye tracking has been used to evaluate visualizations as well (Pirolli, Card, & Van Der Wege, 2001). However, no assessment of information visualization has been done with an emergency department information system. Therefore, there is a need to develop a comprehensive assessment technique for emergency medical information visualizations to construct an informative set of guidelines.

1.4 Research Objectives

The research question guiding this dissertation is "how do information visualization techniques for large medical datasets influence decision-making performance?" The overall objective is to develop a multi-layered testing approach to evaluate various information visualization techniques for an emergency medical patient dataset. The sub-objectives are as follows:

- Develop a test bed information display composed of 6 visualization techniques
- Conduct a heuristic evaluation with usability subject matter experts for the visualization test bed to determine usability issues
- Conduct an usability study and focus groups with novices (controlled lab experiments to obtain quantitative and qualitative user data)
- Conduct a case study with medical professionals (feedback from field users/ domain experts)
- Perform an evaluation of information visualization techniques and create a set of guidelines for developing an emergency department information system

1.5 Organization of Dissertation

The next chapter presents a literature review in the areas of hospital information systems, decision-making models in healthcare, and information visualization. The planning phase of this research, described in Chapter 3, presents the assessment framework, reports the outcomes of user requirements analysis from Wesley Long Hospital emergency department, and describes the test bed development for the graphical user interface with domain and software subject matter experts (SMEs). Chapter 4 outlines the methodology for the heuristic evaluation with usability SMEs, the pilot study, the empirical study, focus groups, and a case study with domain SMEs. The results of the usability testing are provided in Chapter 5. The discussion outlining the implications of the results of the previous chapters, limitations, and guidelines are provided in Chapter 6. Chapter 7 concludes this research and provides recommendations for future work in further developing this assessment technique, and notes the need for further work on the utilization of information visualization for patient medical data.

CHAPTER 2

LITERATURE REVIEW

The focus of this chapter is to provide an overview of literature on healthcare and emergency department information systems, decision-making in healthcare, and information visualization. The first section on healthcare information systems provides a general history of hospital information systems and emergency department information systems. The next section on decision-making in healthcare focuses on the decision-making process in healthcare and point out the underlying cognitive issues. The final section provides a comprehensive literature review on information visualization.

2.1 Healthcare Information Systems

The purpose of this section is to introduce how medical information systems are designed, what types of data they store and display, and how these systems are used as decision support tools in healthcare. Following, is a review of healthcare information systems, electronic patient records, and emergency department information systems. In addition, decision support systems (DSS) are included to highlight how these systems can benefit from DSS tools.

Since the 1950s, there have been continuous strides in the area of healthcare information systems (HIS). These advances are attributable to the improved technology capabilities of computer systems as well as new concepts in work organization. Early HIS were used for communication purposes for example, collecting and routing orders, documenting chargeable services, and accessing lab reports (Tang & McDonald, 2001). Since the 1970s, computerized systems have become typical and standard in hospitals (Ball, 2003). A hospital information system is described as a subsystem of a hospital, which is composed of all information processing, as well as, the associated human or technical actors in their respective information processing roles (Haux, 2002). It has been shown that the major problem areas are user involvement, education, and training. Due to improvements in information technology (IT) to enhance the user experience, there has been an increase in the number of features in HIS. However, "functionality creep", has offset the gain where an increase in function leads to complicating the user experience. IT has been used to improve the administrative tasks and treatment processes by medical professionals in hospitals. Hospital managers value information systems as tools to improve information flow and provide quality services to their patients (Jensen & Aanestad, 2006). There is a desire, as well as need for safe and high quality care, which drives the adoption of IT systems in healthcare. In 1999, the American Medical Informatics Association emphasized that human factors has a valuable role in healthcare informatics for the reduction in medical errors and improvement for patient safety (Ball, 2003).

The federal government introduced new requirements for transitioning to electronic health records in 2002 as part of the Health Information Portability and Accountability Act (HIPAA). Patient related data that was once manually input with pen and paper is now computerized and put into electronic databases. As in other industries, there has been an "information explosion" over the decades in healthcare. Information is abundant and available in many formats, provided from many sources.

Decision-making in hospitals can be supported with various technology by providing accurate and timely accessibility to electronic patient records (Favela, Rodríguez, Preciado, & González, 2004). The many medical professionals involved with the decision-making have various uses and needs for information systems.

Hospitals are technology and data rich environments, which create large collections of data. The workers and the resources (equipment, tools, etc.) are constantly mobile and changing. In addition, the patients continuously arrive and depart at different rates introducing uncertainty into the system. All these above mentioned factors (large amount of data, dynamic characteristics, and uncertainty) are aspects of complex systems. Hospitals are complex information environments and thus have a need for considerable technical and computational infrastructure. With regard to human factors research, many issues exist with modern computer based healthcare systems with information displays. Common issues with paper based records are inaccessibility, static, missing, illegible, and unstructured. Common issues with computer based records are usability, ease of use, ease of learning, and the incorporation into mental and physical work routines.

2.1.1 Electronic Patient Record

In 1910, the Flexner report, encouraged physicians to keep a patient-oriented medical record (Kunitz, 1974). By 1991, Reiser states the “purpose of a patient record is to recall observations, to inform others, to instruct students, to gain knowledge, to monitor performance, and to justify interventions” (Reiser, 1991). The data collected from a single record can be used to develop care plan for a single patient. In other cases, when the records are aggregated they can be used to guide the care plan for an entire population (Tang & McDonald, 2001).

An electronic patient record (EPR) is a “technology that is used to provide examination, treatment, and care of a patient” (Jensen & Aanestad, 2006). The EPR is the electronic version of the original paper based version. Transforming the paper version to an electronic version allows for eased sharing of the record amongst staff. An EPR is typically composed of basic patient data such as vitals, current medications, arrival time, and many others, but also includes complex patient data such as nurses’ notes, progress reports, x-rays, lab reports, and so on. The EPR is usually not the only component of a hospital information system. The record is typically used in conjunction with other information systems such as patient billing and facility resources (Jensen & Aanestad, 2006). These electronic records allow for the integration of patient and clinician data with increased accessibility for all those involved with patient care to improve quality of care and drive down associated costs (Favela, Rodríguez, Preciado, & González, 2004).

Possible uses for computerized patient record systems include queries and surveillance. Staff may use such a system to generate alerts about significant clinical events, retrieve selected characteristics of a patient, or to summarize information statistically. A query can be the retrieval and aggregation of data for groups of similar patients, for example, to identify patients of certain subgroups for clinical research (ex: all males, over 50, and taking anti-hypertensive medication) or patients that received particular recalled drugs. Surveillance can exist across subgroups of patients to detect and flag patient conditions that justify medical attention. These computerized patient record systems are potential support tools for clinical care, clinical research, retrospective studies, and/or administration. Retrospective studies are of particular interest because they can obtain answers to research questions at a fraction of time and cost from already existing data sets.

There are many types of technologies available for use in a healthcare setting, such as databases, graphical representations, charts, graphs, textual, and other data capturing devices. Technology support is important because it can speed up care, reduce cost, improve the quality of care, and provide a double check for decision makers. A computer based patient record is designed to address limitations and provide benefits that are not associated with static view of events (Tang & McDonald, 2001). The workflow and interactions between healthcare professionals and their patients may change with the result of the introduction of new technologies to the workplace. There are often associated human and organizational factors that dominate the technical challenges themselves (Tang & McDonald, 2001). Usually there are several types of patient data each with its own data content terminologies, format, software, etc. which all must be compiled and displayed in one interface. Therefore, the interface needs to be designed with the aim of being intuitive as well as efficient for the user.

2.1.2 Functions of an Emergency Department Information System (EDIS)

Hospital emergency departments are available all day and all night for persons requiring medical attention. The demand varies by day, time, location, population, hospital specialization, etc. Emergency departments have been incorporating new technology to assist them in giving quality patient care (The American College of

Emergency Physicians, 2003). An Emergency Department Information System (EDIS) is an "electronic health record system designed specifically to manage data and workflow in support of emergency department patient care and operations" (Rothenhaus, Kamens, Keaton, Nathanson, Nielson, McClay, Taylor, & Villarin, 2009). These systems typically provide standard functions for emergency healthcare. The Emergency Care Special Interest Group (EC-SIG) of Health Level Seven (HL7) created a functional profile for typical EDIS to standardize the required functions (Rothenhaus, Kamens, McClay, & Coonan, 2007). This profile outlines and describes the required functions of an EDIS as mapped to the following emergency department milestones in healthcare: patient arrival, triage, nursing and physician assessment, orders, results, procedures, ongoing assessments, and admit/transfer/discharge planning. Following are the brief descriptions of those required functions from the EDIS Primer for Emergency Physicians, Nurses, and IT Professionals: tracking, registration, clinical workflow, orders, documentation, and post-disposition management (Rothenhaus, et al., 2009).

Tracking function in an EDIS can refer to the patient's physical location or status of care through the emergency department. Physical location tracking includes manual input updates from users or automatic input with radio frequency identification for the entire visit of each patient. Status of care tracking includes capturing certain milestones, for example when a patient is moved from the waiting room to an assessment area. Included also are department centered-tracking which covers information on emergency department bed use, availability, number of patients waiting to be seen in the waiting-room, patients awaiting bed assignment, etc.

Patient arrival and registration involve clinical and administrative data. Usually an EDIS is used in conjunction with other systems for administrative data and generates a record when the patient enters the ED. An EDIS should provide the triage nurse with a way to search and retrieve previous patient visits to the ED to know their prior problem list, prior listed allergies, and current medications.

There is a significant need in the ED to communicate the status of workflow tasks because an essential functional of an EDIS is to manage and display the tasks that are to be completed. Typical tasks in an EDIS to help with task management include maintaining a master set of tasks, ability to invoke the master tasks, display the progress of tasks, and results report. To track and display tasks optimally, an EDIS must display the status of tasks for patient, and display tasks for specific ED personnel.

In any healthcare setting whether it is an emergency department or clinic, orders for tests, medications, nursing tasks, and materials must be completed in the most efficient manner possible. It is challenging to do so because a single order may encompass a single task or several tasks. An EDIS can support this function by being customizable by role and physical location. Roles may vary with the time of day, day of week, institution, or may be shared between team members.

Clinical documentation is another complex part of emergency healthcare because the data collection is often discontinuous and delayed. An EDIS must provide the functionality to support data collection by all providers, for staggered/delayed/offsite completion of records, and notification of incomplete records. There is a lot of redundancy in clinical documentation and in order for EDIS to be assistive; it needs to allow for pulling of prior records and summarization of those records.

After a patient leaves the emergency department, the treatment is not complete. Emergency department healthcare requires the means to track patient issues after their visit and this can include labs, diagnostics, prescriptions, discharge education, and instructions, follow-up information, and detailed visit information. An EDIS must have the functions required to support these tasks as well as include a way the medical professionals can analyze any conflicts between the final interpretations and the initial diagnostic test results for each patient.

2.1.3 Features to Improve an EDIS

In addition to the aforementioned standardized functionality of an emergency department, information system (EDIS), there is still room for improvements with the system. Standardization of components, an intuitive graphical user interface design, and allowing short-cuts for expert users are three features that could improve an EDIS.

An EDIS may be a composition of several systems, applications, modules, or components developed in different architectures with multiple vendors and/or development teams (Rothenhaus et. al., 2007). This is a problem in many industries that rely on software support. The simplest example of standardization would be an example from Microsoft Office where the "disk" icon means "save" feature in its word processor, database, spreadsheet, and presentation programs. In an EDIS that is built on different system components, standardization would improve user efficiency because this decreases the load on long-term memory by strengthening simple schemas and user mental models (Wickens, Lee, Liu, & Becker, 2003). Consistency across components will assist with learning the technology, adopting the use of new technology, and can increase productivity. A concrete example would be to have the patient search function utilize the same icon or method for each module, no matter what function mode the EDIS is operating in or what role the team member may be.

Research has shown that when a new EDIS is implemented, typically there is a decrease in productivity shortly thereafter (Rothenhaus et. al., 2009). Emergency departments are fast-paced environments where the slightest error can cause the most significant consequence. It is understandable in such an environment that the margin for error is small and that any intervention in the routine methods of the personnel can cause some productivity loss. However, an EDIS can partially overcome this short-term productivity loss by having a well optimized graphical user interface that gives the users what they need and when they need it. This interface should build upon prior existing usability software heuristics that focus on the "ease of learn" and "ease of remembering" (Nielsen, 1994). If a system is easy to learn and easy to remember how to operate (particularly for the infrequent users), it will not have as heavy a negative impact on productivity in the short term for a complex environment such as an ED. An example of this would be to have certain repetitive functions, for example, the act of saving a patient record should be the same no matter the function mode of the EDIS.

Physicians in emergency departments most likely have a range of experience and expertise. Each one has developed a way of working swiftly, accurately, and fully to give quality care to the patients. An EDIS must have the functionality sufficient for documentation and these same skilled physicians will have to use that functionality to support their tasks in documentation (Rothenhaus et al., 2009). Providing short-cuts to assist those more frequent users of an EDIS may help increase productivity and user satisfaction. Short-cuts do not have to be used by everyone, but there are some persons who will benefit from them and those persons can significantly decrease their workload.

2.1.4 Clinical Decision Support System

A decision support system (DSS), describes a computerized system that aids in human decision-making (Hudson, Cohen, & Anderson, 1991). A clinical DSS is any computer program designed to help healthcare professionals to make clinical decisions (Musen, Shahar, & Shortliffe, 2006). There are three different functions of decision support for clinical DSS tools: information management, focusing attention, and patient-specific recommendations (Musen et al., 2006). For all functions, the system needs to have the technology and features inherent to support the underlying cognitive functions of the user. For the purposes of this research, the first two functions are significant. Information management is composed of dedicated workstations that offer an enhanced work environment for the storage, retrieval, and browsing of data. However, the interpretation is left to the clinician. The system itself does not provide “answers” but provides a way to find the answer. Focusing the attention of the user on specific abnormal values is another function of clinical DSS tools. Of course, the system needs to have a clear definition of what is “abnormal” for particular values. Some existing software is geared towards focusing the user's attention with flags or messages about abnormality in data, for example drug interactions.

Underlying cognitive processes such as attention have a significant impact on human information processing and thus decision-making (Patel, Zhang, Yoskowitz, Green, & Sayan, 2008). Determining what information is needed and how it will be processed is important for developers of decision support systems in healthcare. This is significant to situations where not every staff member will need the same information or place as much significance on a particular piece of information.

This section has presented a description and analysis of healthcare information systems, the electronic patient record, the emergency department information system, and clinical decision support systems. Several usability aspects have been highlighted, in particular, information presentation, which affects underlying cognitive processes of users and their decision-making. The following section will offer a discussion and analysis of cognition and decision-making in healthcare.

2.2 Decision Making in Healthcare

The purpose of this section is to analyze the decision-making process in healthcare and point out the underlying cognitive issues. The following review of literature will cover mental workload, human information processing and attention, cognition, and decision-making models. This review is to show that there exists a need to improve decision accuracy, reduce cognitive workload, and reduce decision time in healthcare decision-making.

There are various types of decisions made in healthcare. Some decisions concern specific patients, a group of patients, or are about administration (Musen et al., 2006). Examples include a researcher may want to view lab data to assist designing the next phase of clinical studies or a hospital administrator may use management data to guide decisions about resource allocation. In either type of healthcare related decision, there exists a major challenge when the decision maker is overloaded by a large amount of information. Poor data potentially leads to poor decision-making but poorly presented quality data can still lead to poor decision-making (Musen et al., 2006). If the decision maker's cognitive abilities are less than adequate to deal with the decision, poor decisions can still be made. There is a need to support cognitive abilities of users and reduce mental workload.

2.2.1 Mental Workload

Mental workload measurement is one of many types of measures used to assess human machine systems in various work environments (Wright, Taekman, & Endsley, 2004). It can be measured directly or indirectly and usually coupled with measures of performance. Mental workload can be affected by time, task, information presentation, information processing, the environment, and the available resources. Workload analysis is beneficial because it is an important part of usability analysis and can be used to make inferences about an operator's ability to perform (Wickens et al., 2003). If the operator's ability is enhanced, by relation their decision-making performance should also improve.

Mental workload is a multidimensional subjective measurement, that when applied to healthcare, reflects the medical professional's perception in terms of their own effort exerted (Bertram, Opila, Brown, Gallagher, Schifeling, Snow, & Hershey, 1992). Mental workload as it relates to information processing is a by-product. At any moment, the human information processing system is limited. When performance is reduced and task demand increases, the mental workload increases as a way to offset the performance reduction (Bertram et al., 1992).

Information overload has been recognized as an issue in many information system domains. Information overload is the result of “the inability of living systems to process excessive amounts of information” (Grisé & Gallupe, 1999). When there are many alternatives or large amount of attributes compete for the user’s attentional resources, the decision maker is said to be “mentally overloaded” (Svensson, Angelborg-Thanderz, Sjoberg, & Olsson, 1997). The interaction of high information load, high task complexity, and limitations of human information processing is the contributing factor to information overload. Typically, mental workload increases to offset these limitations (Grisé & Gallupe, 1999). There is no simple, direct correlation between information amount and decision-making performance. Initially, there is a positive correlation between the two variables. However, this trend eventually reverses itself and more information results in less accurate decisions as seen in Figure 2.1 (Eppler & Mengis, 2003; Camoes, 2008). After the peak, the decrease in accuracy is a reflection of failed integration of additional information into the decision-making process (O'Reilly & Caldwell, 1980). After the optimum amount of information is passed, some users may feel as though they can make better decisions with more information, this has been described as an “illusion of knowledge” (Van Raaij, Peek,Vermaat-Miedema, Schonk, & Hautvast, 1988). When the flow or rate of information presented is low, there is decreased cognitive resource needed (Senders, 2009). When the flow increases, resources focus on smaller portions of information presented, and an increase in cognitive resources are required.

As a way to offset the limitations imposed in human information processing, people tend to develop decision strategies to deal with information overload. Some persons may use heuristics-and-biases or fast-and-frugal methods to determine a selection (Reyna, 2004). These can be seen as short-cut methods. It is noted that these methods may have negative consequences. However, this is an area where a well-designed information system for decision support can improve decision-making quality.

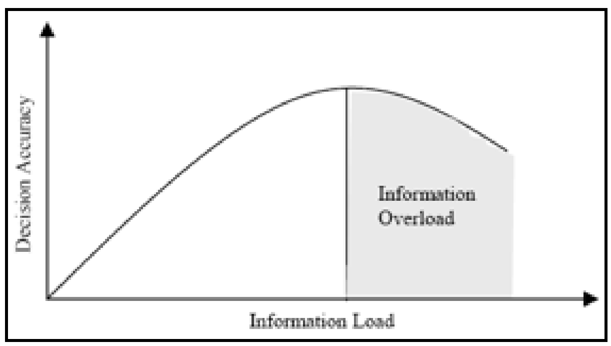


Figure 2.1: Information Load and Accuracy (Eppler & Mengis, 2003)

2.2.2 Human Information Processing Model

To design information systems that support decision-making in healthcare, the strengths and limits of human information processing should be analyzed. One of the most common information processing models proposed by Wickens can be seen in Figure 2.2 (Wickens et al., 2003). The model incorporates perception, memory, attention, and response. Of particular note for this research is the attention resource allocation at the top of the Figure. These resources affect perception, working or short-term memory, decision selection, and response. Examining Figure 2.2, it can be seen that an enhancement in perception, working memory, or attention resources, can improve the decision selection. From a human factors perspective, this is significant in designing decision support tools that will provide assistance with making efficient and accurate decisions.

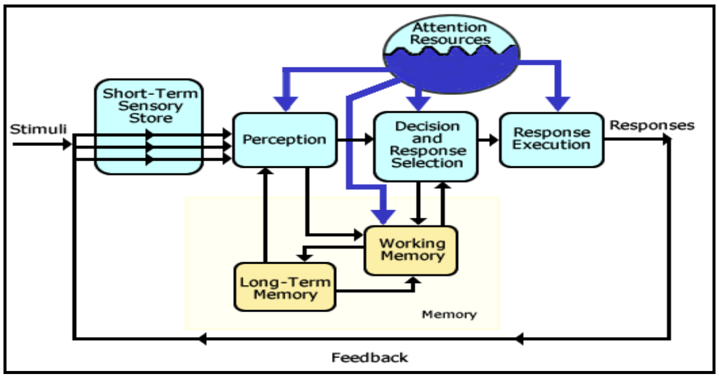


Figure 2.2: Human Information Processing Model (Wickens et al., 2003)

For this research, special focus is on attention. Two stages of information processing, “pre-attentive” and “attentive” are related to attention depicted in Figure 2.3 (Senders, 2009). The pre-attentive stage is where the sensory systems (visual, audio, etc.) detect information. The attentive stage is where the subject has focused on a subset of information and ignored the rest. During the pre-attentive stage, meaning is not associated with raw input from the sensory systems and everything is a set of basic characteristics (color, size, location, etc.). This stage is automatic and occurs without much focus. The information collected during this time is forgotten if it does not get past the information filter seen in the middle of the Figure 2.3. The attentive stage occurs after information has made it through the filter.

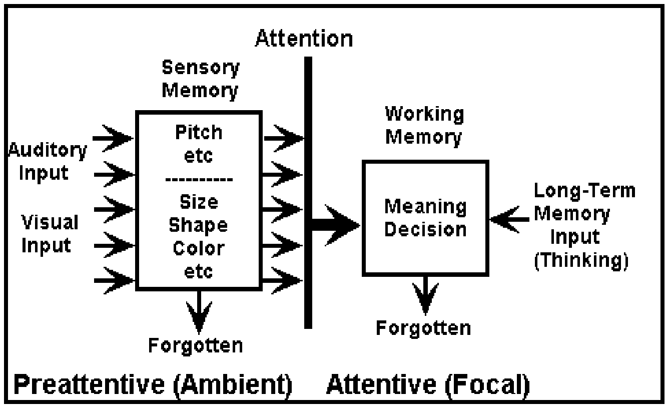


Figure 2.3: Stages of Information Processing (Senders, 2009)

With regard to decision-making, if care is taken to design a decision support system that enhances significant information for a healthcare professional's visual field, then the filter that occurs between pre-attentive and attentive is improved. However, it is still important to note that the filter is dependent upon the individual's knowledge base and environment. If the knowledge base is corrupt or if the sensory input is degraded from the environment, then the decision-making process is likely to be poor.

2.2.3 Cognition

Cognition occurs when humans have gathered data from our sensory system and have attached meaning to the data. In order to make an accurate decision, humans first need to understand the problem. There exist two methods to reduce cognitive load: simplify the task or increase the resources available. In most cases, it is unacceptable and impossible to simplify a complex task, particularly in healthcare. The latter method is an area where human factors can be used to highlight solutions. One way to increase the resources is to supplement cognition with technology. Many problematic issues have been identified in the area of cognition. However, this review will focus on two problematic areas: cognitive tunneling and cognitive friction.

Cognitive tunneling, also called “confirmation bias” results from top-down or expectancy-driven processing (Wickens et al., 2003). Cognitive tunneling usually occurs when the system is very complex, and intermittent failures of a system are difficult to determine (Wickens et al., 2003). Cognitive tunneling happens when a person fixates on a particular hypothesis and only gives weight to the cues that confirm that hypothesis. A decision maker would even interpret ambiguous evidence as supportive even though it was not supportive. Most likely, experts are the type of persons to have productive hypotheses regardless of cognitive tunneling because this issue is based on their long-term memory and experiences (Kendler, 2004). While this is not necessarily true all the time, ignoring or misunderstanding cues is improper and dangerous with potentially disastrous consequences in healthcare decision-making. Decision support systems should be implemented with care so cognitive tunneling is not encouraged. The use of color may encourage users to fixate on a particular element and miss other significant information causing what is known as “inattentional blindness” (Bui, Aberle, & Kangarloo, 2007).

Cognitive friction is “the resistance encountered by a human intellect when it engages with a complex system of rules that change as the problem changes” (Cooper, 2004). An example of cognitive friction is a microwave with a 10-key pad where the keys are used to control two different states on the device. The keys are used to control the radiation level and the duration of cooking. Cooking something longer or at a higher radiation level are two common complaints among microwave consumers. Software design potentially is very high in cognitive friction. Cognitive friction only occurs when the interaction does not fit the user's mental model (Cooper, 2004).

2.2.4 Decision Making Models

There are many types of decision-making models that a physician may utilize for one patient because each situation is dynamic and unique from other situations. There are trial and error methods, educated guess methods, and then there are more complicated methods such as naturalistic or humble decision-making. It is difficult to categorize models since a physician may actually use more than one model simultaneously with a single patient's care. Even so, decision support tools, with special attention on information display can help support the cognitive efforts of medical personnel in emergency healthcare decision-making.

Manufacturing and healthcare are two different domains. However there have been analyses that take a decision-making model from manufacturing and apply it to healthcare. One such model is the "fast" model (Stepanovich & Uhrig, 1999). This model uses the available information, develops alternatives, seeks advice, resolves conflicts, and eventually leads to an integrated decision. This model has its origins in high volume manufacturing, which in some ways is similar to healthcare due to the complexity inherent in such a system. An advantage of applying this particular model is it adheres to the time constraints in a fast-pace, critical system. This model is geared more towards the management and administration of a hospital, not so much dealing with the staff that interacts with patients. A disadvantage of this model is that it does not allow for reflective decision-making or incorporation of feedback.

The concept of "sensemaking" has been previously applied to complex environments. Sensemaking is a method for persons to make sense and process information from their environment. This concept has been applied to healthcare information technology applications (Jensen & Aanestad, 2006). Jensen and Aanestad (2006) applied the sensemaking decision model to the adoption of a new information technology. For this model, understanding reality as it is socially constructed is initiated by a self-conscious sensemaker. This person seeks to define and/or maintain their identity within their environment and team. Particular attention of this model is paid to the way in which people notice and extract cues from the environment. Interesting to note, this model allows people to explore the affordances of a technology in relation to their own practices. This model is interesting because it analyzes a decision-making process in healthcare accounting for more than one decision maker and it takes into account technological affordance. The disadvantage of this model is it can be difficult to apply in situations where information and cues are not carefully organized. In an emergency department, it is somewhat unrealistic to assume that information will always be available, in good quality, and organized.

Humble decision-making is an adaptive approach that is utilized when one has limited knowledge from a situation (Etzioni, 1989). First, a judgment is made and a direction is decided. Next, incremental decisions are made. In healthcare an example would be a doctor that has general knowledge and a sense of what needs to be accomplished. She or he uses what information is available to hone in on a specific decision then tentatively prescribes one treatment and then tries another if the previous treatment fails until the right treatment is found. There are several methods for adapting with partial knowledge, one of which is not to make permanent decisions.

Naturalistic decision-making (NDM) approach focuses on developing detailed, ecologically valid and descriptive models to analyze decision-making performance (Patel et al., 2008). It is based on qualitative and quantitative methodologies. NDM research is conducted in real-world settings where stress, time pressure, and risk are a part of the analysis that leads to a better understanding of human decision-making processes in critical systems.

The involvement of emotions in healthcare decision-making is a difficult issue to assess but it warrants some analysis. The benefits and disadvantages of emotional influence in decision-making are numerous. Channeling emotions can be beneficial to a person needing to persuade a decision maker or the group. However, emotions can also cloud some issues. Of course, there is the question as to whether or not a person can completely take emotions out of the process. Instead of taking them out, how does one come to a balance with a rational approach? The addition of technological support to the decision-making process can exacerbate the impact of emotion, or in contrast, it could help to drown out emotional impact. An example of technology and emotional impact in decision-making is the "evaluative display" design (Hibbard & Peters, 2003). The evaluative display design provides the evaluation of information in simpler format. Evaluability is another approach to decision support system design where the visual display of information is designed to reduce cognitive effort with cues to transform the information to an evaluative good/bad scale (Hibbard & Peters, 2003). If a user is unfamiliar with the exact meaning of a measure (i.e. a measure of quality of care, expressed by the percentage of people satisfied with their care), however they are familiar with the numbers used (i.e. a medication that has a 2% elevated risk of stroke) there is no emotional meaning or understanding for the user. The authors go on to state that when emotion is taken out of the decision-making process, the information is no longer weighted properly (Hibbard & Peters, 2003). The authors conclude by stating that if the evaluability of information can be modified then the decision-making process will be simplified because the information that has been provided is now in an apples-to-apples comparative format.

This section has reviewed the decision-making process and the influences of attention, cognition, mental workload on that process. Cognitive support improves decision-making and can be provided by technologies and the following section will review one such technology: information visualization.

2.3 Information Visualization

The purpose of this section is to provide a comprehensive review of literature on information visualization. The general background, benefits, specific applications will be presented first. A discussion of types of information visualization techniques and visualization task classification follows. In the end, this section concludes with analyzing human factors issues relative to information visualization and evaluation methods.

Visualization is the process of forming a mental picture of something not present to the sight and can be used to explore, analyze, and interpret data (Falkman, 2000). There are different categories of visualization. According to Shneiderman (2002) there are two groups: scientific and information (2002). Scientific visualization deals with continuous variables, volumes, and surface. Information visualization is used to analyze categorical variables, discover patterns, trends, clusters, outliers, and gaps in data. Information visualization is the interdisciplinary study of the visual representation of large-scale collections of non-numerical information, such as files and lines of code in software systems, and the use of graphical techniques to help people understand and analyze data (Plaisant & Schneiderman, 2004). In contrast with scientific visualization, information visualization focuses on abstract data sets, such as unstructured text or points in high-dimensional space. Information visualization may not have an inherent 2D or 3D geometrical structure but it is multi-dimensional. Visualization and display is important to the discipline of human-computer interaction because “the real power comes from devising external aids that enhance cognitive abilities” (Tufte, 1991). This means the brain intuitively processes information and gathers understanding from the presented visual images. According to Tufte (1991), the purpose of visualization is “insight, not pictures.”

For most persons, a visual representation of data is easier to use and/or understand than a textual description or string of numbers, particularly when the dataset is large (Plaisant & Schneiderman, 2004). There are several data types: 1D linear, 2D map, 3D world, Multi-dimensional, Temporal, Tree, and Networks. The visual-information-seeking-mantra is “overview first, zoom and filter, then details on demand” which is associated with one of many task classifications for information visualization (Plaisant & Schneiderman, 2004). These various task classifications are important when developing a system that will support those tasks and utilize information visualizations.

Decision-making in complex systems, such as hospital emergency departments, is dependent upon personnel having the ability to recognize patterns with the large amount information generated from patient care. Without the proper structuring and delivery of this information, information overload becomes a significant issue. The issue of information overload has increased as the growth of accessible information, efficient information filtering, and sharing is needed to face complex decisions (Chen, 2004). Information visualization has been described as the use of interactive visual representations of abstract data to amplify cognition (Card, Mackinlay, & Shneiderman, 1999; Spence, 2001). Visualization has the potential to provide valuable assistance for data analysis and decision-making tasks by amplifying cognition in five categories as seen in Figure 2.4 (Card et al., 1999; Tory & Torsten, 2004).

2.3.1 Benefits of Information Visualization

Information overload has become a significant issue as the steady growth of accessible information and efficient information filtering and sharing is needed to solve a problem (Chen, 2004). Several technologies exist for implementing a solution. Visualization has the potential to provide valuable assistance for data analysis and decision-making tasks by assisting humans (Tory & Torsten, 2004). Representing information visually can support underlying cognitive processes in decision-making. The benefits of information visualization have been described and detailed in many pieces of literature. This sub-section will summarize the overarching themes among the many benefits and give examples from literature to support them.

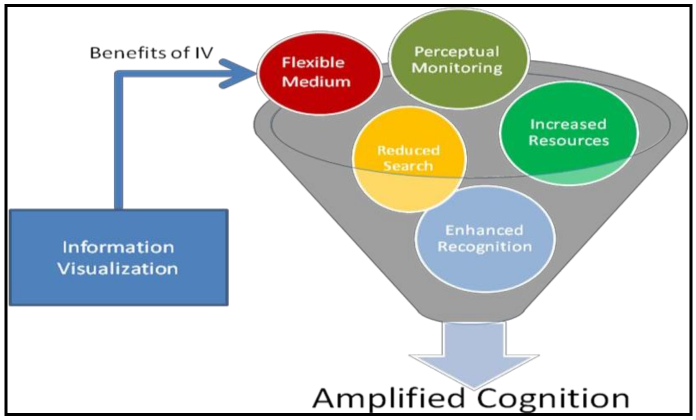


Figure 2.4: Benefits of Information Visualization for Cognition

This technology offers to increase cognitive resources with the benefits of parallel processing, offload of work to the perceptual memory, external memory, and increased storage and information accessibility (Card et al., 1999). Tasks that would typically require complex perceptual operations can be completed using simple perceptual operations with the help of information visualization, reducing the load on perceptual memory. Any time a decrease in the demands on the human memory can be obtained, that is a huge benefit to the user. Another benefit of this technology is its ability to store vast amounts of data in an easily accessible form, increasing productivity in an analysis.

Information visualization offers the advantage of reducing search time by grouping related information, storing large quantity of data in a relatively small space, and offers internal structuring of data and tasks. In an empirical study of InfoZoom, an information visualization tool for data analysts, the researchers learned that the participants preferred the easy extraction method for charts and reports provided by the tool (González & Kobsa, 2003). The users stated that searching for data was easier with this tool because it provided immediate visual feedback.

With information visualization, users can rely on recognition and not have to recall a piece of data from their memory. The abstraction and aggregation of data where the tool will omit irrelevant data and gathers up the relevant data allows increased visibility for higher level patterns that the users can recognize. In the empirical study by Gonzalez and Kobosa (2003), the data analysts were quickly able to discover patterns in their data that were not clearly visible or straightforwardly obtained with their current tools (2003).

Information visualization adds the additional benefit of utilizing pre-attention visual characteristics in data representations. This benefit allows the monitoring of a large number of possible events. If a supervisor had to monitor several hundred data points for anomalies, a visualization tool could use a color-coding scheme to make those anomalies stand out from the rest of the data to the user. This allows humans to see the “little picture” and see it in relation to the “big picture”.

Tools for information visualization allow users to explore interactively data by varying parameter values. This manipulation by the user has the potential to allow different patterns to be recognized (Card et al., 1999). In the empirical study with InfoZoom, the data analysts complimented on how quickly they were able to include or exclude certain data points for visualization which allowed them the ability to concentrate on specific areas finding data quicker (González & Kobsa, 2003). Information visualization can allow users go beyond their routine when necessary because of the technology’s flexibility.

When users view images with a visualization technique, they can compare that image with their own current mental model of the data. If there is a conflict in that comparison, they can adjust either their mental model or their understanding of the image. This is an advantage in diagnostics and hypothesis generation and testing (Card et al., 1999). When an operator has a particular hypothesis about the cause of a problem in his system, based on the features of the visualization, he can either confirm his original hypothesis or generate a new one.

For all the benefits that IV offers, there are some issues. These issues justify why a valid and reliable assessment methodology for IV techniques in healthcare is needed. A set of guidelines can make it less difficult for developers to design effective clinical decision support systems that use visualization *and* that users will want to adopt. Plaisant and Shneiderman (2004) describe several issues:

- Import data: organizing the input data, correcting format, filtering out incorrect items, normalizing attribute values, coping with missing data
- Combine visualization representations with textual labels: labels should be visible without overwhelming the display or confusing the user
- See related information: additional information is often needed to make meaningful judgments
- View large volumes of data: prototypes have problems with more than few thousand items or maintaining real-time interactivity with larger numbers of items
- Integrate data mining: data-mining researchers believe that statistical algorithms and machine learning can be relied on to find interest patterns. Statistical summaries can hide outliers and discontinuities but data mining can point users in

the right direction

- Collaborate with others: users verify assumptions by collaboration with others and also need to share in order to convince others of a significant finding.
- Achieve universal usability: diverse group of users (disabilities, technical background, technical accessibility)

In summary, visualization systems can have several roles for assisting in decision-making and have a diverse range of benefits. The technology extends human cognitive abilities by allowing a decrease in many areas of mental workload like offloading work to perceptual system, acting as an external memory, and enhancing recognition. It allows users to organize and share ideas through grouping and data structuring. Visualization has the possibility to improve mental models and insight. However, it should be noted with all the benefits there are potential downsides to information visualization (IV) if a system is improperly designed. An inadequately designed system can increase mental workload and task duties if human factors design guidelines are not taken into account.

2.3.2 Applications of Information Visualization in Healthcare

Rxplore is a visualization tool by Jon Duke that is used to help reduce the risk of drug side effects and bad interactions in patients (Duke, Faiola, & Kharrazi, 2009). This tool, seen in Figure 2.5 is an example of a visualization tool used for medication tracking and decision support. The application allows for an aggregated view of a single patient's records and flags for drug interactions. The typical hard copy drug databases are usually dense and hard to read making it challenging to look up drug interactions and compare back to a patient record. This tool reduces decision time and increases decision accuracy.

However, there are still limitations in terms of translating medical terminology like “occasionally” or “some” into an actual quantitative data to be displayed with this technique (Whitney, 2010).

Figure 2.5: Screenshot from Rxplore (Duke, Faiola, & Kharrazi, 2009)

A research team at Columbia University used a network mapping visualization technique to study 161 different diseases among 1.5 million patients (Rzhetsky, Wajngurt, Park, & Zheng, 2007). The tool is a combination of bioinformatics and medical informatics. Bioinformatics uses computing to work on molecular-biology problems like analyzing gene expression, while medical informatics uses computing to process patient records (Bourzac, 2007). This tool allowed the researchers to determine correlations among groups of diseases seen in Figure 2.6. In particular, they were able to view a correlation between schizophrenia, bipolar disorder, and autism that may lead to discovering the group of specific genes responsible for those diseases.

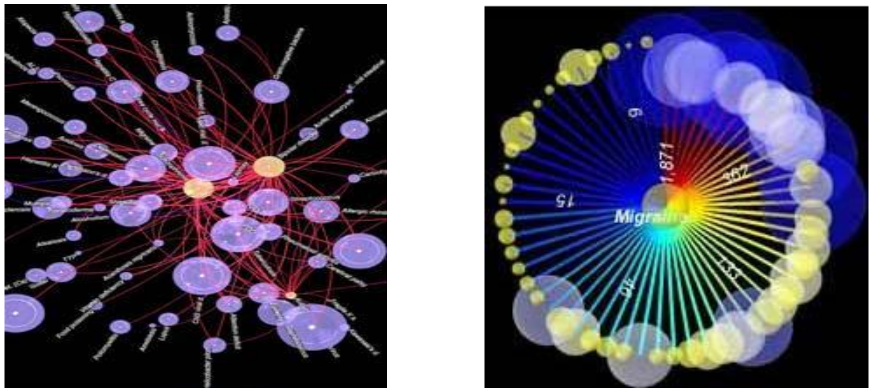


Figure 2.6: Disease Mapping Using Clusters (Rzhetsky, et al., 2007)

Research has been done to indicate the need for a decision support tool for doctors that have to make the decision to take the patients off a particular experimental treatment (Cheng, Shahar, Puerta, & Stites, 1997). The study's main focus is developing a tool that combines a domain structured knowledge base, a domain specific abstraction of raw temporal data, a way to display the abstracted data, and allowing the doctors navigation to see how the program filtered and produced the abstracted data. To build this tool the authors combined pre-existing technologies to create the program and a user task analysis to create the user interface. The methodology used to evaluate this new tool was the evaluation of the prototype by potential medical personnel in a usability pilot study consisting of a training period, a few example tasks, and a survey. Overall the findings were positive in that the users were quick in the tasks and very enthusiastic about the tool. This article gives guidance in user interface development for a health care decision support tool. The article also provides background information on medical decision-making process for seriously ill patients and discusses how to visualize the dimension of time with abstract data for a patient.

There is a vast amount of data available on medical information systems; however, the organization of this data is by source and this does not provide adequate support to the decision-making process of doctors. A particular research study took this concept of data organization and developed a tool that would organize the data by medical diseases and conditions to support higher level cognitive processes in decision-making (Bui, Aberle, & Kangarloo, 2007). The authors created a system that displays the data using a problem-centric temporal pattern. The result of this development process is a comprehensive framework called TimeLine. TimeLine is used for accessing the emergency medical record as a chronological, problem-oriented display. The TimeLine framework maps, reorganizes, and transforms data from multiple heterogeneous sources into higher-level logical views for the purposes of visualization.

Difficulties have been identified in the porting of desktop visualization applications to mobile devices with a small screen in the study by Chittaro (2006). The main focus of the study was to explore current trends in mobile visualization application research in the context of a map-based application. The authors looked at visualization applications with several types of data. The results of the analysis concluded that visualization is a powerful tool that will make many mobile applications more intuitive and productive in healthcare settings. Many doctors at hospitals and clinics already utilize mobile devices to support their tasks and decision-making and the addition of visualization is an area warranting further research.

It has been shown in recent research that there is a slow development between the combination of medical images and electronic medical records despite all the recent advances in technology. A study was completed to develop a set of standards to assist with the integration of digital images and digital text records (Ratib, Swiernik, & McCoy, 2003). The study of Ratib et al. (2003) uses a chronology of picture archiving communication systems (PACS) and its incorporation into the electronic patient record providing a review of current technological trends in electronic healthcare decision support systems.

2.3.3 Information Visualization Techniques and Task Taxonomies

There exist hundreds of visualization techniques and several research studies dedicated to categorizing them. With hundreds of information visualization techniques, there are most likely just as many categorization schemes. Keim groups visualizations by stacked display, dense pixel, standard 2D/3D, geometric display, and iconic display techniques (Keim, 2002). Another study generated a periodic table of visualization methods, using the organizational style of the original periodic table of elements show in Figure 2.7, grouping the techniques by their inherent properties (Lengler & Eppler, 2007). Kosara uses two categories of visualization, pragmatic and artistic, seen in Table 2.1 (Kosara, 2007). Kosara focuses on whether the visualization is presenting an idea (artistic) or whether it is providing new insight into data.

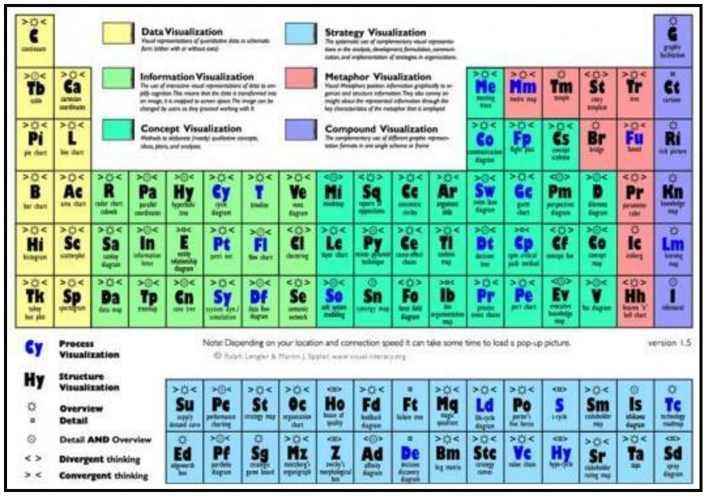


Figure 2.7: Periodic Table of Visualization Methods (Lengler & Eppler, 2007)

Table 2.1: Description of Visualization Categories from Kosara (2007)

Visualization Type	Description	Properties
Pragmatic	Visually represents data to gain new insights.	*Communicate data*. The main point of this type of visualization is to effectively and efficiently make the user understand the data.
		Visual efficiency. The means by which visualization works, and what makes it so interesting, is that it uses the visual channel to convey a lot of information.
		Data is given. Pragmatic visualization is not concerned with collecting data, though it often requires cleaning and/or interpolating data, and preprocessing it in various other ways.
Artistic	The main goal is to show an idea.	*Communicate concern*. The data is a vehicle to communicate deeper concerns or ideas.
		Visual effectiveness is not an issue. Many artistic visualizations are not designed to be effective, but are either strongly based on metaphors (to an extent that can hurt perception) or about the exploration of a form.
		Data is collected. Because the existence of the data is often a part of the message, the collection of the data is an important part of the work.

Several studies introduced a visualization task classification. Some task taxonomies sort visualizations by purpose, others by nature of the data, or domain, etc. Shneiderman's is one of the most popular and can be seen in Table 2.2 (Shneiderman, 2002). This particular taxonomy is based on the functionality of what the visualization tool can provide to a user. Wehrend and Lewis (1990) developed a taxonomy that was more domain-dependent, yet still allowed a high level of generalizability over other domains (Wehrend & Lewis, 1990). Their task taxonomy can be seen in Table 2.3. Zhou and Feiner (1998) developed an extension of their taxonomy.

Table 2.2: Task Taxonomy by Shneiderman (2002)

Task	Description	Examples
Overview	Gain an overview of the entire collection	Zoomed-out view with adjoining detail view
Zoom	Zoom in on items of interest	Users have some interest in a particular portion
Filter	Filter out uninteresting items	Sliders, buttons, control widgets, brushing, and linking
Details-On-Demand	Select an item or group and get details when needed	Once a collection has been trimmed a few times, it's easier to browse details about a group or individual items
Relate	View relationships among items	Can show relationships by proximity, containment, connected lines, and color coding
History	Keep a history of actions to support undo, replay, and progressive refinement	Typically users need several actions to get the desired outcome
Extract	Allow extraction of sub-collections and of the query parameters	Users will want to share the item/set of items with others via email, print, publish, etc.

Table 2.3: Task Taxonomy by Wehrend and Lewis (1990)

Task	Description
Locate	Interaction techniques that allow the user to find special data entries.
Identify	The user is asked to describe an object that was not necessarily known previously.
Distinguish	This action allows distinguishing between different values of the same variable
Categorize	Define divisions that can be used for sorting displayed objects
Cluster	Techniques that allow us to determine whether data entries are clustered or not.
Distribution	User needs to describe the overall pattern while cluster merely asks that the set be detected.
Rank	Users could be asked to indicate the best and worst cases in a display.
Compare within entities	Describes tasks in which a user is called upon to decide something based on the attributes of similar objects.
Compare between relations	When different entities are used as the basis of comparison, the comparison between relations operator is used.
Associate	User needs to form relationships between objects in a display.
Correlate	Discern which objects share attributes in a multiple object display.

2.3.4 Human Factors Assessment in Information Visualization

A potential role of human factors in visualization research is to provide assessment tools and design guidelines based on user-centered designs. The availability of a new technology or imaging technique is not always a suitable match with the user's tasks or goals. Although human factors design is neglected in visualization designs, its purpose is to design artifacts that are usable and useful to people. Human factors can provide designs and prototyping tools, assessment tools, usability testing guidelines, and more for information visualization developers.

A particular study on human factors and technology adoption into a community focused on evaluation techniques for information visualization theorizing the use of quantitative and qualitative data (Carpendale, 2008). A qualitative examination is better for determining the interaction among factors from a holistic approach, for example, interviews, unstructured observations, think-aloud protocol, or participant opinion. Typically, quantitative examination is more related to realistic settings and can be applied to all types of studies, for example, performance measures such as time and error are very common quantitative measurements. This is not to mention one type is better than another, in fact, both should be used together to come up with an assessment that will provide relevant information for improving a visualization technology. When quantitative metrics, such as time and error, contradict user opinion then a re-evaluation of what is most important to the users is appropriate. This re-evaluation can lead to more information about how to design a more suitable tool.

Care should be taken when evaluating information visualization techniques because they can encourage cognitive friction. An intuitive information display may still cause friction with the user’s cognition because it does not conform to the user’s past experience with the variables represented, for example, a technique may not correspond visually to the variables it is displaying. In order to discourage cognitive friction, the developers must have a clear understanding of the users’ mental models (Faiola & Matei, 2010). A particular study completed a usability evaluation of a computer-based patient record information system (Thyvalikakath, Monaco, Thambuganipalle, & Schleyer, 2008). The results of their study explain that the organization of the clinical information was not optimal because it did not represent the users' mental model of how the system should be organized, and this resulted in an unnecessary degree of cognitive friction. As a solution to cognitive friction, Cooper suggests the goal-directed method where the behaviors, patterns, and modes of product use are analyzed (Canossa & Drachen, 2009).

There are some research studies that have evaluated multiple techniques at the same time; however they are not based on medical data, typically business or financial domains (Pillat, Valiati, & Freitas 2005; Wei, Shi, Tan, Sun, Lian, Liu, & Zhou, 2010). In another case, the study evaluated a single medical visualization technique and its design elements or novelty (Pohl, Rester, & Wiltner, 2007). In this dissertation, a usability evaluation comparing multiple visualization techniques was conducted collecting quantitative subjective and objective data from an empirical study, as well as qualitative data from a case study to produce a set of guidelines for use of medical information visualization tools.

Usability and effectiveness of an information display is important to supporting cognitive efforts and decision-making. If the display adds to mental workload or makes a task more daunting, it becomes a part of the problem, not a part of the solution. By developing an evaluation methodology for information visualization techniques, the design of the display for a clinical decision support system can be improved before it is put into the field.

User testing is a significant part of evaluation. There exist several challenges in improving user testing. Most empirical evaluations of visualization techniques include basic tasks (Plaisant, 2004). The experiments typically are composed of "locate and identify"; however, more complex tasks such as comparison, association, correlate, etc. are unusual. Other issues with evaluating information visualization include users looking at the same data from different perspectives or having to create questions before looking at the visualization.

Performance measurements like time and accuracy are commonplace while subjective and learnability measurements are gaining in popularity. Other possible measurements of performance include abandonment rate, which is the rate the subjects choose not to answer a question for a visualization task (Cawthon & Moere, 2007). Research also indicates that overall performance on a set of tasks is less accurate than measuring performance on each individual task (Plaisant, 2004). Plaisant's (2004) research indicated that different techniques performed differently for tasks. A bias was introduced when a composition of tasks favored one technique over another technique (this only applied to performance measurement). However, it is still important to note that performance does not equate to usability or effectiveness. A visualization tool may produce high performance rates, and still rate poorly, for example high mental workload.

Mental workload and its relationship with the decision-making process has been discussed previously in this research. Mental workload has three dimensions: perceptual, cognitive, and motor (Wu & Liu, 2006). It can be measured after a task is completed; therefore is non-intrusive and can be a complementary usability measurement for performance measures. Surveys are typically the most economically method and have been utilized in visualization evaluation studies (Wu & Liu, 2006). In prior visualization studies, it has been measured with multi-dimensional scales and with a univariate scale where a single number was used to correspond to the overall cognitive effort (Huang, Eades, & Hong, 2009). Mental workload is also inter-related to task complexity. Speier and Morris hypothesized that subjective mental workload varies between different display designs when task complexity is low; and when task complexity is high, the mental workload will be lower for visual query display than text-based query display (Speier & Morris, 2003). In their study, mental workload was measured with a NASA-TLX survey. Interesting to note, Huang et al. (2009) explained that an inappropriate visualization might actually increase mental workload. However, those authors also stated it is not typical that a visualization results in improved decision accuracy and performance time. Usually, there is a speed-accuracy trade-off where the user's accuracy increases but takes more time.

Cognitive tasks often have an impact on the users. A particular research study concluded that some cognitive tasks cause drivers to change their visual search behavior while other tasks do not produce those changes (Recarte & Nunes, 2003). It is possible to extend this theory to this research in examining the visual search behavior of users when interacting with different information visualization techniques. Eye tracking would be assistive in this regard. Eye tracking is another evaluation tool employed in usability studies and has also been utilized in assessing information visualization techniques (Huang, 2007). Eye tracking technology provides many quantitative metrics such as fixations, sequence, observation time, etc. for nonverbal signs of user attention. A research study by Pirolli, Card, and Van Der Wege (2001) used eye tracking to analyze visual search patterns for interfaces with different information scent. The visual search patterns were used to interpret users' attention spotlight. This is an example of how eye tracking describes users' underlying cognitive process: attention. Attention is not something that can be measured directly, so eye tracking technology provides the opportunity to investigate what researchers cannot see under the surface. Other studies in this area measured the number of fixations between different visualizations (Pirolli, Car, & Van Der Wege, 2000; Ziegler, 2002). Eye movements provide an abundance of user data however, there have been some issues with utilizing them to study visualization techniques such as hardware/software challenges (noise, calibration) or large variation in personal search strategies (Salojärvi, Kojo, Simola, & Kaski, 2003).

The purpose of this section was to provide a broad analysis of literature on information visualization. Applications, techniques, and tasks taxonomies were highlighted, in addition to the benefits of information visualization. A review of evaluation studies and methods for information visualizations was also presented.

2.4 Conclusion

This chapter has provided a review of literature on healthcare information systems, decision-making in healthcare, and information visualization. Particular attention was given to information visualization in medical domain and their evaluation. Assessment methods for IV are not new concepts; however, there has been very little emphasis on comparing multiple medical data visualization techniques collecting qualitative and quantitative data. An assessment methodology is necessary to develop the set of guidelines for this research's overall objective. As evidenced in this chapter, an adequate assessment should be multi-layered and cover a variety of data collection including objective, subjective, quantitative, and qualitative. The next chapter will present the planning phase of this research, which includes a user requirements analysis and the test bed development.

CHAPTER 3

EVALUATION METHODOLOGY FRAMEWORK

This chapter presents the framework for the evaluation methodology used in this research, the user requirements analysis, and the development of the test bed. The evaluation methodology framework has four phases as seen in Figure 3.1. Phase-I is the planning phase, which includes the user requirements analysis, and test bed development steps. The requirements analysis was an on-site case study of an emergency department. The final design iteration of the test bed was developed in Visual Basic. Phase-II is the testing phase comprised of a heuristic evaluation, pilot study, empirical study, focus groups, and a case study. The third phase is the analysis of the data gathered in phase-II. The fourth phase concludes with creating the guidelines for information visualization developers based on the data analysis.

3.1 Planning

A user requirements analysis was completed with an analysis of patient information flow at a local emergency department to understand the tasks and decisions of a healthcare professional supported with a clinical decision support system (Balogun, Chenou, Jenkins, & Park, 2009). Activity theory was applied to the information flow analysis to determine the relationship between an activity, the actions, and the operations at each point in the flow of patient information. With the requirements information gathered, the test bed for evaluating the visualizations was first designed using a simple prototyping tool to examine the features. This interface underwent several iterations in design improvements throughout the planning and testing phases.

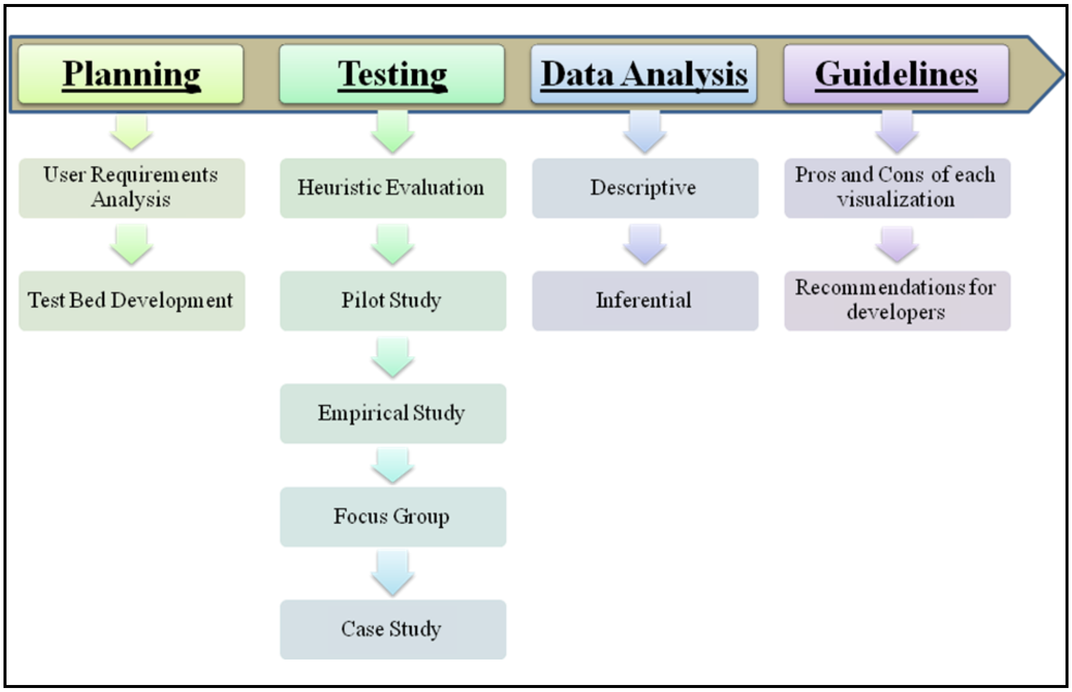


Figure 3.1: Framework for Evaluation Methodology

3.1.1 Information Flow in Emergency Department

An information and activity analysis were conducted at Wesley Long Hospital in Greensboro, NC. The results of that report are used in this research to examine information flow patterns and work analysis in an emergency department.

The emergency department (ED) located within a hospital is responsible for initially examining and taking care of patients' initial treatment for a variety of categories of illnesses and injuries; some of which may be life threatening while others require non-emergency care. The purpose of the ED is to make a rapid assessment and manage critical illnesses. In order to deliver a rapid assessment, patients need to be prioritized. This process can be illuminated by providing the staff with an information system where they can see all necessary data in one place, and not in several places. Errors in medical care are less frequent, but can potentially be very costly. Errors can potentially occur at any point during a patient's visit, but they can be related to the staff's access to accurate and reliable patient information. The staff from an emergency department needs to be able to detect, track, and fix errors in patient information.

Figure 3.2 depicts the flow of a patient going through an emergency department (Balogun, Chenou, Jenkins, & Park, 2009). Patients can arrive in one of three ways: walk-in, by ambulance, or by helicopter. Patients that arrive from ambulance or helicopter go straight to assessment and do not go through registration and triage. For walk-in patients, they first go through registration and triage where they are given a preliminary assessment. The resources during the registration process include the patient information forms, medical personnel available, and the patient information database. The information is transferred from patient to secretary or administrative assistant, to the nursing assistant, and then to the computer. A possible delay could be incurred during information processing. The resources available during triage include the licensed practical nurse, lab technician, and computer. During triage, the patient information from registration is verified and new information concerning their reason for visiting is recorded. Possible delays during triage include assessing the patient status and processing information. The information is used to create an initial preliminary assessment and give each patient a priority status of high or low.

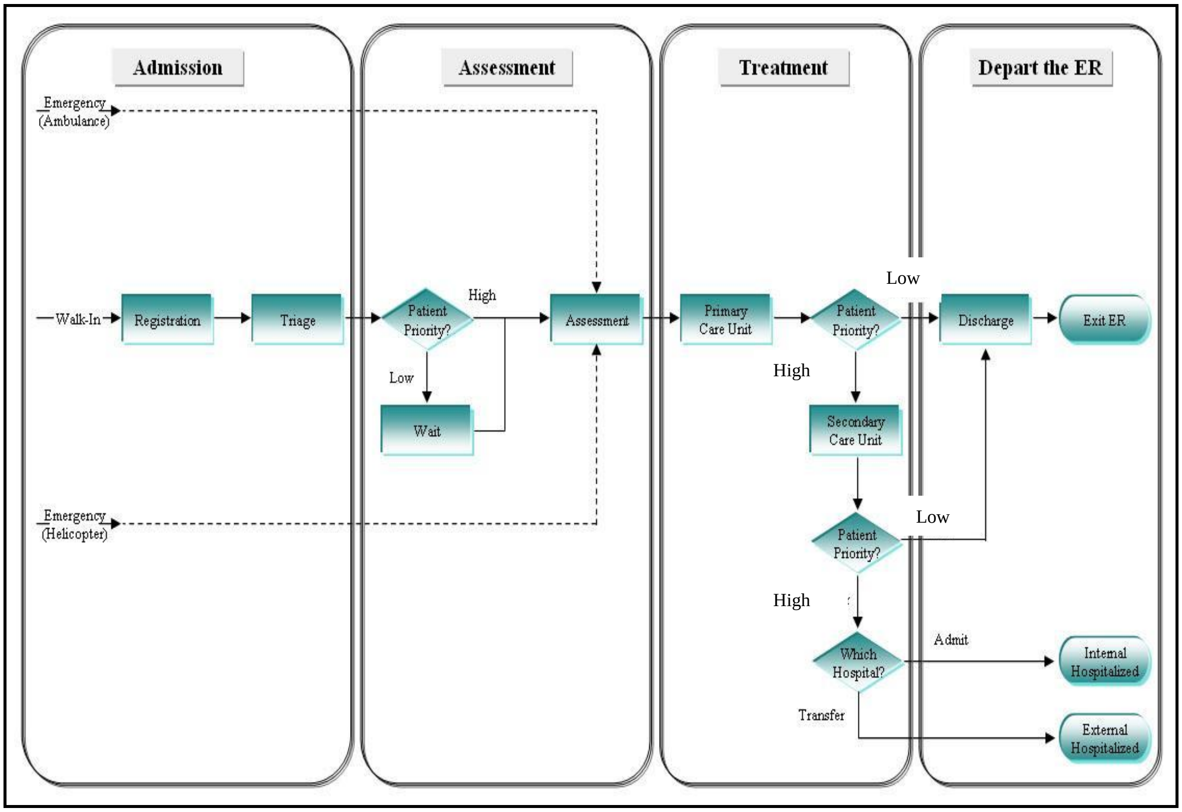


Figure 3.2: Patient Information Flow for Emergency Department (Balogun, Chenou, Jenkins, & Park, 2009)

Low priority patients are sent back to the waiting room for the next available doctor while high priority patients are sent to the next phase of assessment. The resources of the waiting room include the square footage of the room and the number of chair available. During this time, a patient waits until asked to move to the primary care unit. Associated delays in the waiting room include patient priority and the medical staff availability.

High priority patients are sent to the assessment area where they may be seen by either a doctor or nurse practitioner. In the assessment area patient vitals are measured, more patient information is recorded, and labs may be ordered. Possible delays in assessment are due to unoccupied care units and room availability. Resources of assessment include room capacity and availability of medical staff.

After an assessment is made by either a doctor or nurse practitioner and a diagnosis is given, the patient is sent to the primary care unit where they are treated by a physician. In this treatment area, they are given medications or treatments for their diagnosis. Possible delays in the primary care unit are processing and waiting for lab results, waiting for prescriptions or other treatments, and accessibility of the physician. Resources for the primary care unit include equipment, beds available, medical staff, and on-hand medication. At this stage, if the physician decides that the patient's priority is still high, the patient can be sent to a secondary care unit for further treatment. The resources, transactions, and delays of the secondary care unit are similar to the primary care unit.

If the patient status is changed to low priority, the patient is discharged and sent home. From the secondary care unit, there are three paths for the patient. He or she can be discharged and sent home if their health status has improved, admitted to long term care in the hospital if their health status has worsened, or transferred to another hospital for further specialized treatment. Resources of the discharge process include the payment clerk and computer. At this time, the patient pays for services rendered, receives print out of services, and discharge instructions. Possible delays during discharge (or admitting/transferring) include processing information and availability of staff.

The medical personnel involved throughout a patient ED visit have different roles and responsibilities. The admissions staff is responsible for taking the patient's basic information and assigning them a patient identification number. These persons can be general administrators or nurses (CNA, LPN, RN, nurse practitioner). In the admissions stage, patient information includes name, contact information, primary physician information, insurance information, current allergies, and medication. This information is then transferred to various members of the medical team such as doctors and other nurses. In triage, the medical staff needs to know what medications the patient is currently taking to avoid dangerous medication interactions.

In the next stage, patient assessment, the staff is responsible for evaluating the patient's condition. These persons can be a nurse practitioner or doctors. They are assisted by nurse assistants and lab technicians. In the assessment stage, patient information includes more details about their particular reasons for coming to the ED that day. This information includes history of the problem, problem location (head, arm, chest, etc.), lab orders and results, possible diagnosis, list of symptoms, and any notes possibly made by other medical staff about the patient's condition. This information then transfers to a care unit where the patient will be treated according to their diagnosis.

The treatment staff, comprised of primarily doctors and nurses, is responsible for treating the patient's injury or illness. In the treatment stage, patient information includes details of medication given to the patient in the ED, symptoms, and diagnosis. This information can transfer to a secondary care unit if the doctors in the primary care unit determine that the patient needs further treatment; otherwise, patient information would next transfer to discharge.

The discharge staff is responsible for billing related to the patient's current ED visit and scheduling follow up appointments. These persons can be general administrators or nurses (CNA, LPN). In the discharge stage, patient information includes a list of charges for tests, medications, and any other activities related to their visit.

3.1.2 Activity Networks of Patient Information in Emergency Department

Activity theory is a multi-disciplinary framework for examining different types of human practices as developmental processes, with both individual and social levels interacting simultaneously (Kuutti, 1996). It is used in human-computer interaction as a frame to assess a design by analyzing "computer-supported activity of a group or organization" (Kaptelinin, 1996). As described by Leontiev, an activity "satisfies a need, it provides a motive and is driven by the goals of persons" (Leontiev, 1978). An activity is comprised of actions and is the basic unit of analysis in activity theory.

Activities that are characterized by physical or cognitive processes are set in order to obtain a specific goal. From an activity, several actions can be derived. From an action, several operations can be derived (Nardi, 1996). An operation, or method of operating, is comprised of the conditions for how an action is carried out. For the context of designing an interface, understanding the hierarchal relationship between activity, actions, and operations of a user for a system assists the designer in developing a metaphoric representation of the interface objects and actions (Plaisant & Schneiderman, 2004). The three units of analysis in activity theory are activity, action, and operation. Each is related to the other in a hierarchal association.

Activity theory was applied to each process in the emergency department to further illustrate the structure of activity for information flow. Within each activity network, the subject, tools, rules, community, division of labor, and the outcome are described to illustrate the activity associated with processing patient information. Each network includes resources, transactions, delays, actions, and operations during each phase of the emergency department process for a patient. All the information gathered throughout the user requirements analysis was used to understand the users' tasks to assist in developing the test bed interface and tasks for the empirical study.

3.1.3 Test Bed Development

The test bed is the interface that was used in the evaluation of the visualizations. This interface will need to display the visualization, allow interaction, and include the typical user tasks. It will also include the usability and mental workload surveys. The development includes collecting data, determining visualizations, and designing the graphical user interface (GUI).

3.1.3.1 Data Collection. As was previously discussed, there exist large quantities of healthcare information available to medical personnel in a hospital. When a patient arrives in an emergency department at a hospital, the information obtained is relative to the patient's current reason for the visit. In regards to this research, the interest here is in the type of information. Patient records are a mixture of numerical and textual data formats. Numerical data can be nominal, ordinal, or metric. Monitoring and identifying trends are the most common uses of visualizations of medical data. RODS, a biosurveillance system was developed and evaluated in 2004 (Chapman, Dowling, Ivanov, Gesteland, Olszewski, Espino, & Wagner, 2004). The data came from over 100 hospitals in four states. The admission data included age, sex, zip code, and triage chief complaint. The medical data was used to determine public risks. In other studies, health care research has been supported by an integration of data types. In one such study there was an integration of geographic information system and spatial analytical methods (Moore & Carpenter, 1999). The system was developed in attempt to support decisions by epidemiologists that need to identify trends in health databases, such as, those used by the CDC. Another study examined environmental factors like air pollution from the EPA in connection with asthma rates from CDC (Li, Xu, Jeng, Naik, Allen, & Frontini, 2008). This study developed a system that integrated the two different data sets (environmental and medical) and used SAS to manipulate and analyze the trends in the sets.

To explore potential future outbreaks of respiratory syndromes across the state of Indiana, a predictive analytics system was developed in 2010 (Maciejewski, Hafen, Rudolph, Larew, Mitchell, Cleveland, & Ebert, 2010). The system focused on categorical spatiotemporal event data such as financial data, crime reports, and emergency department logs. The events consisted of a location in time and/or space and fit into a hierarchal categorization structure. Then, these categories were filtered by linked data and events were mapped to a specific location. The categories of data were commonly processed in two methods. The first method was a time series aggregated over some spatial location like a county, zip code, or collection station. The second method was a spatial snapshot of a small time unit such as a day or week. In either method, the aggregations were analyzed using a cumulative summation or moving average.

Visualization has not only been used to analyze healthcare data from a hospital information system but also for scientific publications monitoring disease outbreaks.

MacEachren, Stryker, & Turton (2008) developed a system to visualize articles from PubMed. The features of the system were developed based on the theory of situation awareness. The goal of the system was to support decision-making in the surveillance network of the avian flu from various agencies such as World Health Organization (WHO), World Animal Health Organization (WAHO), and Food and Agriculture Organization (FAO) of the United Nations.

***3.1.3.2 National Hospital Ambulatory Medical Care Survey*.** The Center for Disease and Control and Prevention (CDC) has a vast repository of healthcare data collected from surveys with regard to public health importance. The data are collected in collaboration with many other agencies such as the National Center for Health Statistics (NCHS). The data collected ranges from such diverse topics as aging to youth obesity. Beginning in 1973, data on ambulatory patient visits to physicians' offices has been collected through the National Ambulatory Medical Care Survey (NAMCS). Visits to hospital emergency and outpatient departments are not included. This raises two issues. First, these persons represent a significant portion of total ambulatory medical care. Second, these patients are known to differ from office patients in their demographic characteristics and medical aspects. The National Hospital Ambulatory Medical Care Survey (NHAMCS) filled this data gap from the original NAMCS (CDC-NCHS, 2009). The NHAMCS was designed to collect data on the utilization and provision of ambulatory care services in hospital emergency and outpatient departments in ambulatory surgery centers (CDC-NCHS, 2008). It was initiated to learn more about the ambulatory care rendered in hospital emergency and outpatient departments in the United States.

There are two components of the NHAMCS: outpatient departments (OPD) and emergency departments (ED). For the purposes of this dissertation, the focus was on the latter, the emergency department data. The form is shown in Appendix A: NHAMCS Patient Data Form. The dataset is based on a national sample of visits to the emergency department in non-institutional general and short-stay hospitals, exclusive of Federal, military, and Veterans Administration hospitals, located in the 50 States and the District of Columbia. This survey was not based on a sample of the population but on a sample of visits.

Before a facility participated in the study, trained interviewers from Ambulatory and Hospital Care Statistics Branch of the National Center for Health Statistics, Centers for Disease Control and Prevention visited to explain the survey procedures, verify eligibility, develop a sampling plan, and train staff in data collection procedures. The main survey instrument utilized was the Patient Record Form seen in Appendix A. The medical personnel completed the patient record forms through a random sample of patient visits during a randomly assigned 4-week reporting period. The actual visit sampling and primary data collection was the responsibility of the hospital staff for several reasons; lack of a standard form in hospitals made it difficult to train CDC field representatives, hospitals did not want field representatives to see patient identifying data, and hospital staff was more knowledgeable of the domain.

A representative from the U.S. Bureau of the Census collected the survey data at the end of the four week period as an agent for the NHAMCS. Completeness checks were done by the hospital staff and again by field staff before central processing. All medical and drug coding and keying operations were performed centrally by Constella Group, Inc. and subject to quality control procedures.

Data was obtained on demographic characteristics of patients, expected source(s) of payment, patients' complaints, diagnoses, diagnostic/screening services, procedures, medication therapy, disposition, types of providers seen, causes of injury (emergency department and ambulatory surgery center only), and certain characteristics of the facility, such as, geographic region and metropolitan status.

To ensure confidentiality for the patients, the top section of each form, which contains the patient’s name and record number, was separated from the bottom section, which contains all the data. The top section remained attached to the bottom until the entire Patient Record form was completed. Prior to collecting the completed Patient Record forms, the top section was detached and given to hospital staff. This portion was kept for 4 weeks, in case it was necessary to fix clerical errors.

The most recently available NHAMCS is 2008. In 2008, there were 34,134 Patient Record forms provided by emergency departments and 33,908 Patient Record forms provided by outpatient departments that participated in the survey (CDC-NCHS, 2009). The 2008 NHAMCS was conducted from December 31, 2007 through December 28, 2008, and consisted of a sample of 475 hospitals. Of the sampled hospitals, 79 were found to be ineligible due to closing or other reasons. Of the 396 hospitals that were eligible for the survey, 357 participated, for an unweighted hospital sampling response rate of 90.2 percent (89.8 percent weighted). Of interesting to note in the emergency department data set is that that the patient race was missing 15.3% of the time and patient ethnicity was missing 23.8% of the time. The total variable count for this data set is 403 (CDC-NCHS, 2008).

The data set was publicly available. It was accessed on June 9, 2011 from Inter-University Consortium for Political and Social Research (CDC-NCHS, 2008). The data set was downloaded in addition to a user guide, patient record form, and other supplementary documents. The format was originally in SAS and was eventually exported to Excel to be analyzed in other software programs.

The NHAMCS data set has been used in previous healthcare information studies. One particular study used the survey results to identify whether US physicians' practice patterns in treating tobacco use at ambulatory visits improved over the past decade (Thorndike, Regan, & Rigotti, 2007). Another study used the dataset to estimate the number of and the risk for emergency department visits for adverse events involving medications for senior citizens (Budnitz, Shehab, Kegler, & Richards, 2007). In a similar study, the NHAMCS was used in conjunction with another medical survey for injury surveillance and data from the Toxic Exposure Surveillance System to identify trends in self-inflicted harm medical visits. In 2008, a research team used the survey results as part of a study to evaluate an electronic emergency department system versus paper system (Buehler, Sonricker, Paladini, Soper, & Mostashari, 2008). Additionally, the NHAMCS data has also been used to examine electronic health record usage in hospital care (Linder, Ma, Bates, Middleton, & Stafford, 2007).

With respect to visualization, there have been two studies from the same research group that utilized this dataset. The researchers in that group applied their work on correlating content from multiple text/data fields using interactive visualization technologies to analyzing the NHAMCS dataset (Wei, Shi, Tan, Sun, Lian, Liu, & Zhou, 2010). This survey was used to reveal healthcare-related data patterns through the correlations between unstructured data fields (e.g., cause of injury and diagnosis) and between structured/ unstructured fields (e.g., gender and cause of injury). The same research group conducted a usability study of their particular visualization technology with business professionals. Their use of the NHAMCS data was to demonstrate the visualization interaction possible and did not examine any human factors issues related to the visualization (Wei et al., 2010). This dissertation investigated human factors issues with various visualization techniques utilizing the NHAMCS dataset.

This data set was chosen for its availability, complexity, applicability and relevance. Most of the CDC surveys are available for public use. In some cases, there may be datasets with sensitive information and those seeking use of those sets will have additional screening and protocol to access those variables. Many times in academic research, the opportunity to access data from local healthcare facilities is difficult so this was an appropriate alternative. This particular CDC data set is very complex with over 30,000 records covering several hundred variables. This type of complexity is typically necessary for using information visualization techniques. In addition, this data has relevance to the field of emergency department information systems since the data was procured from emergency departments in hospitals and clinics across the country.

3.1.3.3 Visualization Techniques. Different software was used to generate the visualization techniques to avoid bias from using a single software program to develop all techniques. SAS JMP is an extension to statistical analysis software that links datasets to visualizations for the purpose of data exploration. The software allows for the user to create several types of charts, graphs, and models. For this research, the following visualizations in SAS JMP were used from the NHAMCS dataset: 3D scatter plot, scatter plot matrix, and tree map seen in Figures 3.3-3.5. This software was acquired through university license. The tree map visualization could have also been created in Many Eyes by IBM, a publicly available visualization tool. This tool is located on a website and there is a 5MB limit to data sets. Due to the size of the NHAMCS 2008 data set, this other option was not feasible.

VOS Viewer was developed by the Centre for Science and Technology Studies (CWTS) of Leiden University. It is a publicly available computer program developed on Java platform that can be used to create, explore, and analyze maps based on network data. It is easy to use and can create several types of maps. However, for this dissertation, only the density chart (Figure 3.6) and network diagram (Figure 3.7) was used. The density chart could only be produced in VOS Viewer. The network diagram could have been produced in VOS Viewer or in Many Eyes by IBM. However, as previously mentioned, Many Eyes has a data limit of 5MB and this did not suit the needs of this research.

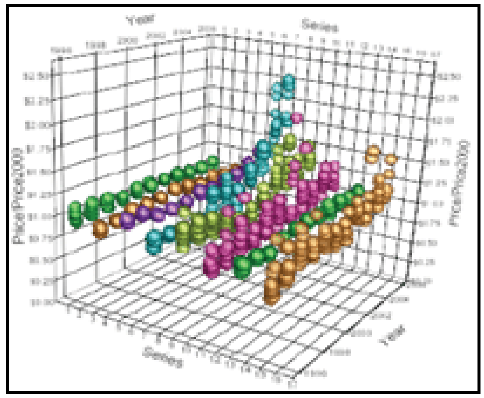


Figure 3.3: 3D Scatter Plot in SAS JMP

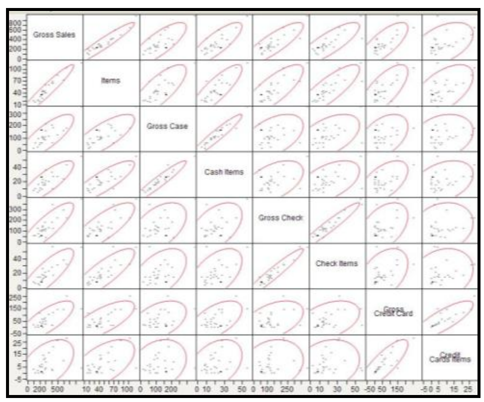


Figure 3.4: Scatter Plot Matrix in SAS JMP

Figure 3.5: Tree Map in SAS JMP

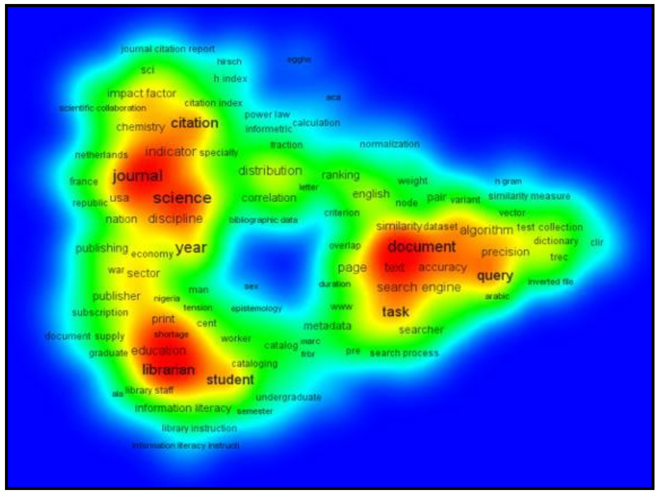


Figure 3.6: Density Chart in VOS Viewer

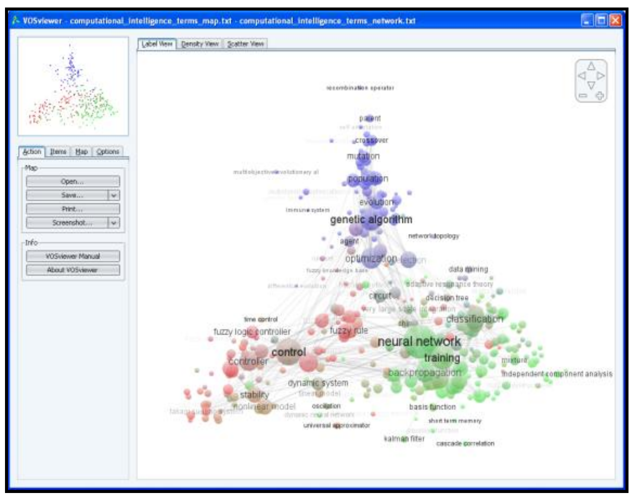


Figure 3.7: Network Diagram in VOS Viewer

XDAT is another publicly available visualization tool that allows for multidimensional analysis of a dataset using the parallel coordinate technique to display multivariate data. This tool was downloaded and installed on a computer whereby the user could freely create their own parallel coordinates using their own data set as seen in Figure 3.8. The parallel coordinates' visualization could be created in SAS JMP or XDAT. The visualization would look similar in either software. However, XDAT was chosen to produce this particular visualization because producing all the visualizations within SAS JMP would introduce software bias.

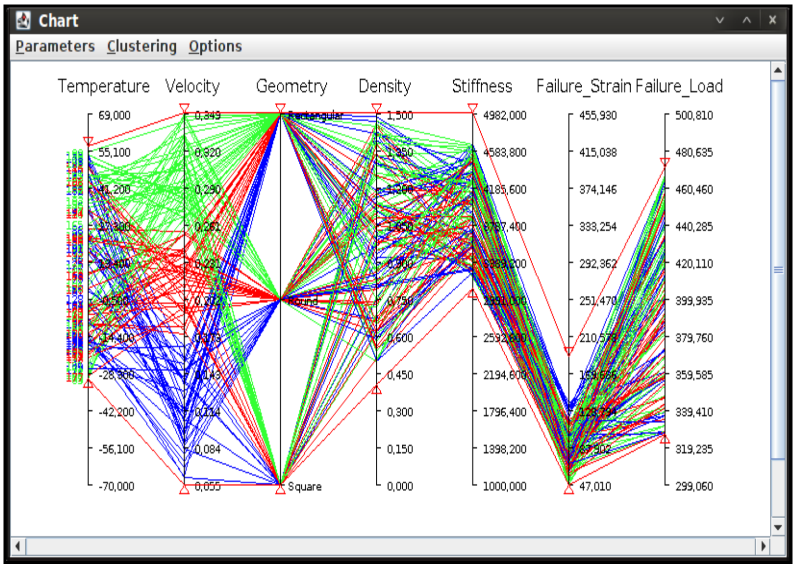


Figure 3.8: Parallel Coordinates in XDAT

<u>3.1.3.4 Graphical User Interface Prototype</u>. The software mentioned earlier was used to produce the visualizations, but the techniques were embedded in a graphical user interface (GUI) that allowed the user to interact with the visualizations for the experiments. The GUI development went through several iterations corresponding to each step of the testing phase: the heuristic evaluation, the pilot study, and the empirical study. The initial GUI was designed in GUI Design Studio by Caretta Software, a simple prototyping tool. This GUI, shown in Figure 3.9 was used for the heuristic evaluation by usability experts. Details about the heuristic evaluation will be described in Section 3.2.1.

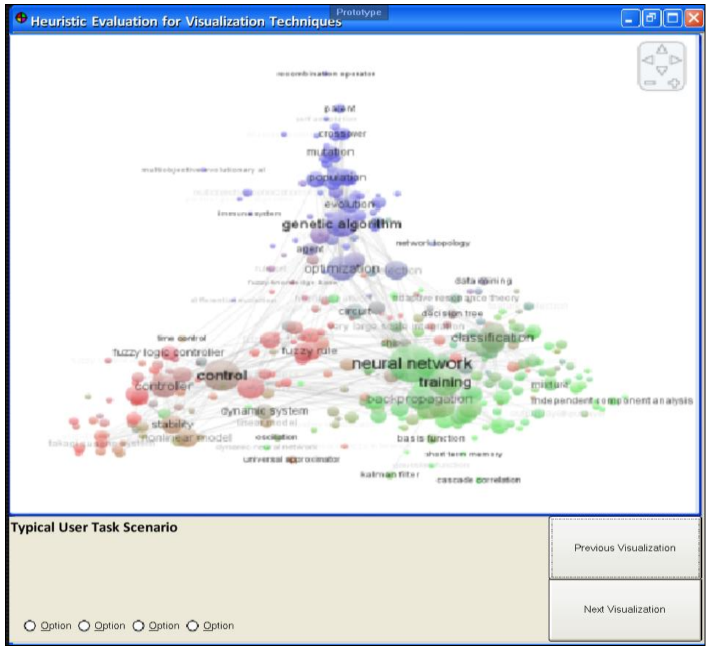


Figure 3.9: GUI in GUI Design Studio

The GUI was reviewed by usability experts before being used in the testing. The feedback from the usability experts was used to improve the GUI for the pilot study. Issues found included font size, lack of a "back" button, and clarity of visualization images. These issues were addressed in the pilot study version of the GUI. The font size was changed to size 11 and a "back" button was added to every page except the Start page, and the visualizations were saved as JPEG to reduce blurriness. For the pilot study, to add functionality, the GUI was built in NetBeans with the Java language as seen in Figure 3.10 and 3.11.

Figure 3.10: Start Screen in NetBeans

Following the heuristic evaluation, the GUI was designed using NetBeans as an integrated developer environment for Java. A start screen, seen in Figure 3.10, was created to input the subject code and session date. This was done to link the data to the subject using an anonymous code. A start button, to go to the first visualization task was added in the right hand corner. An example of a visualization task screen is seen in Figure 3.11. The visualization appears in the upper left hand corner with a button the user can press to enlarge the visualization. Below the visualization is the task scenario description. To the right of the visualization, the question text appears. Below the question, the user has the option to pick one of four radio button answers. The bottom right corner is a 'next' button when clicked will take the user to the survey screen.

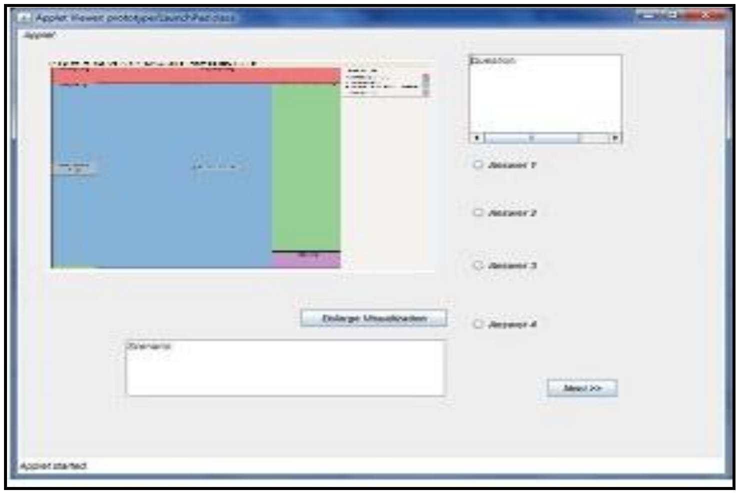


Figure 3.11: Visualization Task Screen for the Pilot Study in NetBeans

During the pilot study, issues were discovered with the GUI. Details about the pilot study will be described in Section 3.2.2. The button to enlarge the visualization frequently malfunctioned and the screen would not accommodate the enlarged image. The data collected would not properly save to a text record and often would be erased when the next visualization task appeared. Due to these issues, the GUI was rebuilt in Visual Studio using the Visual Basic language, shown in Figure 3.12. An instruction screen and finish screen were also added in the re-built version. The task scenario, question, and answers were all placed below the visualization image as seen in Figure 3.13. The survey screen depicted in Figure 3.14 followed each visualization task screen. For the case study, the survey was modified to add open-ended questions shown to the right of the visualization seen in Figure 3.15. These open-ended questions will be discussed in the case study methodology in Section 3.2.5. Although the GUI was designed in different applications, the fundamental structure and elements were all kept consistent. With each test, the GUI was improved.

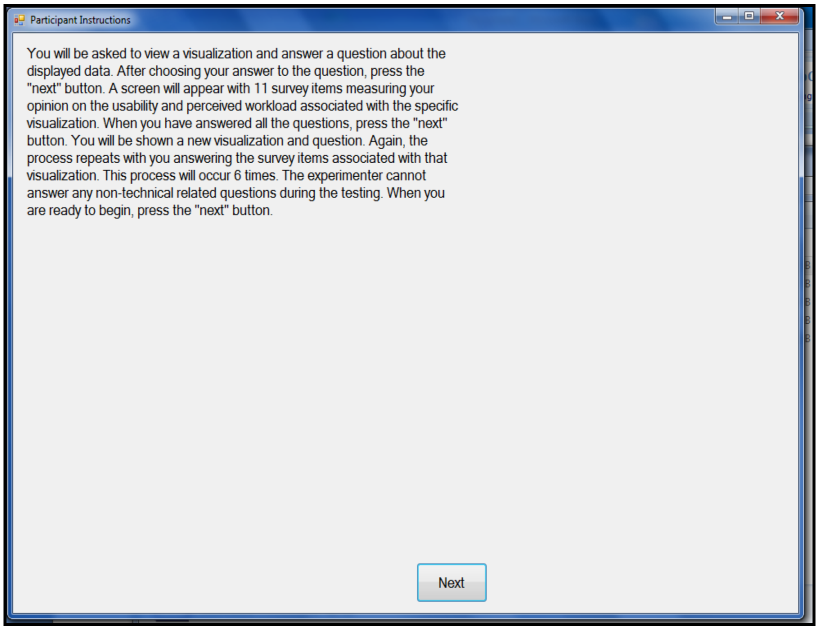


Figure 3.12: Instruction Screen for Visualization Experiment

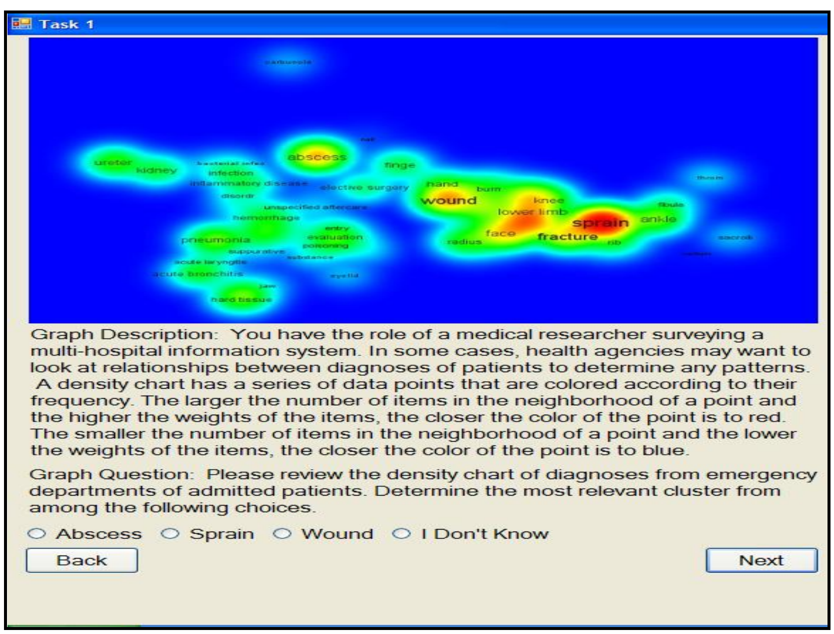


Figure 3.13: Visualization Task Screen for Empirical Study in Visual Studio

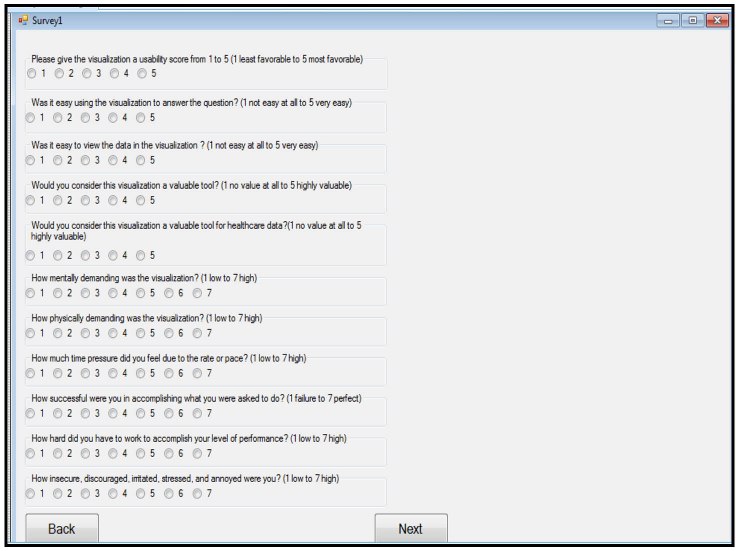


Figure 3.14: Survey Screen for the Empirical Study

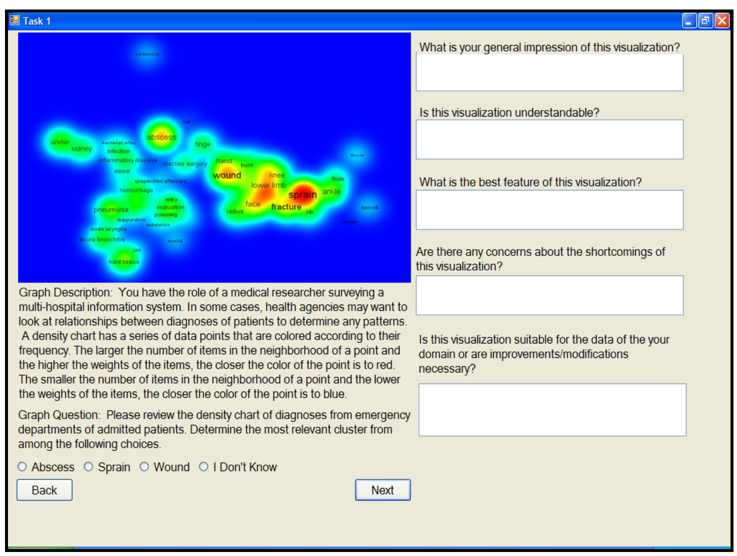


Figure 3.15: Visualization Task Screen for the Case Study in Visual Studio

3.2 Testing

The testing phase includes the following: a heuristic evaluation, a pilot study, an empirical study, and a case study. A heuristic evaluation with usability experts to evaluate the visualizations and the interface for usability issues was done. Next, a pilot study with novices was done to determine the expected ranges for the empirical study and to clarify the testing protocols with the designed tasks. Following the pilot study, an empirical study with novices was conducted to compare the visualizations for accuracy, performance time, abandonment rate (rate for users intentionally choosing not to complete a task), usability measurements with 5 survey items using Likert Scale, mental workload using the NASA-TLX survey, and eye tracking measurements. The following eye tracking measurements were collected: time to first fixation, fixation length, first fixation duration, fixation count, and number of fixations before. In addition, demographic data was collected. Focus groups with the novices from the empirical study were used to gather qualitative data about the users' interaction with the visualizations. Lastly, a case study with domain experts was used to gather qualitative and quantitative interaction data (performance, usability, and workload) from persons who are knowledgeable of the medical domain. All this data was analyzed to look for patterns.

3.2.1 Heuristic Evaluation Methodology

The literature review on healthcare information systems, decision-making in healthcare, and data visualization identified a need for the development of a set of guidelines for determining the most appropriate visualization techniques for specific emergency healthcare decision-making situations. A heuristic evaluation was conducted on the prototype interface in GUI Design Studio.

Various methods can be utilized to evaluate the usability of graphical user interfaces. From a heuristic evaluation, or usability inspection, significant usability issues can be identified. This specific method is typically quick, inexpensive, and utilizes the expertise of usability subject matter experts. The total time to complete an evaluation is influenced by size and complexity of the interface, issues identified, and competency of the evaluators.

An assembly of evaluators completed an independent assessment of an interface based on a standard set of usability principles. There are several widely used set of usability principles and standards that can be potentially used in heuristic evaluations. When the assessment is finished, each evaluator has a list of problems identified. Each problem is associated with a specific usability heuristic(s) and estimated impact on usability, or the severity rating. This estimated impact assists with prioritizing the issues. The evaluation is usually done within the context of user tasks, and while more evaluators can identify more problems, there is an ideal range of evaluators that are dependent on a cost-payoff function. Figure 3.16 illustrates this function. The minimum should be about three and the maximum is about ten. Of course this is dependent on the benefits and risks associated with a particular system having usability issues. When it is a critical system, where a small usability problem can cause a huge loss, then all efforts should be made to identify 100% of the problems.

There are several widely used sets of usability heuristics. The most common are Nielsen and Bastien and Scapin (Bastien & Scapin, 1993; Nielsen, 1994). Nielsen's 10 heuristics are:

- Visibility of system status
- Match between system and the real world
- User control and freedom

- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

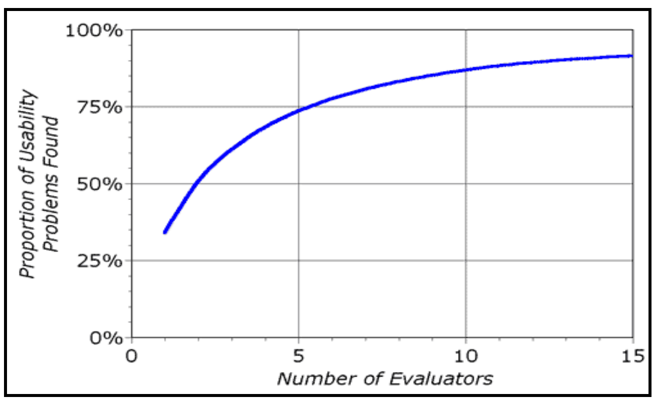


Figure 3.16: Diminishing Returns with Heuristic Evaluators (Nielsen, 1994)

Table 3.1 shows the Bastien and Scapin set of ergonomic criteria that contain 18 heuristics grouped into eight major areas. The latter set has been used in previous studies evaluating usability of visualizations (Pillat et al., 2005).

Purpose. The purpose of this heuristic evaluation of six different visualizations was to determine the usability issues with the use of these techniques when applied to emergency patient medical data.

Table 3.1: Bastien and Scapin's 18 Ergonomic Principles

Criteria No.	Group	Heuristic
1	Guidance	Prompting
2		Grouping / Distinction by location
3		Grouping / Distinction by format
4		Immediate feedback
5		Legibility
6	Workload	Concision
7		Minimal actions
8		Information density
9	Explicit Control	Explicit user action
10		User control
11	Adaptability	Flexibility
12		User Experience
13	Error Management	Error Protection
14		Quality of Error messages
15		Error Correction
16	Consistency	Consistency
17	Significance of Codes	Significance of Codes
18	Compatibility	Compatibility

Participants. The three evaluators were current doctoral candidates in Industrial and Systems Engineering at North Carolina A. & T. (N.C.A.T.) with a focus in human factors engineering. They had a high skill in technology and expertise in the area of usability engineering and testing. There were two females and one male. The evaluators were not users or developers.

Equipment. The visualization techniques were displayed on a Dell laptop using GUI Design Studio. The reports were done with paper and pen.

Stimulus Materials. The evaluators were given a list with descriptions of the Bastien and Scapin ergonomic principles, description of the severity ratings as seen in

Table 3.2, the GUI embedded with the six visualizations, and a user scenario associated with medical data.

Table 3.2: Severity Ratings by Nielsen (Nielsen, 1994)

Description	Rating
No usability problems at all	0
Cosmetic problem only: Need be fixed unless extra time is available	1
Minor usability problem: Fixing this should be given low priority	2
Major usability problem: Important to fix, so should be given high priority	3
Usability catastrophe: Imperative to fix this before product can be released	4

Procedure. For round one, each evaluator assessed all the visualizations independently and was reminded that they were to assess the visualization technique, not the software interface. Evaluators were allowed to study the heuristic list and ask any questions to clarify their understanding of criteria. The overall task for the evaluators was to record every usability issue they saw with the visualization. This round was conducted without discussion or interaction among group members. Each evaluator had unlimited time to complete the evaluation to accommodate for their unfamiliarity (if at all) with the heuristics list or the visualization techniques. For round 2, the evaluators came together for a group discussion and agreed on a final list of issues and severity ratings. Repetitious problems were combined or eliminated to create a comprehensive list. This procedure has been done in a prior evaluation study for information visualization (Pillat, Valiati, & Freitas, 2005). The data collected included a description of the issue, the associated heuristic violation, and severity rating. Additionally, the evaluators gave recommendations for the GUI's appearance in round 2. For round 3, the evaluators came together to assess the changes to the GUI after the recommendations from round 2.

Data Collection. Usability experts were given a blank heuristic report to record all their findings. The report is displayed in Figure 3.17. There is space for a description of the issue, the associated heuristic violation, and severity rating. The evaluators analyzed three components of an identified issue to determine its severity: the frequency (how often the problem occurs), persistency (whether or not the problem can be easily overcome in the future), and the impact of the problem (how much the problem undermines the goals of the user).

Evaluator's Name:		Session Date:		
Evaluator's Age:		Background:		
Evaluator's Gender:				
Visualization	**Heuristic Violated**	**Usability Defect Description**	**Inspectors Comments Regarding Defect**	**Severity Rating**

Figure 3.17: Blank Heuristic Report for Individual Evaluation

3.2.2 Pilot Study Methodology

Following the heuristic evaluation, a pilot study was conducted with a small number of novices to gather information before the larger empirical study. The main goals were to reveal deficiencies in the design of the experiment or procedures and determine time ranges for the empirical study. The following sub-section outlines the methodology.

Purpose. The purpose of this pilot study was to collect data about the usability of the data visualizations in a comparative analysis. The data collected was to determine appropriate ranges and experiment issues for the subsequent empirical study.

Participants. Participants were recruited via campus postings and classroom announcements. The subjects were students from North Carolina A&T State University, totaling five graduate engineering students. Five subjects, two females and three males, participated in the pilot study. Three of the subjects were between 18 and 25 years of age, one was between 26 and 36 years of age, and one was over 45 years of age. None were color blind and all were right handed. Four subjects rated their daily computer usage as high/frequently and rated their attitude to new technology as highly favorable.

Experimental Design. All subjects saw all six visualizations, making this a within subject design. The subjects saw the visualizations in a random order. The independent variable was the type of visualization. The dependent variables were decision time, accuracy, abandonment, and user opinion on usability.

Equipment. All subjects completed the experiment on the Tobii Eye Tracker X120, which has a standard monitor. The visualizations and accompanying questions were embedded into a GUI developed in Java**.**

Stimulus Materials. Stimulus materials include the healthcare decision-making task scenarios. To develop a task list, an analysis was done using studies prior information visualization evaluation, display comparisons for emergency department information systems, and the counsel of subject matter experts. The visualizations were created by SAS JMP, XDAT, and VOS Viewer.

Procedure. Subjects were introduced and welcomed to the study. The experimenter explained the study's purpose, benefits, and risks. The subjects gave their informed consent after asking any questions about the experiment. Next, the experimenter asked the subject to complete a demographic questionnaire, seen in Appendix B: Demographic Survey, and measured their color blindness using an Ishihara color blind test. Afterwards, the experimenter demonstrated the software and provided a training session on how to use the software. Subjects were allowed to ask any questions at this time about using the interface and completing the tasks. Subjects were informed that their tasks would be timed; therefore they must answer the task questions as quickly as possible. For the testing session, each visualization technique was shown to the subject accompanied with a question about what was shown in the image. Each question had four possible responses with the fourth response always “I Don’t Know”, allowing the subject to “abandon” the question if they wanted. After the subjects answered the task question, they completed a five question usability survey about the visualization. This process repeated for all six visualizations.

Data Collection. There were three categories of data collected from this study: objective metrics, subjective metrics, and demographic information. The objective metrics include time to complete task, percent of correct answers, and the abandonment rate (the rate a subject will select “I Don’t Know” as a response for a task). The time was measured with a standard stopwatch. The number of correct answers was measured by the predefined answer key for the questions. The usability of each visualization type was measured using a post-test questionnaire using a Likert scale from 1 to 5, including one item that asked the user to give the visualization an overall usability score. This survey can be seen in Appendix C: Usability Survey. The demographic information to be collected in the pre-test questionnaire included age, gender, field, color-blind, computer interaction and literacy, and attitude toward technology.

3.2.3 Empirical Study

An empirical study, consisting of two parts, was completed with novices to gain knowledge from direct observations. The main goals were to collect performance data, usability data, mental workload, and eye tracking metrics. This empirical study consisted of two parts. Part 1 compared 6 different visualizations (only one independent variable) and part 2 compared 2 different visualizations at 2 different difficulty levels (two independent variables).

3.2.3.1 Part 1. The following sub-section outlines the methodology for part 1 of the empirical study.

Purpose. The purpose of this empirical study was to collect data about the usability of the data visualizations in a comparative analysis. The following hypotheses are included for this study:

- Subjective mental workload will vary between the techniques
- Performance measures will vary between the techniques
- Eye tracking metrics will vary between the techniques

Participants. Fifty-nine subjects were recruited from North Carolina A&T State University and from nursing programs at local universities. Of the participants, 41 were 18-25 years of age, 12 were 26-35 years of age, and 6 were 36-45 years of age. There were 28 males and 31 females. Table 3.3 below summarizes the demographic statistics of the participants. Of the participants, 60% considered themselves more than average computer users and 60% were at least somewhat favorable towards new technology. In addition, 40% had worked in a hospital, and of those, 75% had experience using hospital information systems (either for scheduling or admission tasks).

Experimental Design. A within subject design was used. The subjects saw the visualizations in a randomized order. The independent variable was the type of visualization. The dependent variables were performance measurements (decision time, accuracy, and abandonment), eye tracking metrics, user opinion on usability, and mental workload.

Table 3.3: Demographic Analysis for Empirical Study (Part 1)

Discipline	Engineers	Nurses	
	73%	27%	
Ages	18-25 years	26-35 years	36-45 years
	70%	20%	10%
Gender	Male	Female	
	48%	52%	
Primary Hand	Right-Handed	Left-Handed	
	92%	8%	
Visual Acuity	20/20 or better	Less than 20/20	
	92%	8%	

Equipment. All subjects completed the experiment on the Tobii Eye Tracker X120, which has a standard monitor.

Stimulus Materials. Stimulus materials included the healthcare decision-making task scenarios. The visualizations were created by SAS JMP, XDAT, and VOS Viewer.

Procedure. Similar to the pilot study, subjects' consent was obtained after the experimenter welcomed and explained the study's purpose. The experiment then followed with the collection of demographic data, visual acuity, and color blindness. Next, the subject was walked through a demonstration and training on how to interpret the visualizations. Subjects could ask any questions at that time about using the interface and completing the tasks. Following the training session, the testing session began where the subjects viewed a visualization technique, read a scenario and answered a question about the data. Each question had four responses including "I Don't Know". After the subjects answered the task question, they completed a five question usability survey about the visualization and a six question mental workload survey, both seen in Appendix C and Appendix D. This process repeated for all 6 visualizations.

Similar approaches to this methodology for empirical studies evaluating information visualization have been used (Cawthon & Moere, 2007; Morse, Lewis & Olsen, 2000). The approach in this dissertation is different because it is analyzing decision time, decision accuracy, abandonment rate, usability, mental workload, and eye tracking metrics simultaneously.

Data Collection. There were four categories of data collected from this study: objective metrics, subjective metrics, eye tracking data, and demographic information. The demographics survey can be seen in Appendix B: Demographic Survey. The objective metrics included time to complete task, percent of correct answers, and percent of abandoned questions. The time was measured with a standard stopwatch. The number of correct answers was measured by a pre-defined answer key for the questions. The subjective metrics included user opinion about the usability of the visualizations and mental workload. The user opinion of each visualization technique's usability was measured using a post-test questionnaire seen in Appendix C: Usability Survey, using a Likert scale from 1 to 5. The mental workload associated with each visualization technique was measured using the NASA TLX survey seen in Appendix D: NASA-TLX Mental Workload Survey. The Tobii Eye Tracker X120 was used to collect data about user gaze patterns throughout the experiments. These metrics included first fixation duration, time to first fixation, and total fixation length. The demographic information collected in the pre-test questionnaire included age, gender, professional/academic field, degree level. The subjects' visual acuity was measured using a Snellen Eye Chart and color blindness was tested using the Ishihara test.

The National Aeronautic and Space Administration Task Load Index (NASA-TLX) is one of the most frequently used subjective mental workload scales and is based on multidimensional property of mental workload (Wu & Liu, 2006). It measures mental workload employing six rating scales: mental demand (MD), physical demand (PD), temporal demand (TD), performance (PE), effort (EF), and frustration (FR) levels. NASA-TLX has been successfully applied in a number of human factors system assessments (Yost and North, 2006; Fischer, Lowe, & Schwan, 2008).

3.2.3.2 Part 2. In part 1, task difficulty was not investigated. Therefore, a part 2 was deemed necessary to analyze the differences in task difficulty. In this part, there were two independent variables: visualization type (scatter plot matrix and parallel coordinates) and difficulty level (3 and 4 variables). For the context of this experiment, task complexity was determined by the number of variables visualized. This approach has been utilized in several information visualization evaluation studies (Svensson et al., 1997; Speier & Morris, 2003).

Purpose. The purpose of this empirical study was to collect data about the usability of two visualizations in a comparative analysis with two different task difficulty levels. The following hypotheses are included for this study:

- Subjective mental workload will vary between the techniques and the difficulty levels
- Performance measures will vary between the techniques and the difficulty levels

Participants. Ten subjects were recruited from North Carolina A&T State University and from nursing programs at local universities, six were psychology students and four were nurses. Of the participants, six were 18-25 years of age, two were 25-35 years of age, and two were 35-45 years of age. Three males and seven females participated. Table 3.4 summarizes the demographic statistics of the participants.

Table 3.4: Demographic Analysis for Empirical Study (Part 2)

<table>
<tr><td rowspan="2">Discipline</td><td>Psychology</td><td colspan="2">Nurses</td></tr>
<tr><td>40%</td><td colspan="2">60%</td></tr>
<tr><td rowspan="2">Ages</td><td>18-25 years</td><td>26-35 years</td><td>36-45 years</td></tr>
<tr><td>40%</td><td>30%</td><td>30%</td></tr>
<tr><td rowspan="2">Gender</td><td>Male</td><td colspan="2">Female</td></tr>
<tr><td>20%</td><td colspan="2">80%</td></tr>
<tr><td rowspan="2">Primary Hand</td><td>Right-Handed</td><td colspan="2">Left-Handed</td></tr>
<tr><td>80%</td><td colspan="2">20%</td></tr>
<tr><td rowspan="2">Visual Acuity</td><td>20/20 or better</td><td colspan="2">Less than 20/20</td></tr>
<tr><td>90%</td><td colspan="2">10%</td></tr>
</table>

Experimental Design. A 2x2 factorial design was used.

Equipment. All subjects completed the experiment on the Tobii X120, which has a standard monitor.

Stimulus Materials. Stimulus materials included the healthcare decision-making task scenarios and the visualizations created in SAS JMP and XDAT.

Procedure. The procedures for Part 2 are the same as the procedures described for Part 1 in Section 3.2.3.1. The subjects completed the tasks with two task difficulty levels for the scatter plot matrix and parallel coordinates. Afterwards, completing the usability and mental workload surveys for each task.

Data Collection. There were three categories of data collected from this study: objective metrics, subjective metrics, and demographic information. Details were described previously in Section 3.2.3.1.

3.2.4 Focus Groups Methodology

Focus groups were used in this research to gather qualitative data from the novices of the empirical study. The following outlines the methodology and the discussion format.

Purpose. The purpose of the focus groups was to collect qualitative data about the visualization techniques that was used to supplement the quantitative data collected in the empirical study. This approach has been used in prior studies evaluating medical data visualizations (Rester, 2007).

Participants. The participants were 14 nurses from the empirical study. They were split into two groups (a group of nine and a group of five). Eight of the participants were 25-35 years of age, two were 35-45 years of age, and four were over 45. One male and 13 females participated.

Equipment. The only equipment used for this study was a projector screen and an audio recorder.

Stimulus Materials. Stimulus materials included the visualization video replay and the questions from the moderator. The guidelines for the focus group followed the agenda in Table 3.5. The goal of the sessions was to gather qualitative data from the participants in regards to the following:

- Ease of use and usefulness of the technique for gaining insights
- Overall confidence in insights gained with the technique
- Major strength and weakness of the technique
- Similarity and difference of gained insights using different techniques
- Assumed comprehension rates of the complex matter with each technique

Table 3.5: Focus Group Agenda

Component	Topic	Time
Introduction	Focus group methodology	
	Recap of the visualization techniques	10 minutes
Overall Usability of the Visualizations	Can the technique be used intuitively?	
	What was the most severe problem?	
	What was the best feature of the visualization from a user perspective?	20 minutes
Visualization Features	Understandability	
	Structure or Position of Elements	20 minutes

Procedure. The guideline for the discussion began with an introduction that included the focus group methodology and a re-cap of the visualization techniques. This assisted in keeping the subjects focused and helped them recall the experiment. During that time, the subjects' consent was collected. Screenshots of the visualizations were played on a projector screen. Next, the discussion was directed at the overall usability of the visualizations than the strengths and weaknesses of each. Following this discussion, various components of each of the techniques were subject to discussion in regards to their understandability.

Data Collection. Qualitative free form data was collected from the focus group transcripts of both sessions.

3.2.5 Case Study Methodology

Following the empirical study, a case study was conducted using domain experts. The heuristic evaluation was performed with usability experts while the empirical study and focus groups were conducted with the same group of novices. A case study with domain experts is expected to give more complementary data. These types of studies can produce significant information about the assessment of a visualization technique because the domain experts are familiar with the realistic settings that the technique would be used.

Purpose. The purpose of the case study was to provide data from domain experts that will complement the data from the usability experts and novices, thus giving some validity to the findings in this dissertation.

Participants. The subjects included a neurology/neuroscience researcher, a bioinformatics researcher, and a nurse practitioner from a large hospital's triage unit. Two males and one female comprised the group. The demographic summary can be seen below in Table 3.6.

Table 3.6: Demographic Analysis for Case Study

Discipline	Medical Research	Nursing
	40%	60%
Ages	26-35 years	36-45 years
	33%	66%
Gender	Male	Female
	66%	33%
Primary Hand	Right-Handed	Left-Handed
	66%	33%
Visual Acuity	20/20 or better	Less than 20/20
	66%	33%

Experimental Design. This was a within-subject design. This is a validation study with field subjects that are experts in the domain but novices to the visualization techniques as applied to emergency medical data. The dependent variables included accuracy, performance time, abandonment rate, usability, and mental workload. The independent variable was visualization type.

Equipment. For this study, the experiment took place in the workplaces of the subjects or wherever they were comfortable (i.e. a conference meeting room, office, etc.) The visualizations were set up on a Dell Vostro laptop loaded with the Visual Basic GUI seen in Figure 3.15.

Stimulus Materials. Stimulus materials included the healthcare decision-making task scenarios, the visualizations (density chart, tree map, network diagram, 3D scatter plot matrix, and parallel coordinates), and a list of the heuristic principles by Bastien and Scapin for reference throughout the session.

Procedure. Table 3.7 shows the agenda for the case study. Subjects were greeted and introduced to the experimenter. They were then briefed on the research project, research question, and procedures. Their consent was obtained prior to beginning the case study. The experimenter then demonstrated for them the five visualization techniques (density chart, network diagram, tree map, scatter plot matrix, and parallel coordinates). Subjects could ask any questions at this time and start interacting with the interface. When the subjects were comfortable and ready, they started the usability test. The procedures for the testing session are similar to the procedures described in the empirical study seen in Section 3.2.3.1. After the last test was completed, the experimenter interviewed the participants to collect their opinion and feedback on the visualizations. Lastly, the interviewer debriefed and thanked subjects for their time.

Data Collection. The quantitative data collected included time, accuracy, abandonment, usability, and mental workload. The qualitative data collected included participants' feedback on the usability and utility of the visualizations.

3.3 Data Analysis

Both descriptive and inferential statistics were provided. Prior to the inferential statistical analysis (analysis of variance), an outlier analysis, correlation analysis, and simple descriptive statistics were done on the performance data. For the usability data and the mental workload analysis, simple descriptive statistics and a profile analysis were completed. An outlier analysis, simple descriptive statistics, a correlation analysis and multivariate analysis of variance were completed for the eye tracking measurements. These analyses support the guidelines that are created for the utilization of these visualization techniques. All data was checked for normality assumptions.

Table 3.7: Case Study Agenda

Component	Topic	Time
Introduction	Introducing interviewers, research project, research question, and interview procedure. Get permission to record interviews.	5 minutes
Training	Initial training (i.e. presentation of most basic features and interactions with software): Demonstrate each of the five visualization techniques.	10 minutes
Testing	Complete the usability test for each of the five techniques with the two task scenario questions Record the % correct, % abandoned, completion time	20 minutes
Interviewing	Interviewing of subjects on usability and utility of different visualizations • Was the general impression of the visualization positive or negative? (Give reasons) • Was the visualization understandable? (Give reasons) • Are there any concerns about shortcomings of the visualization? • What was the best feature? • How was the learnability? • Is the visualization technique suitable for the data of the subjects' domains or are improvements and/or modifications necessary?	25 minutes

3.4 Guidelines

Guidelines were developed based on the results of the above mentioned studies.

When developing the guidelines, the data analysis revealed some trends. It is not the goal to say a visualization technique is better than another is; rather the goal is to understand when and how a particular visualization may be more effective. A particular visualization may not always be the best in performance because some situations may warrant a further analysis. These guidelines clarify some situations for the developers so they may create useful, effective information systems that employ visualizations. In the overall framework, developing guidelines is the final step.

3.5 Conclusion

This chapter has presented the framework for evaluating an information visualization technique for emergency department medical data and summarized the planning phase of this research, which included a user requirements analysis done with an information flow analysis and activity theory. In addition, the test bed development method was composed of the data collection, visualization techniques, visualization software, and the GUI design. The test bed was developed carefully so that the results of the usability studies were accurate and relevant. The methodology for the heuristic evaluation, a pilot study, an empirical study, focus groups, and a case study was outlined in this chapter. Chapter 4 will discuss the results and findings from these experiments.

CHAPTER 4

RESULTS

The quantitative data gathered from the studies presented in chapter 3 was analyzed to compare the visualizations and their effects on user performance, perceived usability, and mental workload. These results are presented in this chapter in this order: heuristic evaluation, pilot study, empirical study, focus groups, and case study.

4.1 Heuristic Evaluation Results

The heuristic evaluation results of the three usability experts are summarized below for each of the three rounds. Round 1 was an independent round whereas round 2 and 3 were group rounds. The visualization techniques were assessed in both round 1 and round 2. The following visualization techniques were implemented in a graphical user interface: tree maps, network diagram, parallel coordinates, density chart, scatter plot matrix, and 3D scatter plot. The graphical user interface (GUI) was evaluated in all three rounds. Round 3 was to discuss the changes made to the GUI after round 2.

4.1.1 Results from Round 1

Overall, 34 problems were found using the Bastien and Scapin's 18 heuristic principles. The results can be seen in the following Table 4.1. The 3D scatter plot visualization had the most violations with 9. The scatter plot matrix and the density chart each had 3 violations, the least among all visualizations. Legibility had the most number of heuristic violations with 6. Several heuristic categories did not have any violations:

concision, explicit user action, and user control.

Table 4.1: Heuristic Evaluation Results of Round 1

Heuristic Criteria	3D Scatter Plot	Density Chart	Network Diagram	Parallel Coordinates	Scatter Plot Matrix	Tree Map
Prompting	1			1		2
Grouping / Distinction by Location			1		1	
Grouping / Distinction by Format		1				1
Immediate feedback	1					
Legibility	2		1	1	1	1
Concision						
Minimal Actions				1		1
Information Density	2		2	1		
Explicit User Action						
User Control						
Flexibility			1	1		1
User Experience						
Error Protection	1		1			
Error Messages						
Error Correction	1				1	
Consistency			1			1
Significance of Codes	1	1				
Compatibility		1				

4.1.2 Results from Round 2

In round 2, 30 problems were identified by the group as seen in Table 4.2. The 3D scatter plot had 7 violations, the most among all techniques. In terms of heuristics, significance of codes had the most violations with 6. The specific violations that had severity rating of 4 are seen in Table 4.3. The average severity rating score by visualization is as follows: 3.56 (standard deviation = 0.50) for 3D scatter plot, 3.50 (standard deviation = 0.58) for scatter plot matrix, 3 (standard deviation = 1.0) for density chart, 2.86 (standard deviation = 0.69) for network diagram, 2.43 (standard deviation = 0.49) for tree map, and 2.4 (standard deviation = 0.89) for parallel coordinates.

Table 4.2: Heuristic Evaluation Results of Round 2

Heuristic Criteria	3D Scatter Plot	Density Chart	Network Diagram	Parallel Coordinates	Scatter Plot Matrix	Tree Map
Prompting	1					
Grouping / Distinction by Location		1		1		
Grouping / Distinction by Format		2	1			2
Immediate Feedback	1					1
Legibility	1					
Concision						
Minimal Actions						
Information Density	1		1	1		1
Explicit User Action						
User Control						
Flexibility		1	1			
User Experience						
Error Protection	1		1			1
Error Messages						
Error Correction	1				1	
Consistency				1		
Significance of Codes	1	1	1	1	1	1
Compatibility				1		

Table 4.3: Severity Ratings for Heuristic Evaluation Results of Round 2

Visualization Type	Usability Defect Description	Inspectors Comments Regarding Defect	Severity Rating
3D Scatter Plot	Legibility	Titles should be centered, labels should be displayed in upper case letter; With multiple variables it is difficult to differentiate headings and values even with the different colors	4
3D Scatter Plot	Information Density	With multiple variables it is difficult to differentiate headings and values even with the different colors which heightens ambiguity for the user	4
Density chart	Significance of Codes	Assumes user understands color scheme	4
Density chart	Guidance	Font colors versus color scheme, font size relativity, shadow effects on particular boundaries, confusing locations of the words	4
Network Diagram	Error Protection	If there were a large number of nodes, the user would have issues detecting and preventing data entry errors, or actions. It would be difficult to protect field labels and input values with such a convoluted diagram.	4
Scatter Plot Matrix	Error Correction	If the variables are correlated together it could be very difficult to determine which headings and values are associated causing issues in error mitigation and correction practices	4
Tree Map	Guidance	Color shades are too similar, font sizes, divisions are not noticeable, cannot see smaller portions	4

4.1.3 Results from Round 3

After round 2, the following changes were made to the GUI. The font size was determined to be too small and thus was increased to size 11 to increase readability for the users. In addition, there were issues with clarity of the visualization images. This was fixed by saving the images as JPEGs and embedding them within the interface. As for error management, a back button was added to every page (excluding the Start page).

Details of the GUI designs were previously discussed in the Chapter 3.

4.2 Pilot Study Results

The descriptive statistics for time, accuracy, and abandonment rate for the pilot study are shown in Table 4.4. Among the visualization techniques, the 3D scatter plot and parallel coordinates had the longest average completion times. The visualizations with the lowest average completion times were density chart and network diagram. The density chart, tree map, and network diagram had the best average accuracy while parallel coordinates had the worst. The highest average abandonment rate was for scatter plot matrix and parallel coordinates. Density chart had the best overall usability score with 4.4 (standard deviation = 0.55) as shown in Figure 4.1. The lowest usability score was found by the 3D scatter plot at 2 (standard deviation = 1.22).

Table 4.4: Descriptive Statistics for Pilot Study

Visualization	Time (seconds)	Accuracy Rate	Abandonment Rate
Density Chart	86.72 (29.37)	1.00 (0.00)	0.00 (0.00)
Tree Map	116.10 (50.99)	1.00 (0.00)	0.00 (0.00)
Network Diagram	102.80 (43.78)	1.00 (0.00)	0.00 (0.00)
3D Scatter Plot	144.61 (36.40)	0.40 (0.55)	0.00 (0.00)
Scatter Plot Matrix	114.60 (17.61)	0.40 (0.55)	0.40 (0.55)
Parallel Coordinates	138.10 (58.56)	0.20 (0.45)	0.40 (0.55)

Before any large scale empirical study, a small pilot study is usually done to reveal problems before evaluation begins. Results from the pilot study revealed, the 3D scatter plot had the longest average completion time and lowest average usability score.

This is in agreement with the fact that it received several severity 4 ratings from the heuristic evaluation. The density chart had lowest average completion times and the best average accuracy. However, results from the heuristic evaluation also indicated that the usability issues identified by the evaluators were associated with the color scheme but none of the subjects were color blind. A high abandonment rate was found for the scatter plot matrix and this may give more insight to the heuristic violation noted by the evaluators for confusion with axis titles and headings. The parallel coordinates had the lowest accuracy and highest abandonment rate, yet it did not receive a severity rating of 4 by the heuristic evaluators. This may indicate a contradiction in regards to usability issues identified by evaluators and the performance of users with the visualization technique.

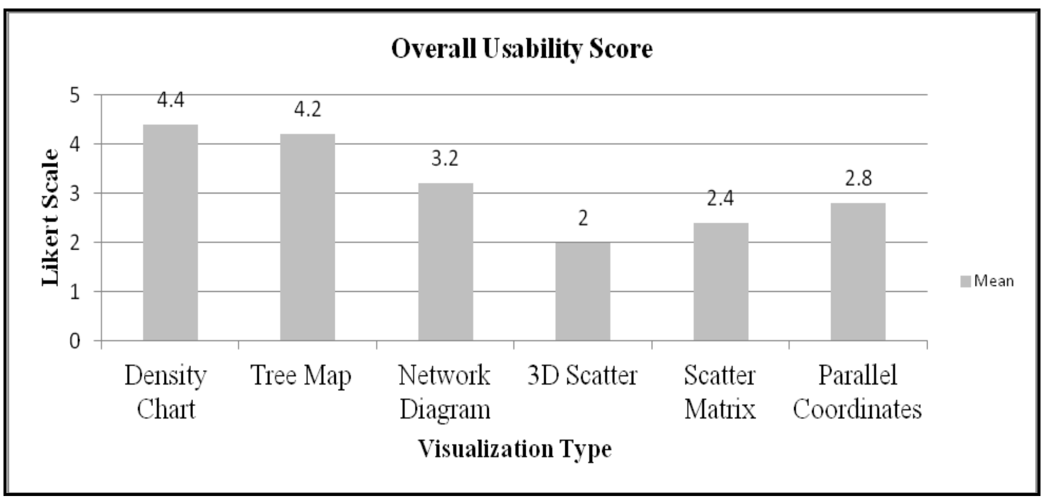


Figure 4.1: Mean Usability Scores from Pilot Study

The time ranges for the tree map, the network diagram, and the parallel coordinates are longer than the other visualizations. Due to this, the tasks were analyzed and redefined for the empirical study. The accuracy rate was 40% correct and 40% abandoned for scatter plot matrix. In addition the accuracy rate was 40% incorrect and 40% abandoned for parallel coordinates. These percentages indicate subjects were making guesses. To mitigate this issue in the empirical study, the visualizations were modified in attempt to resolve the subjects of any 'guessing'.

The findings of this pilot study are very limited by the sample size of the study and the functionality of the GUI. The sample size of the study was only 5 and it is difficult to draw definite conclusions from the small sample. It is unlikely that the pilot study alone will provide adequate data on variability in measurements. However, this data still has significance in refining and preparing for the empirical study. The pilot study also indicated that the GUI needed improvement before the empirical study.

The results of the pilot study were mainly used to develop procedures and identify issues for the empirical study and to determine expected time ranges for the subjects for the training phase and the testing phases. Using the results, the following changes were made: the order of the surveys (conduct usability survey before mental workload survey) and addition of the color blindness and visual acuity tests. According to the pilot study, the 3D scatter plot, scatter plot matrix, and parallel coordinates are the three visualizations that may have the most usability issues surfacing in the empirical study.

4.3 Empirical Study Results

An empirical study consisting of two parts was conducted. The following subsections present the results of both parts.

4.3.1 Part 1

This section presents the performance measurements, usability measurements, mental workload measurements, and eye tracking metrics from part 1 of the empirical study.

<u>4.3.1.1 Performance Measurements.</u> The performance measurements gathered include completion time, accuracy rate, and abandonment rate.

<u>Data Pre-Processing</u>. Box plots were constructed to identify outliers on completion time, which can be seen in Figure 4.2 There were 9 data points that were out of range across the 6 visualizations, which were removed.

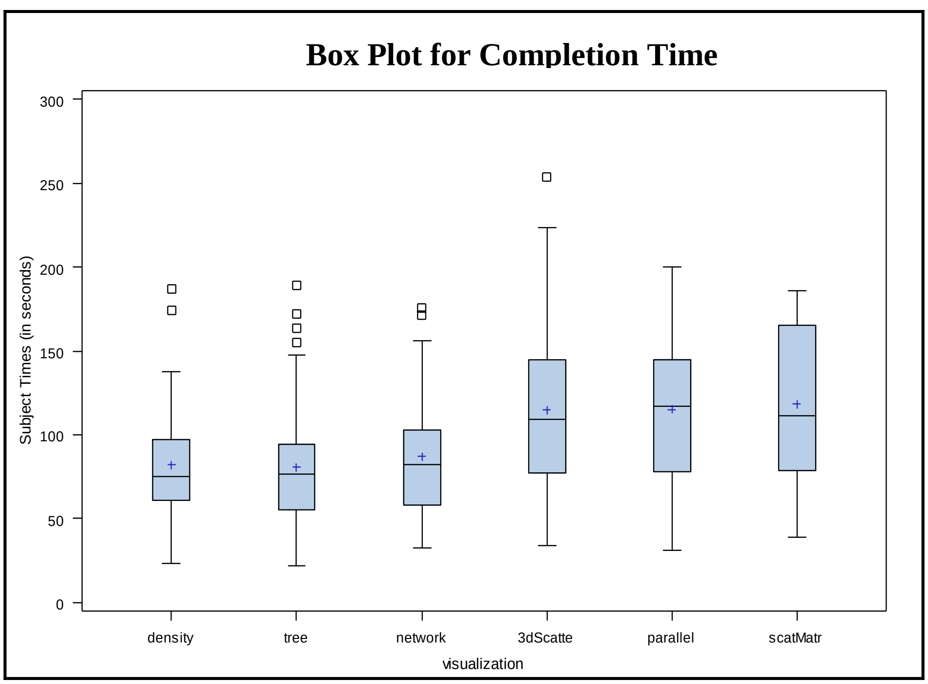


Figure 4.2: Outlier Analysis for Completion Time

Descriptive Statistics. The descriptive statistics for completion time, accuracy, and abandonment for all subjects can be seen in Table 4.5. The highest mean time to completion for 3D scatter plot was 118.27 (standard deviation = 45.63) seconds while the tree map had the shortest mean completion time of 80.57 (standard deviation = 35.69) seconds. For engineers, 3D scatter plot had the highest mean completion time and the network diagram had the shortest mean completion time. For nurses, parallel coordinates had the highest mean completion time while the tree map had the shortest mean completion time. The figure for comparing the completion time of nurses and engineers can be seen in Figure E.1 of the appendix. The network diagram had the highest accuracy rate at 79% while the parallel coordinates had the lowest with 21%. For engineers, density chart had the highest accuracy and parallel coordinates had the lowest accuracy. For nurses, tree map and network diagram had the highest accuracy while parallel coordinates had the lowest accuracy. Parallel coordinates also had the highest abandonment rate at 39%. The tree map visualization only had a 3% abandonment rate. For engineers, parallel coordinates had the highest abandonment rate. For nurses, density chart had highest abandonment rate.

Table 4.5: Descriptive Statistics for Performance from Empirical Study (Part 1)

Visualization	Time (seconds)	Accuracy Rate	Abandonment Rate
Density Chart	82.31 (30.04)	0.75 (0.39)	0.23 (0.36)
Tree Map	80.58 (35.69)	0.72 (0.44)	0.03 (0.18)
Network Diagram	86.96 (35.28)	0.79 (0.39)	0.09 (0.28)
3D Scatter Plot	114.90 (47.75)	0.49 (0.54)	0.24 (0.42)
Scatter Plot Matrix	118.27 (45.63)	0.46 (0.50)	0.37 (0.48)
Parallel Coordinates	115.21(44.55)	0.21 (0.39)	0.39 (0.48)

Inferential Statistics. In this study, there is one independent variable with 6 levels and three dependent variables. A moderate correlation was shown between accuracy, time, and abandonment as see in Table 4.6. Due to this moderate correlation a multiple analysis of variance (MANOVA) needs to be used to determine the variance between the performance variables for the 6 different visualizations.

Table 4.6: Correlation of Performance Metrics from Empirical Study (Part 1)

	Time	**Accuracy**	**Abandonment**
Time	1		
Accuracy	-0.462	1	
Abandonment	0.310	-0.630	1

A model adequacy check needs to be done to check the assumptions of normality, randomness, independence, and homogeneity of variance. Residual plots and a normality plot were created in SAS for time, which can be seen in Figure E.2 of the appendix. Visual inspection showed no major violations. Similar tests were run for accuracy and abandonment and none of those plots showed major violations.

The results of the MANOVA indicated a significant difference was present for visualization technique when all performance measurements were analyzed simultaneously (Wilks' Lambda = 0.6590, $F_{15, 930.71} = 10.12$, $p < 0.001$). In regards to each individual ANOVA for the performance measurements of time, accuracy, and abandonment there were significant differences present in the data. A significant difference was found for time ($F_{5,339} = 9.88$, $p<0.0001$). A significant difference was also found for accuracy rate and abandonment rate ($F_{5,339} = 19.96$, $p<0.0001$; $F_{5,339} = 7.34$,

$p<0.0001$).

Tukey post-hoc comparisons of the completion time for the 6 visualizations indicated that the tree map (mean = 80.58, standard deviation = 35.69) had significantly lower times than the network diagram (mean = 86.96, standard deviation = 35.28) and density chart (mean = 82.31, standard deviation = 30.04). The scatter plot matrix (mean = 118.27, standard deviation = 45.63) had significantly higher completion times than the network diagram and density chart. The parallel coordinates (mean = 115.21, standard deviation = 44.55) had significantly higher completion times than the network diagram and density chart. The 3D scatter plot (mean = 114.90, standard deviation = 47.75) had significantly higher completion times than the network diagram and density chart. The network diagram (mean = 86.96, standard deviation = 35.28) had significantly lower times than the scatter plot matrix, parallel coordinates, and 3D scatter plot. The density chart (mean = 82.31, standard deviation = 30.04) had significantly lower times than the tree map, scatter plot matrix, parallel coordinates, and 3D scatter plot. No comparisons of means for completion time were statistically significant at $p < .05$.

Tukey post-hoc comparisons of accuracy rate for the 6 visualizations indicated that the scatter plot matrix (mean = 0.46, standard deviation = 0.50) was significantly lower than the density chart (mean = 0.75, standard deviation = 0.39), the tree map (mean = 0.72, standard deviation = 0.44), and network diagram (mean = 0.79, standard deviation = 0.39). The 3D scatter plot (mean = 0.49, standard deviation = 0.54) had significantly lower accuracy than the density chart and network diagram. The density chart (mean = 0.75, standard deviation = 0.39) had significantly higher accuracy than the scatter plot matrix, parallel coordinates, and 3D scatter plot. The tree map (mean = 0.72, standard deviation = 0.44) had significantly higher accuracy than the scatter plot matrix, parallel coordinates, and 3D scatter plot. The network diagram (mean = 0.79, standard deviation = 0.39) had significantly higher accuracy rate than all other visualizations. The parallel coordinates (mean = 0.21, standard deviation = 0.39) had significantly lower accuracy than all other visualizations.

Tukey post-hoc comparisons of abandonment rate for the 6 visualizations indicate that the scatter plot matrix (mean = 0.37, standard deviation = 0.48) was significantly higher than network diagram (mean = 0.09, standard deviation = 0.28) and tree map (mean = 0.03, standard deviation = 0.18). The network diagram (mean = 0.09, standard deviation = 0.28) has significantly lower abandonment rate than scatter plot matrix (mean = 0.37, standard deviation = 0.48) and parallel coordinates (mean = 0.39, standard deviation = 0.48). The density chart (mean = 0.23, standard deviation = 0.36) had significantly lower abandonment rate than parallel coordinates (mean = 0.39, standard deviation = 0.48). The tree map (mean = 0.03, standard deviation = 0.18) had significantly lower abandonment rate than the scatter plot matrix and parallel coordinates. The parallel coordinates (mean = 0.39, standard deviation = 0.48) had significantly higher abandonment rate than tree map (mean = 0.03, standard deviation = 0.18).

4.3.1.2 Usability Measurements. The usability survey had 4 items measuring overall usability, ease of use, ease of viewing, and value as a healthcare tool. The following sections investigate the descriptive and inferential statistics associated with the usability items for the visualization techniques.

Descriptive Statistics. Table 4.7 shows the descriptive results for all usability items for empirical study (Part 1). For overall usability, the density chart scored the highest with 3.88 (standard deviation = 1.05) and 3D scatter plot scored the lowest with 2.32 (standard deviation = 1.11). Nurses rated the parallel coordinates with lowest usability while engineers rated the 3D Scatter Plot the lowest. Detailed descriptive statistics comparing nurses and engineers can be seen in Figures E.3-E.6. For the "ease of use" item, the density chart scored the highest with 3.57 (standard deviation = 1.35) and the 3D scatter plot scored the lowest with 1.98 (standard deviation = 1.05). Nurses favored the network diagram while engineers preferred the density chart. Density chart rated the lowest with nurses while engineers rated the 3D scatter plot the lowest. The tree map was rated the highest with 3.73 (standard deviation = 1.09) and the 3D scatter plot was rated the lowest with 1.98 (standard deviation = 1.05) for the "ease of viewing" item. Nurses favored the scatter plot matrix while engineers preferred the tree map for "ease of viewing". For "value as healthcare tool", the density chart was rated the highest with 2.29 (standard deviation = 0.74) and the 3D scatter plot was rated the lowest with 1.94 (standard deviation = 0.87). There is less variation in this item compared to the other usability items. Nurses favored the tree map while engineers favored the density chart. Parallel coordinates scored the lowest with nurses. 3D Scatter Plot scored the lowest with engineers.

Inferential Statistics. A correlation analysis, depicted in Table 4.8 revealed there is moderate correlation among all the usability survey items. Model adequacy tests revealed no major violations as seen in Figure E.7 of the appendix.

Since all four variables were measured at the same time in a survey, a profile analysis was conducted and three research questions were addressed: overall difference, parallelism, and flatness. An overall difference among the groups was found indicating there was a significant difference for visualization type ($F_{5, 278} = 16.26$, $p < 0.0001$). The test for parallelism examines whether the distance between scores for the techniques on any of the dependent variables differs. Results revealed that the profiles were not found to be parallel. Figure 4.3 shows a graphical depiction of the profiles (Wilks' lambda = 0.78464 and $F_{15, 762.32} = 4.67$, $p < 0.0001$). The last step is to check for the flatness of the profiles, which indicates that the visualizations show the same significance across the usability items. The results revealed that the profiles were not flat (Wilks' lambda = 0.645576 and $F_{3, 276} = 50.51$, $p < 0.0001$).

4.7: Descriptive Statistics for Usability Survey from Empirical Study (Part 1)

Usability Survey Item	Density Chart	Tree Map	Network Diagram	3D Scatter	Scatter Matrix	Parallel Coordinates
Ease of Use	3.57 (1.35)	3.2 (1.35)	2.84 (1.33)	1.98 (1.17)	2.33 (1.35)	2.29 (1.25)
Overall Usability	3.88 (1.05)	3.43 (1.27)	3.24 (1.28)	2.32 (1.11)	2.48 (1.36)	2.44 (1.32)
Ease of Viewing	3.33 (1.16)	3.73 (1.09)	2.82 (1.15)	1.98 (1.05)	2.63 (1.37)	2.31 (1.35)
Value for Healthcare	2.29 (0.74)	2.18 (0.81)	2.14 (0.76)	1.94 (0.87)	2.11 (0.80)	1.96 (0.90)

Table 4.8: Correlation of Usability Items from Empirical Study (Part 1)

	Ease of View	Ease of Use	Healthcare Value	Usability
Ease of View	1			
Ease of Use	0.415	1		
Healthcare Value	0.650	0.755	1	
Usability	0.480	0.812	0.780	1

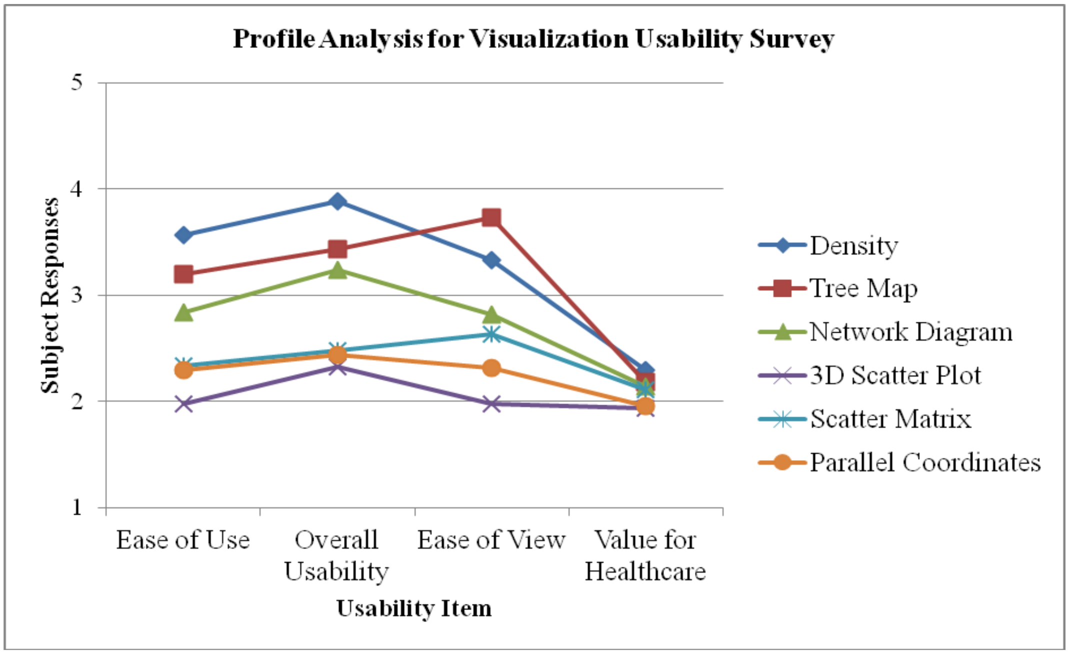


Figure 4.3: Profile Analysis Results for Usability Items

4.3.1.3 Mental Workload Measurements. The NASA TLX survey had 6 items measuring effort, frustration, performance, physical demand, mental demand, and temporal demand. The following sections investigate the descriptive and inferential statistics associated with the mental workload items for the visualization techniques.

Descriptive Statistics. The descriptive statistics are shown in Table 4.9. On the "effort to complete task" item, the parallel coordinates was rated the highest at 4.74 (standard deviation = 1.84) while the density chart rated the lowest with 3.21 (standard deviation = 1.57). For nurses, the parallel coordinates rated the highest with 6.75 and the density chart the lowest with 4.5. For engineers, the highest rated visualization was the 3D scatter plot with 4.68 and the lowest rated was the density chart with 3.12. All nurses to engineer comparisons are shown in Figures E.8-E.13. On the "frustration" item, the parallel coordinates was rated the highest at 3.72 (standard deviation = 2.17) while the tree map scored the lowest at 1.96 (standard deviation = 1.32). For nurses, the parallel coordinates was rated the highest with 6.5 and the tree map the lowest with 3.25. For engineers, the highest rated visualization was the 3D scatter plot with 3.53 and the lowest rated was the tree map with 1.84. For “mental demand”, the parallel coordinates was rated the highest at 4.7 (standard deviation = 2.49) while the density chart rated the lowest at 3.02 (standard deviation = 1.48). For nurses, the parallel coordinates rated highest with 6.75 and density chart the lowest with 3.5. For engineers, the highest rated visualization was 3D scatter plot with 4.28 and the lowest rated was the density chart with 2.98. On the “performance” item, the density chart rated the highest at 5.66 (standard deviation = 1.34) while the 3D scatter plot rated the lowest at 3.55 (standard deviation = 1.66). For nurses, the density chart rated highest with 6.5 while the scatter plot matrix and parallel coordinates were the lowest rated at 3.75. For engineers, the highest rated visualization was the density chart with 5.58 and the lowest rated was 3D scatter plot with 3.5. On the “physical demand and “temporal demand” items the trends are similar. The 3D scatter plot and parallel coordinates rated the highest while the density chart rated the lowest. For nurses, the parallel coordinates were the highest rated and density charts the lowest rated. For engineers, the highest rated visualization was 3D scatter plot and the lowest rated was the density chart.

Inferential Statistics. A correlation analysis, depicted in Table 4.10 revealed there was moderate correlation among all the mental workload survey items. Model adequacy tests revealed no major violations as seen in Figure E.14 of the appendix

Since all six variables were measured at the same time in a survey, a profile analysis was conducted and three research questions were addressed: overall difference, parallelism, and flatness. An overall difference among the groups was found indicating there was a significant difference for visualization ($F_{5, 256} = 5.64$, $p < 0.0001$). The test for parallelism examines whether the distance between scores for the techniques on any of the dependent variables differs. Results revealed that the profiles were not found to be parallel. Figure 4.4 shows a graphical depiction of the profiles (Wilks' lambda = 0.73365 and $F_{25, 937.64} = 3.26$, $p < 0.0001$). The last step is to check for the flatness of the profiles, which indicates that the visualizations show the same significance across the usability items. The result indicated that the profiles were not flat (Wilks' lambda = 0.42156 and $F_{5, 252} = 69.16$, $p < 0.0001$).

Table 4.9: Descriptive Statistics for NASA TLX from Empirical Study (Part 1)

Mental Workload Item	Density Chart	Tree Map	Network Diagram	3D Scatter	Scatter Matrix	Parallel Coordinates
Effort	3.21 (1.57)	3.77 (1.66)	3.4 (1.78)	4.80 (1.69)	4.7 (1.92)	4.74 (1.84)
Frustration	2.4 (1.3)	1.96 (1.32)	2.51 (1.65)	3.75 (1.93)	3.51 (2.04)	3.72 (2.17)
Mental	3.02 (1.48)	3.66 (1.62)	3.43 (1.81)	4.4 (2.07)	4.13 (2.44)	4.7 (2.49)
Performance	5.66 (1.34)	4.87 (1.76)	4.87 (1.73)	3.55 (1.66)	3.65 (2.19)	3.67 (2.04)
Physical	2.2 (1.68)	1.96 (1.350	2.36 (1.63)	3.42 (2.04)	3.12 (2.04)	3.12 (2.13)
Temporal	2.62 (1.57)	1.98 (1.26)	2.35 (1.46)	3.0 (1.7)	2.81 (1.92)	3.02 (2.01)

Table 4.10: Correlation of Mental Workload Survey Items

	Effort	Frustration	Mental	Performance	Physical	Temporal
Effort	1					
Frustration	0.889	1				
Mental	0.680	0.656	1			
Performance	0.749	0.419	-0.270	1		
Physical	0.615	-0.020	0.310	-0.780	1	
Temporal	0.192	0.450	0.765	-0.569	0.780	1

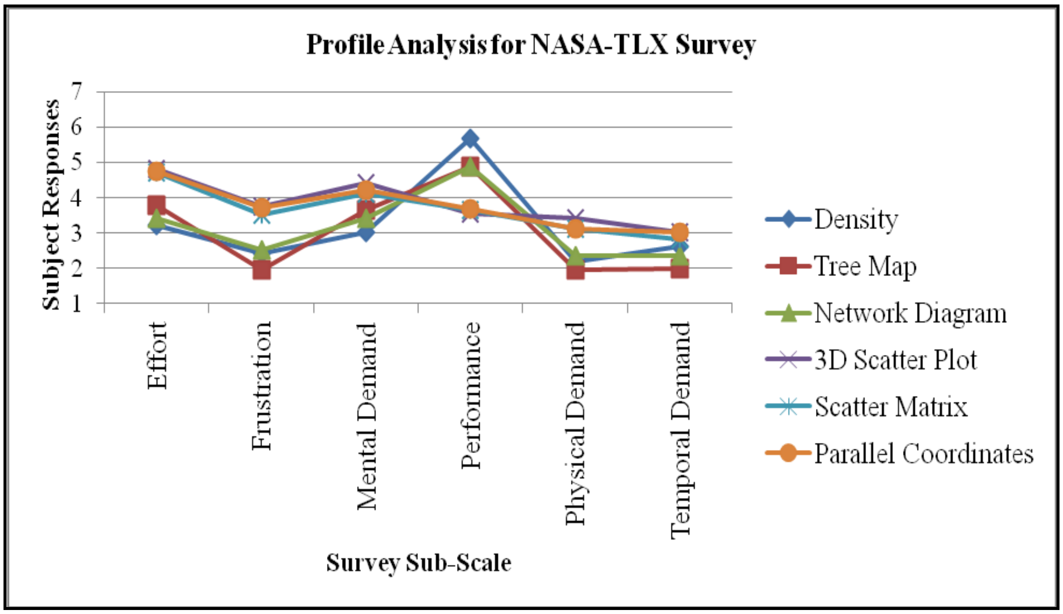


Figure 4.4: Profile Analysis of Mental Workload Items

4.3.1.4 Eye Tracking Results. The Tobii Eye Tracker was used to collect time to first fixation, first fixation duration, and total fixation duration. Using the Tobii Eye Tracker, 4 areas of interest for each visualization technique were measured: main, title, legend, and axis. The following sections provide the descriptive and inferential statistics associated with the eye tracking measurements for the visualization techniques.

Descriptive Statistics. The simple statistics for three eye tracking measurements can be seen below in Figures 4.5-4.7. The time to first fixation is displayed in Figure 4.5. Users took the longest amount of time before they looked at the axis for the 3D scatter plot (mean = 66.04, standard deviation = 3.67). For all visualizations, the shortest time to first fixation is for the main area of the visualization. In Figure 4.6, the first fixation duration is displayed, revealing that the axis of the 3D Scatter Plot received the longest mean time (mean = 0.45, standard deviation = 0.079) for duration of the first fixation. In Figure 4.7 depicting the total fixation duration, the 3D scatter plot, scatter plot matrix, and parallel coordinates have a large portion of their time associated with the axis.

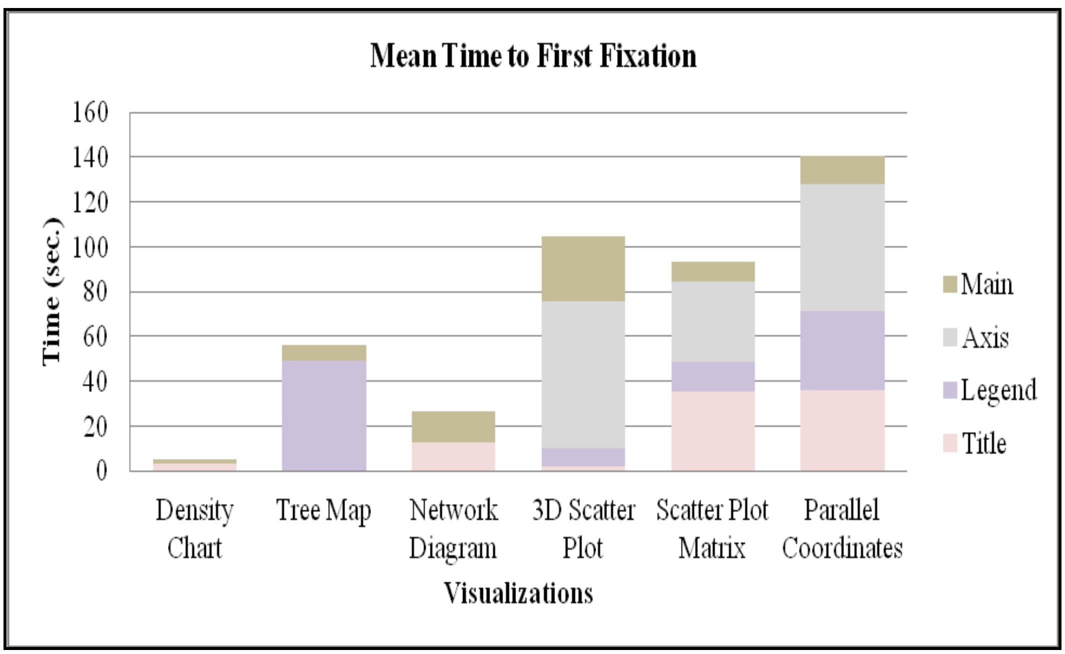


Figure 4.5: Mean Time to First Fixation

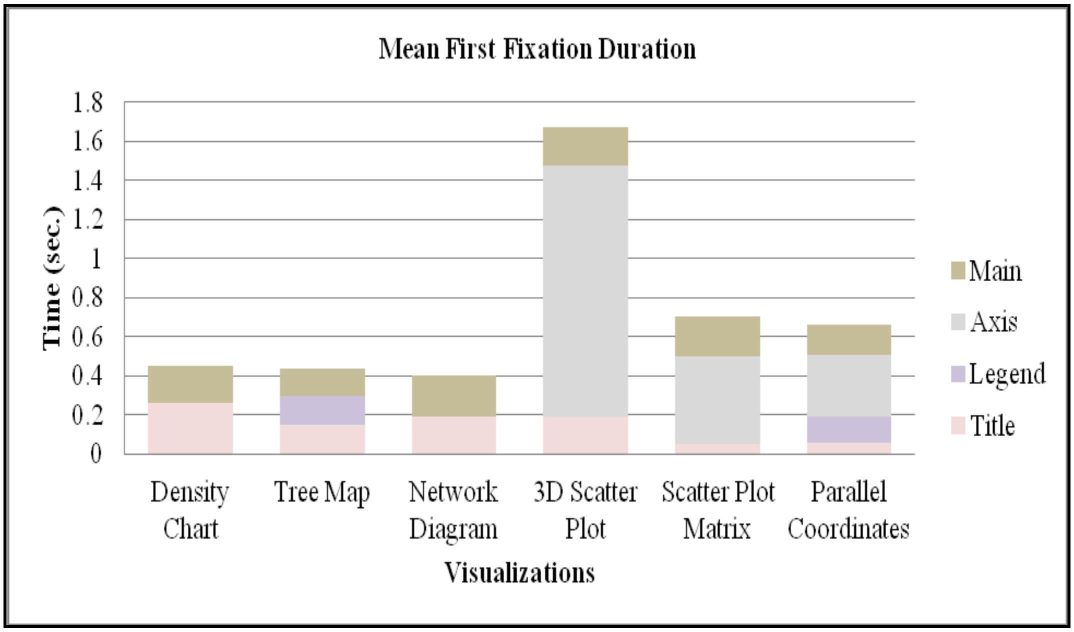


Figure 4.6: Mean First Fixation Duration

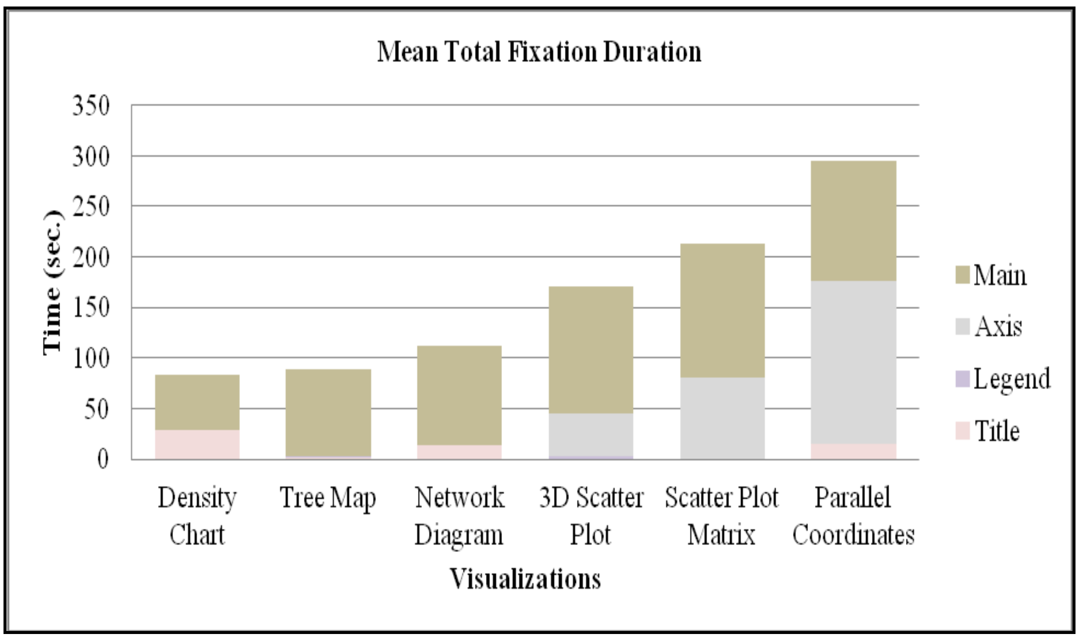


Figure 4.7: Mean Total Fixation Duration

Inferential Statistics. In this study, there is one independent variable with 6 levels and three dependent variables. A slight to moderate correlation was shown between accuracy, time, and abandonment as see in Table 4.11. Due to this moderate correlation a multiple analysis of variance (MANOVA) needs to be used to determine the variance between the eye tracking metrics for the 6 different visualizations.

Table 4.11: Correlation Matrix of Eye Tracking Metrics

	Time to First Fixation	**First Fixation Duration**	**Total Fixation Duration**
Time to First Fixation	1		
First Fixation Duration	0.012	1	
Total Fixation Duration	0.419	0.117	1

A model adequacy check was done with residual plots and a normality plot for time to first fixation, which can be seen in Figure E.15 of the appendix. Visual inspection showed no major violations. Similar tests were run for first fixation duration and total fixation duration and none of those plots showed major violations.

The results of the MANOVA indicated a significant difference was present for visualization technique when all three eye tracking metrics were analyzed simultaneously (Wilks' Lambda = 0.79119 and $F_{15,\ 955.56} = 5.54$, $p < 0.001$). In regards to each individual ANOVA for the eye tracking measurements, there were significant differences present in the data for time to first fixation and first fixation duration ($F_{5,344} = 14.84$, $p<0.0001$; $F_{5,344} = 2.33$, $p = 0.422$). There was no significant difference found for total fixation duration ($F_{5,344} = 0.53$, $p = 0.7512$).

Tukey post-hoc comparisons of means for total fixation duration were not statistically significant at $p < 0.05$. However, a comparison of means for time to first fixation for the 6 visualizations indicate that the 3D scatter plot (mean = 29.31, standard deviation = 39.6) was significantly higher than all other visualizations. The network diagram (mean = 13.89, standard deviation = 24.55) had significantly higher times than tree map (mean = 1.24, standard deviation = 1.84) and density chart (mean = 0.169, standard deviation = 0.22). The scatter plot matrix (mean = 8.6, standard deviation = 16.6) had significantly higher times than tree map (mean = 1.24, standard deviation = 1.84) and density chart (mean = 0.169, standard deviation = 0.22). The parallel coordinates (mean = 5.92, standard deviation = 17.81) had significantly higher times than tree map and density chart. The tree map (mean = 1.24, standard deviation = 1.84) had significantly lower times than 3D scatter plot (mean = 29.31, standard deviation = 39.6), network diagram (mean = 13.89, standard deviation = 24.55), scatter plot matrix (mean = 8.6, standard deviation = 16.6), and parallel coordinates (mean = 5.92, standard deviation = 17.81). The density chart (mean = 0.169, standard deviation = 0.22) had significantly lower times than 3D scatter plot (mean = 29.31, standard deviation = 39.6), network diagram (mean = 13.89, standard deviation = 24.55), scatter plot matrix (mean = 8.6, standard deviation = 16.6), and parallel coordinates (mean = 5.92, standard deviation = 17.81).

4.3.2 Part 2

This section presents the performance measurements, usability measurements, mental and workload measurements for Part 2 of the empirical study.

4.3.2.1 Performance Measurements. The performance measurements gathered include completion time, accuracy rate, and abandonment rate.

Data Pre-Processing. Box plots were constructed to identify outliers on completion time, which can be seen in Figure 4.8. No outliers were identified.

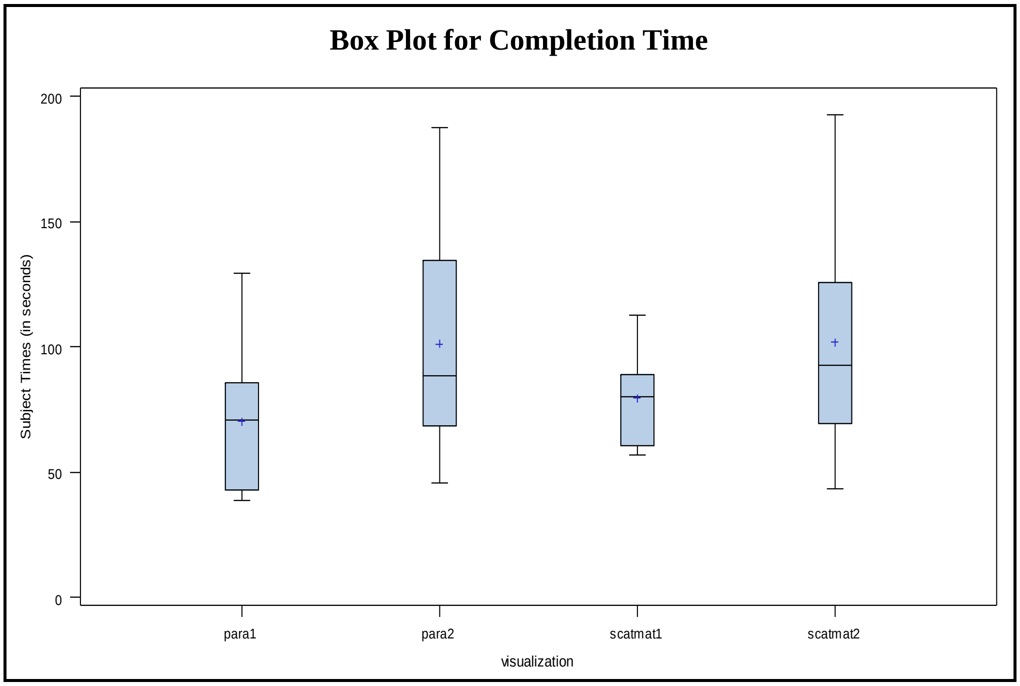


Figure 4.8: Outlier Analysis for Completion Time

Descriptive Statistics. The descriptive statistics for completion time, accuracy, and abandonment rate can be seen in Table 4.12. This was a 2x2 factorial design. There were 2 visualization types, scatter plot matrix and parallel coordinates. There were two levels of difficulty, level 1 is 3 variables and level 2 is 4 variables. The highest mean time to completion was for the scatter plot matrix, level 2 at 101.8 (standard deviation = 45.32) seconds. The shortest mean time to completion was for the scatter plot matrix, level 1 at 84.66 (standard deviation = 17.74) seconds. Scatter plot matrix at level 1 difficulty had the highest accuracy at 0.90 (standard deviation = 1) seen in Table 4.12. Parallel coordinates had the highest abandonment rate at 0.20 (standard deviation = 1), for both difficulty levels.

Table 4.12: Descriptive Statistics for Performance from Empirical Study (Part 2)

Visualization		Time (seconds)	Accuracy Rate	Abandonment Rate
Level 1	Parallel Coordinates	63.88 (22.75)	0.7 0 (0.48)	0.20 (0.32)
	Scatter Plot Matrix	84.66 (17.74)	0.9 0 (0.32)	0.00 (0.00)
Level 2	Parallel Coordinates	101.20 (46.55)	0.7 0 (0.48)	0.20 (0.32)
	Scatter Plot Matrix	101.80 (45.32)	0.8 0 (0.42)	0.00 (0.00)

Inferential Statistics. In this study, there are two independent variables with 2 levels and three dependent variables. Little to no correlation was found between time, accuracy, and abandonment rate as seen in Table 4.13. A two-way analysis of variance (ANOVA) was used to determine the impact on user performance.

Table 4.13: Correlation of Performance Metrics from Empirical Study of (Part 2)

	Time	Accuracy	Abandonment
Time	1		
Accuracy	0.166	1	
Abandonment	0.035	0.006	1

A model adequacy check was done with residual plots and a normality plot for completion time, which can be seen in Appendix F. Visual inspection showed no major violations. Similar tests were run for accuracy rate and abandonment rate and none of those plots showed major violations.

Two-way ANOVA was conducted to examine the effect of visualization and difficulty level on time, accuracy rate, and abandonment rate respectively. No significant interaction effect between visualization and difficulty level was found on time ($F_{1,36}$ = 0.13, p = 0.7172), accuracy ($F_{1,36}$ = 0.13, p = 0.7161), or abandonment ($F_{1,36}$ = 2.06, p = 0.1600). A significant task difficulty main effect was found on time ($F_{1,36}$ = 5.18, p = 0.0289), but not on accuracy ($F_{1,36}$ = 0.13, p = 0.7161) or abandonment ($F_{1,36}$ = 0.00, p = 1.00). No significant difficulty level main effect was found on time ($F_{1,36}$ = 0.17, p = 0.6758), accuracy ($F_{1,36}$ = 1.21, p = 0.2788), or abandonment ($F_{1,36}$ = 2.06, p = 0.1600).

4.3.2.2 Usability Measurements. The usability survey had 4 items measuring overall usability, ease of use, ease of viewing, and value as a healthcare tool. The following sections investigate the descriptive and inferential statistics associated with the usability items for the visualization techniques.

Descriptive Statistics. Table 4.14 shows the descriptive results for the usability survey items. For the "overall usability" item, the scatter matrix at level 1 rated the highest with 4.4 (standard deviation = 0.82) and the parallel coordinates level 2 rated the lowest with 2.30 (standard deviation = 1.16). On the "ease of use" item, the scatter matrix at level 1 rated the highest with 4.3 (standard deviation = 0.82) and parallel coordinates level 2 rated the lowest with 2.20 (standard deviation = 0.92). For the "ease of viewing" item, the scatter plot matrix at level 2 difficulty was rated the highest with 4.3 (standard deviation = 0.82), while the parallel coordinates at level 2 difficulty had the lowest with 2.2 (standard deviation = 0.92). Overall, the participants did not have a large range for the scores on the "value as healthcare tool" usability item. Scatter plot matrix at level 2 and parallel coordinates at level 2 both rated 2.8 (standard deviation = 0.42) while the lowest rating was for parallel coordinates at level 1 with 2.6 (standard deviation = 0.42).

Table 4.14: Descriptive Statistics for Usability Survey from Empirical Study (Part 2)

Visualization		Ease of Use	Overall Usability	Ease of Viewing	Value for Healthcare
Level 1	Scatter Matrix	4.3 (0.82)	4.4 (0.84)	3.3 (1.16)	2.7 (0.67)
	Parallel Coordinates	3.1 (1.45)	3.1 (1.52)	3 (1.15)	2.6 (0.42)
Level 2	Scatter Matrix	3.9 (0.99)	3.8 (1.23)	4.3 (0.82)	2.8 (0.70)
	Parallel Coordinates	2.2 (0.92)	2.3 (1.16)	2.2 (0.92)	2.8 (0.42)

Inferential Statistics. A correlation analysis, depicted in Table 4.15 revealed there was little to moderate correlation among all the usability survey items. A model adequacy check was done with residual plots and a normality plot for all usability items. Visual inspection showed no major violations.

Table 4.15: Correlation of Usability Items from Empirical Study (Part 2)

	Ease of Use	Usability	Ease of view	Value
Ease of Use	1			
Usability	0.153	1		
Ease of view	0.358	0.075	1	
Value	0.917	0.014	0.438	1

Similarly, two-way ANOVA was conducted to examine the effect of visualization and difficulty level on usability survey items. Results revealed no interaction effect between visualization and difficulty level ($F_{1,36}$ = 0.26, p = 0.7692), visualization main effect ($F_{1,36}$ = 0.11, p = 0.7469), or difficulty level main effect ($F_{1,36}$ = 0.42, p = 0.5195) on "ease of use". However, there was a significant interaction effect between visualization and difficulty level on "overall usability" ($F_{1,36}$ = 9.48, p = 0.0005). When difficulty level was fixed at level 1, a significant visualization simple main effect was found ($F_{1,36}$ = 8.39, p = 0.0064). When difficulty level was fixed at level 2, a significant visualization simple main effect was found ($F_{1,36}$ = 9.84, p = 0.0034). However, no significant difficulty level simple main effect was found for parallel coordinates ($F_{1,36}$ = 0.23, p = 0.6322) or scatter plot matrix ($F_{1,36}$ = 0.06, p = 0.8106). Similarly, there was a significant interaction effect between visualization and difficulty level on "ease of viewing" ($F_{1,36}$ = 0.70, p = 0.4080). When difficulty level was fixed at level 1, a significant visualization simple main effect was found ($F_{1,36}$ = 14.36, p = 0.0006). When difficulty level was fixed at level 2, a significant visualization simple main effect was found ($F_{1,36}$ = 6.79, p = 0.0133). However, no significant difficulty level simple main effect was found for parallel coordinates ($F_{1,36}$ = 0.22, p = 0.6386) or scatter plot matrix ($F_{1,36}$ = 0.50, p = 0.4820). The same trend was also found for "healthcare value". There was an interaction effect for visualization and difficulty level on value for healthcare ($F_{1,36}$ = 6.86, p = 0.0029). When difficulty level was fixed at level 1, a significant visualization simple main effect was found ($F_{1,36}$ = 7.80, p = 0.0083). When difficulty level was fixed at level 2, a significant visualization simple main effect was found ($F_{1,36}$ =

5.58, p = 0.0236). However, no significant difficulty level simple main effect was found for parallel coordinates ($F_{1,36}$ = 0.18, p = 0.6700) or scatter plot matrix ($F_{1,36}$ = 0.00, p = 1.000).

<u>4.3.2.3 Mental Workload Measurements</u>. The NASA TLX survey had 6 items measuring effort, frustration, performance, physical demand, mental demand, and temporal demand. The following sections investigate the descriptive and inferential statistics associated with the mental workload items for the visualization techniques.

<u>Descriptive Statistics</u>. The descriptive statistics for the NASA-TLX survey items are shown in Table 4.16. For "effort to complete task", parallel coordinates at level 2 difficulty was rated the highest at 4.4 (standard deviation = 1.65) while the scatter plot matrix at level 1 rated the lowest with 2.5 (standard deviation = 1.51). A similar trend is seen in mental demand, temporal demand, and physical demand. For the item "frustration with task", parallel coordinates at level 2 was rated the highest score again with 3.0 (standard deviation = 1.83) but the scatter plot matrix at level 2 was rated the lowest at 1.3 (standard deviation = 0.67). For "performance" scatter plot matrix at level 1 was rated highest with a value of 6 (standard deviation = 1.25) while the parallel coordinates at level 2 were rated the lowest with 3.7 (standard deviation = 1.16).

<u>Inferential Statistics</u>. A correlation analysis, depicted in Table 4.17 revealed there are several little to moderate correlations among all the mental workload survey items. Normality plots of the residuals were used to check the assumption of normality for the workload survey items. There were only minor violations with the data.

Despite the correlations, due to the small sample size, a univariate ANOVA was adequate. Two-way ANOVA was also conducted to examine the effect of visualization and difficulty level on mental workload survey items. There was no interaction effect between visualization and difficulty level on effort ($F_{1,36}$ = 3.02, p = 0.0611), frustration ($F_{1,36}$ = 2.84, p = 0.0711), performance ($F_{1,36}$ = 1.27, p = 0.2936), or physical demand ($F_{1,36}$ = 2.98, p = 0.0632). There was a significant visualization main effect on effort ($F_{1,36}$ = 4.71, p = 0.0364) and physical demand ($F_{1,36}$ = 5.19, p = 0.0286). No significant visualization main effect was found on frustration ($F_{1,36}$ = 0.57, p = 0.4557), or performance ($F_{1,36}$ = 2.28, p = 0.1395). A significant difficulty level main effect was found on frustration ($F_{1,36}$ = 5.11, p = 0.0297), but no significant difficulty level main effect was found on effort ($F_{1,36}$ = 1.32, p = 0.2578), performance ($F_{1,36}$ = 0.25, p = 0.6177), or physical demand ($F_{1,36}$ = 0.77, p = 0.3866). A significant interaction effect between visualization and difficulty level was found on mental demand ($F_{1,36}$ = 5.43, p = 0.0086). No significant visualization simple main effect was found for difficulty level 1 ($F_{1,36}$ = 0, p = 1.000) or level 2 ($F_{1,36}$ = 2.83, p = 0.1012). No significant difficulty level simple main effect was found for scatter plot matrix ($F_{1,36}$ = 1.81 p = 0.1868). However, a significant difficulty level simple main effect was found for parallel coordinates ($F_{1,36}$ = 9.17, p = 0.0045).A significant interaction effect between visualization and difficulty level on temporal demand was found ($F_{1,36}$ = 7.25, p = 0.0022). A significant visualization simple main effect was found when difficulty level is fixed at level 1 ($F_{1,36}$ = 28.62, $p<0.0001$), but no significant visualization simple main effect was found when difficulty level is fixed at level 2 ($F_{1,36}$ = 1.01, p = 0.3225). A significant difficulty level simple main effect was found for parallel coordinates ($F_{1,36}$ = 32.31, $p<0.0001$), but no significant difficulty level simple main effect was found for the scatter plot matrix ($F_{1,36} = 0.45$ $p = 0.5079$).

Table 4.16: Descriptive Statistics for NASA-TLX from Empirical Study (Part 2)

	Level 1 Difficulty		Level 2 Difficulty	
Item	**Scatter Matrix**	**Parallel Coordinates**	**Scatter Matrix**	**Parallel Coordinates**
Effort	2.5 (1.51)	3.0 (1.94)	3.5 (1.27)	4.4 (1.65)
Frustration	2.1 (1.29)	2.4 (1.96)	1.3 (0.67)	3.0 (1.83)
Mental	2.3 (0.95)	3.1 (1.85)	3.4 (1.17)	4.4 (1.17)
Performance	6.0 (1.25)	4.8 (1.81)	5.2 (1.32)	3.7 (1.16)
Physical	1.78 (1.09)	1.89 (0.93)	1.67 (0.87)	2.89 (1.36)
Temporal	2.44 (1.42)	2.22 (0.71)	1.67 (1.30)	2.67 (1.41)

Table 4.17: Correlation of Mental Workload Survey Items

	Effort	**Frustration**	**Performance**	**Physical**	**Mental**	**Temporal**
Effort	1					
Frustration	0.316	1				
Performance	0.588	0.342	1			
Physical	0.047	0.142	0.287	1		
Mental	0.423	0.474	0.329	-0.133	1	
Temporal	0.176	0.126	0.036	0.108	0.400	1

4.4 Focus Groups Results

From the focus group sessions, the audio recording was transcribed and analyzed. The detailed results of this analysis can be seen in Appendix H: Focus Group Transcript Summary. Each session's discussion was broken down to comments (repetitious or similar comments were combined) and scanned for trends. From these comment lists, popular comments were categorized into four areas, each area representing a "theme" in the comments. The four themes seen in Figure 4.9 are visual comprehension, compatibility, technical support, interaction. Visual comprehension was used to describe comments relative to the visualization allowing the user to see a quick glimpse of multiple records. Compatibility was used to describe comments relative to colors, font sizes, etc. where the user's mental model was met. During the sessions, a participant noted that they knew that red meant a high frequency on the density chart. Technical support comments were associated with the location and structure of axes, legends, labels, titles. The fourth theme, interaction, was created for comments regarding a user manipulation like highlighting, zoom in or out, etc. A fifth area named "Other" was created for some comments that did not fit into the above mentioned themes. There were five positive comments, seven were neutral, and eight were negative. It is interesting to note that during some of the sessions, a particular aspect of visualization would be discussed as both a positive and a negative. The box sizes on the tree map made that visualization seem good for some situations but poor for other. This was the case with other aspects of the visualizations. Overall, more than 70% of the persons in the groups expressed frustration with using the visualizations. However, more than 70% of the persons agreed that these visualizations could be used to assist in emergency medical care. The frustration is evident, but there is still favorability for use of these visualizations for emergency medical data.

4.5 Case Study Results

This sub-section will explain the results of the usability testing with 3 domain experts. This is a small sample and thus has limitations that will be discussed further in Chapter 6.

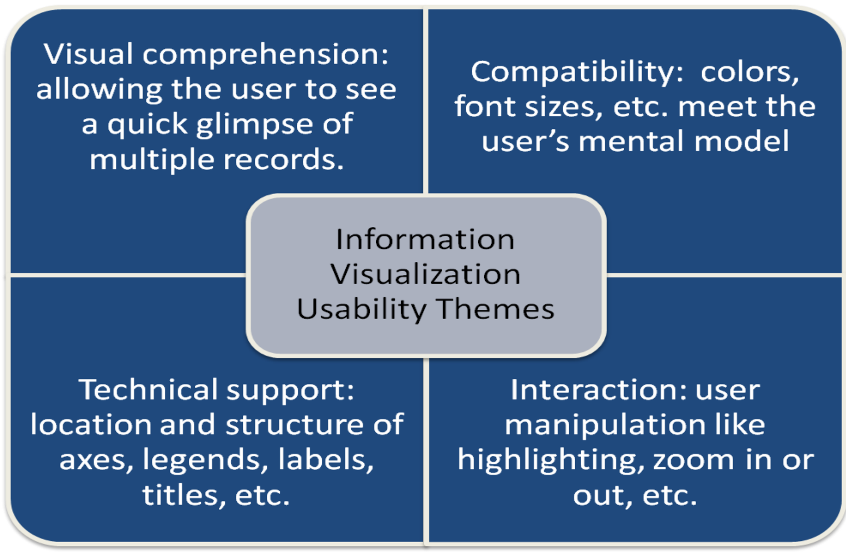


Figure 4.9: Themes from Focus Groups

4.5.1 Performance Measurements

Descriptive Statistics. The descriptive statistics for mean completion time are shown in Table 4.18. Following the trend from the empirical study results, the density chart has the shortest performance time with a mean of 71 seconds. Scatter plot matrix has the highest mean completion time at 98.8 seconds. Density chart, scatter plot matrix, and parallel coordinates had 100% accuracy as depicted in Table 4.18. Tree map and network diagram had 66%. There were no tasks abandoned. Since no tasks were abandoned, no further analysis was done for abandonment rate.

Inferential Statistics. Residual analysis charts and a normality plot were created in

SAS for time, which can be seen in appendix H. Visual inspection showed no major violations and an ANOVA was performed. In regards to completion time, there was no significant difference found ($F_{4,10}$ = 0.91, p = 0.4955). For accuracy, there was no significant difference found ($F_{4,10}$ = 0.75, p = 0.5801).

Table 4.18: Descriptive Statistics for Performance from Case Study

Visualization	Time (seconds)	Accuracy Rate	Abandonment Rate
Density Chart	71.0 (23.12)	1.00 (0.00)	0 (0.00)
Tree Map	78.5 (24.70)	0.66 (0.58)	0 (0.00)
Network Diagram	88.4 (23.15)	0.66 (0.58)	0 (0.00)
Scatter Plot Matrix	98.8 (15.99)	1.00 (0.00)	0 (0.00)
Parallel Coordinates	96.9 (20.26)	1.00 (0.00)	0 (0.00)

4.5.2 Usability Measurements

The usability survey had 4 items measuring overall usability, ease of use, ease of viewing, and value as a healthcare tool. The following sections describe the descriptive and inferential statistics of the case study results associated with the usability items for the visualization techniques.

Descriptive Statistics. The descriptive statistics for the usability survey items are show in Table 4.19. Density chart had the highest rating for “ease of use”, with 4 (standard deviation = 1.0) while the parallel coordinates and scatter plot matrix had the lowest ratings at 2.67(standard deviation = 0.58). The same trend applied to “overall usability”. For “ease of viewing”, the tree map had the highest rating with 4.33 (standard deviation = 1.15) while both parallel coordinates and scatter plot matrix had the lowest ratings of 2.67 (standard deviation = 0.58). The network diagram had the highest rating with 3.67 (standard deviation = 1.53) for "value as healthcare tool" item while parallel coordinates had the lowest at 2.07 (standard deviation = 2.07).

Table 4.19: Descriptive Statistics for Usability Survey from Case Study

Usability Survey Item	Density Chart	Tree Map	Network Diagram	Scatter Matrix	Parallel Coordinates
Ease of Use	4.00 (1.00)	3.67 (0.58)	3.67 (1.15)	2.67 (0.58)	2.67 (0.58)
Overall Usability	4.33 (1.15)	4.00 (1.0)	3.67 (1.15)	2.67 (1.53)	2.67 (1.53)
Ease of Viewing	3.67 (1.15)	4.33 (1.15)	3.67 (1.15)	2.67 (0.58)	2.67 (0.58)
Value for Healthcare	3.33 (1.53)	2.4 (1.04)	3.67 (1.53)	2.73 (1.42)	2.07 (0.90)

Inferential Statistics. Model adequacy tests with normality and residuals plots revealed no major violations for the usability survey items. For the analysis of variance, no items from the survey showed a significant difference. There was no significant difference for "ease of use" ($F_{4,10}$ = 1.75, p = 0.2154), "overall usability" ($F_{4,10}$ = 1.06, p = 0.425), "ease of viewing" ($F_{4,10}$ = 1.68, p = 0.2306), or "value for healthcare" ($F_{4,10}$ = 0.77, p = 0.5699).

4.5.3 Mental Workload Measurements

The NASA TLX survey had 6 items measuring effort, frustration, performance, physical demand, mental demand, and temporal demand. The following sections investigate the descriptive and inferential statistics from the case study data associated with the mental workload items for the visualization techniques.

Descriptive Statistics. Table 4.20 shows the descriptive statistics for the NASA-TLX survey items. For "effort to complete task", the scatter plot matrix rated the highest with 5.33 and the density chart rated the lowest with 1.67. A similar trend is seen in "frustration" and "mental demand". Density chart rated the highest for "performance" with 6.33 while parallel coordinates rated the lowest with 4.33. For "physical demand" and "temporal demand" the scatter plot, network diagram, and parallel coordinates rated the lowest with 1.33 while density chart and tree map rated the highest at 1.67.

Table 4.20: Descriptive Statistics for NASA-TLX from Case Study

Mental Workload Item	Density Chart	Tree Map	Network Diagram	Scatter Matrix	Parallel Coordinates
Effort	1.67 (1.15)	4.67 (1.15)	2.67 (1.53)	5.33 (0.58)	3.67 (1.15)
Frustration	2.33 (0.58)	4.00 (1.0)	3.00 (1.0)	4.33 (0.58)	4.33 (1.53)
Mental	2.00 (1.0)	4.67 (1.53)	2.67 (1.53)	4.67 (1.15)	4.00 (1.0)
Performance	6.33 (0.58)	5.00 (1.73)	5.67 (1.15)	4.67 (1.53)	4.33 (0.58)
Physical	1.67 (0.58)	1.67 (0.58)	1.33 (0.58)	1.33 (0.58)	1.33 (0.58)
Temporal	1.67 (0.58)	1.67 (0.58)	1.33 (0.58)	1.33 (0.58)	1.33 (0.58)

Inferential Statistics. Model adequacy tests with normality and residuals plots revealed no major violations for the mental workload survey items. Of the 6 items from the NASA-TLX survey, only "effort to complete task" showed a significant difference for visualization types ($F_{4,\ 10} = 4.93$, $p = 0.0187$). Frustration ($F_{4,\ 10} = 2.4$, $p = 0.1192$), performance ($F_{4,\ 10} = 1.32$, $p = 0.3281$), physical demand ($F_{4,\ 10} = 0.3$, $p = 0.8714$), mental demand ($F_{4,\ 10} = 2.75$, $p = 0.0886$), and temporal demand ($F_{4,\ 10} = 0.04$, $p = 0.9964$) had no significant difference.

4.6 Conclusion

The results of heuristic evaluation, pilot study, empirical study, focus groups, and case studied revealed a variety of data and trends regarding the impact of visualization technique on user performance, usability opinion, perceived mental workload, and eye tracking measurements. In addition, qualitative data from the focus groups have given further insight into the visualization techniques' features and usage. Chapter 5 will discuss these results in detail and examine the contributions and implications of these findings within a global context.

CHAPTER 5

DISCUSSION

This chapter discusses the impact of this research in academia on evaluating information visualizations for emergency medical data, the strengths and weaknesses of each visualization technique used in this research, and a comparison of the visualizations for various purposes. In addition, recommendations are made for the developers of emergency department information systems that incorporate visualization techniques. Lastly, the limitations of these results are presented.

5.1 Implications of Results

One of the most significant contributions of this research is the development of a rigorous, comprehensive assessment methodology to evaluate information visualizations for emergency department information systems. This multi-phased methodology included a comprehensive literature review, user requirements analysis, development of a test bed interface, and studies with usability experts, novices, and domain experts. Other methods to evaluate information visualization techniques were compared to assist in the creation of this assessment method. This research provides a significant contribution in the academic conversation on healthcare information systems because it provides a method for evaluating the usability of a decision support tool. The methodology can be expanded to domains where visualization techniques are utilized in information systems that support complex decision-making, such as traffic control or cyber security. The following subsections will provide a detailed discussion about each information visualization technique.

5.1.1 Density Chart

The heuristic evaluators pointed out that the density chart while seemingly very straightforward had issues with the color scheme. A legend would be a necessity for using this visualization technique because without it, different people can interpret the same image differently. Without the legend, it is assumed that the user understands the color scheme. If users come from different fields or cultural backgrounds, chances are that the red-yellow-green-blue scheme may mean something different. One of the evaluators also pointed out the lack of a scheme to accommodate for color-blindness. This could cause disastrous consequences for healthcare decision makers. Two of the evaluators actually decided that this specific visualization should not be used in the medical field due to the color-blind issue.

In the empirical study and the case study, the density chart performed well with engineers for performance metrics, usability, and mental workload. There were some inconsistencies with the performance and usability among the nurses. During the focus groups with nurses, one of the positives pointed out was that the density chart provided "quick understanding" and was "very easy to learn". This comment implies that training to use this technique would be minimal for novices. However, the nurses did point out that there was “cluttering” and “words would hide other words on the image”. In addition, it is indicative that this type of visualization needs to have a quality data set for it to be effective. This technique is primarily used for the display of qualitative data, much like the network diagram and tree map. Of the three qualitative visualizations, this one in particular had the highest abandonment rate, implying that when users are unable to answer the question they will not guess, which may be related to the task itself and not the visualization.

5.1.2 Tree Map

In the heuristic evaluation study the tree map had six violations, one of the larger violations. However, the average severity rating was only 2.43 for these violations. This result indicates that there are a number of relatively small usability issues. The tree map, similar to the density chart and network diagram, performed well with the engineers in part one of the empirical study for performance metrics, usability, and mental workload. For the case study, the positive trend was similar except that this visualization had the lowest accuracy. During part one of the empirical study, of the three qualitative visualizations, the tree map technique had the lowest accuracy and the highest mental workload. This result supplements the findings of the heuristic evaluation and implies that within that group of qualitative visualization techniques, there are usability issues with this technique. This problem may be relative to the task or data set. There were some inconsistencies with nurses diverging from this trend. In general their performance was strong with the tree map but they rated the tree map with lower usability scores.

During both focus group sessions with the nurses, the difference in box sizes was brought up as a positive and a negative for this technique. On the one hand, the box sizes make it obvious and salient to the user, which variables are more frequently occurring. However, it is not obvious how much bigger a particular box is versus another providing more ambiguity. A follow up comment/suggestion to this was that this technique could be improved by adding a “roll-over pop-up” display to see more details when a user selects a particular box. A negative comment that came up was in regards to the distinguishability of the colors. Users could not tell the difference between two closely related colors in the legend (i.e. orange and peach). This issue was popular throughout both focus groups with all the visualizations and their respective legends.

5.1.3 Network Diagram

The usability experts noted five violations with the network diagram during the heuristic evaluation, a moderate number in comparison to other techniques. The average severity rating was 2.86, also a moderate value. For the network diagram, the evaluators were concerned with error protection. An increase in the number of nodes could cause an increase in user workload and create issues with detecting and preventing data entry. The addition of field labels and input values would further add to a convoluted diagram. The evaluators suggested a software feature to allow for magnification of particular clusters and areas of nodes to mitigate these issues.

The network diagram, similar to the density chart and tree map performed well in part one of the empirical study with novices for performance metrics, usability, and mental workload. Of the three qualitative visualizations, this technique scored the highest accuracy but lowest usability. This inconsistency is not a surprise given that it was previously discussed in this dissertation that a low correlation exists between performance and usability. During the focus groups, the nurses commented that this technique was “easy to grasp and start using right away”. The negative issues mentioned by the nurses mirror what has already been described in regards to the density chart and tree map i.e. lack of legend for color coding of groups and “words hiding other words”. The recommendation by the nurses for the addition of a highlighting feature was unique to this visualization.

For the case study, the trend somewhat reversed for the network diagram among the three qualitative techniques. This visualization had the highest mean time to completion, lowest accuracy, and lowest usability in comparison to the density chart and the tree map. This trend most likely differs from the trend found in part one of the empirical study due to the use of domain experts and not novices. Therefore, persons more familiar with this area may be more likely to give lower usability scores for network diagram.

5.1.4 3D Scatter Plot

The heuristic evaluators noted the 3D scatter plot had the largest number of violations, nine, and to have the highest average severity rating, 3.56, among all the visualization techniques. The legibility issue described with the 3D scatter plot was given a severity rating of four by the evaluators. Specifically, the concern was that the titles, labels, and legend were difficult to read due to location and font. Important information is conveyed by titles, labels, and the legends. If these features are not legible, users can misinterpret the data and cause errors in their decision-making processes. Furthermore, in the category of legibility, the different colors attributed to a specific data series may also contribute to user confusion. It was suggested to mitigate this issue that a legible font be used, titles should appear in the top center, and legends should be intuitive and placed in a clearly identifiable area. When a large, complex dataset is used, an increase in information density occurs with a 3D scatter plot. This information density could increase the user's workload and affect their decision-making performance. Without the proper software, the user would not be able to magnify certain parts of the plot to investigate patterns and outliers.

In part one of the empirical study with novices, the 3D scatter plot performed the worst of all the visualizations. This visualization's features (displays quantitative data) and its performance warrant that it is grouped with the scatter plot matrix and parallel coordinates when observing trends. This technique had the highest performance time, lowest usability, and highest workload of all visualizations. Among the group of three quantitative data visualizations, it scored the highest accuracy and lowest abandonment rate. These results imply that overall there are serious usability issues with the 3D scatter plot however it has some qualities over the other quantitative visualizations.

Noteworthy during the experiment, several of the engineers remarked that the limited functionality of the GUI impacted their perception of the use of this technique. The GUI's limitations will be discussed in a subsequent section. Notable comments from the focus groups were with respect to the axes and information clutter. The nurses felt that the axes were in poor locations towards the edge of the screen. This issue may be relative to the GUI itself and not the technique's usability. However, with regard to the visualization, the participants commented that it was "too cluttered" and difficult to zoom in and view details of the data. Due to the interaction issues brought up and their severe impact on gathering appropriate user data from part one of the empirical study, the 3D scatter plot was removed from the comparison of techniques for the case study.

5.1.5 Scatter Plot Matrix

The usability evaluators identified two violations (error management and significance of codes) with the scatter plot matrix, the least of all visualizations. However, the average severity score for the violations was 3.5. The heuristic evaluators noted that the design of the scatter plot matrix made it difficult for users to manage errors. The placement of the headings and associated axis values are not user-friendly. The headings and axes could distract users causing them to make errors when making decisions based on what they see in the matrix. According to one of the evaluators, the best way to mitigate this issue is to use the least number of variables. However, the number of variables is associated with the task at hand, which this implies that this visualization may be suitable for situations where there are fewer variables to analyze. Later in the focus groups with nurses, the issue with axes was brought up again; this time in regards to the axes “crowding each other” and “hiding values”.

In part one of the empirical study, the scatter plot matrix performed similar to the 3D scatter plot and the parallel coordinates. This visualization scored in the middle of the other two quantitative data visualizations (still low overall) for the performance measurements. Of these three, the scatter plot matrix had the highest usability score and lowest workload measurement. These findings suggest that while the three quantitative visualizations perform more poorly than the qualitative visualizations, the scatter plot matrix was the best among them. This suggests that the scatter plot matrix visualization may be better suited for some tasks over others, such as tasks with a lower number of variables. However, the trend in the results from the case study suggest differently. The results of the case study reported this visualization had the highest mean time and highest workload among all visualizations. Of course, these results are from a small sample and should be discussed with caution; particularly since this visualization had 100% accuracy for the case study.

In the second part of the empirical study, the scatter plot matrix was used to investigate the impact of visualization type and difficulty level as determined by number of variables. Overall, the trend for scatter plot matrix (and for the parallel coordinates) of the descriptive statistics suggest that an increase in the number of variables (higher difficulty level) to be viewed in a visualization will also increase performance time, decrease accuracy, and lower usability. However, with the scatter plot matrix, workload decreased with a higher difficulty level. This finding, with the results of the heuristic evaluation and focus groups, potentially indicates that this visualization may be suitable for large numbers of variables *if* the axes and labels are designed with clarity. Additionally, the second part of the empirical study was completed with only ten participants and the results should also be discussed with caution due to sample size.

5.1.6 Parallel Coordinates

The heuristic evaluators identified five violations for the parallel coordinates, with an average severity rating of 2.4. Although the evaluators agreed that this visualization had minor issues, one of the evaluators commented that this visualization did have a potential to be a significant usability problem with an increase in variables and a combination of text and numerical data variables. Another evaluator noted that a change in the order of the vertical axes could dramatically change how an image looks, requiring that users may have to look at the same set of data in several different ways to finally see a pattern.

The parallel coordinates performed similarly to the 3D scatter plot and scatter plot matrix in part one of the empirical study. This visualization had the lowest accuracy and the highest abandonment rate for all subjects. As for usability and workload, this technique was in the middle of the 3D scatter plot and scatter plot matrix. As for nurses, this visualization had the highest performance time, the lowest usability, and the highest workload. Eye tracking results indicated that users spent more time looking at this visualization. To note, the eye tracking metrics reported in Chapter 4 did not account for the time that the user spends looking and/or reading the question and answers. This result could indicate that this task required the user to look less at the text and concentrate more on the images.

In the focus groups, the nurses were given the opportunity to shed light on their experiences with this visualization since their group performed overwhelmingly poor. Each vertical axis on the parallel coordinates is associated with the range for a particular variable. Therefore, in a given circumstance each axis has a completely different scale with the axis next to it. The participants commented that with so many axes it is "too easy to get lost." One of the features of this visualization allows the user to filter data and add colors to identify subsets of data. However, the groups did not think the coloring assisted with identifying subsets and that this visualization “shouldn't be used to look for subsets” but just for “anomalies or outliers.”

The descriptive results of the empirical study (part 2) indicated that the parallel coordinates visualization technique, in comparison to the scatter plot matrix, performed worst at both difficulty levels for all metrics. The analysis of variance indicated a statistical significant difference on the impact of visualization type on time but not for difficulty level. This finding may indicate that for this technique, an increase in variables does not change performance, usability, or mental workload measurements. This would make this technique unlike the scatter plot matrix whose qualitative results suggested the number of variables does influence user opinion. As previously mentioned, the sample size is small and the results should be interpreted with caution.

The results of the case study indicated the domain experts had higher accuracy with the parallel coordinates than did the novices in the empirical study. In contrast, this visualization had the lowest usability scores and low to moderate mental workload results from the case study results. This trend has been seen before when a technology has improved performance but still had low user opinion. This finding potentially indicates this technique may need a detailed user-adoption plan, i.e. the users do not like it because they do not know enough about it.

5.1.7 Summary of Trends

The general trend emerging from the results presented from the studies in Chapter 4 indicates the density chart, tree map, and network diagram have lower times, higher usability scores, and lower mental workload ratings than the 3D scatter plot, scatter plot matrix, and parallel coordinates. There were a few exceptions to the trends when examining the differences between engineers and nurses. However, these exceptions were usually given highlight in the focus groups. This trend could be a result of either the tasks or the data format. The tasks associated with the density chart, tree map, and network diagram asked users to compare variables, specifically by frequency or amount of a variable. However the tasks associated with the 3D scatter plot, scatter plot matrix, and parallel coordinates asked the users to compare correlations of variables. Another possible reason for this trend is that the 3D scatter plot, scatter plot matrix, and parallel coordinates are based on numerical data, for example, blood pressure or pulse rate. However, the parallel coordinates does allow for a user to look at correlations of textual and numerical data, for example, the correlation between gender (male or female) and pulse rate. The other three visualizations were developed from textual data sources. The symptoms and diagnoses of patients in the emergency department were used to build the network diagram and density chart.

When analyzing all the visualizations, the results of empirical study part one clearly indicate that visualization type has an impact on performance measurements for emergency medical data tasks for novices. Previous research studies that measured performance on an individual task (instead of a set of tasks) have shown similar results that indicated that different visualization techniques performed differently for various tasks (Plaisant, 2004).

The results of empirical study part two indicated the same impact from visualization type on performance measurements, which are in contrast to previous studies (Speier & Morris, 2003). A statistical difference was not found for difficulty level on performance; however, this study was limited and only analyzed two visualizations.

Speier and Morris (2003) hypothesized that subjective mental workload varies between different display designs when task complexity is low, and when task complexity is high, the mental workload will be lower for visual query display than text-based query display. There was no statistical difference between difficulty levels for mental workload for this research. If this study were repeated issues to address include modifying the task, the data, the visualization type, or increasing sample size.

The eye tracking results gave more light to the interaction of users with the visualization as they completed the task. This was the only data that was taken during the task as opposed to being measured after the user completed the task. The results indicated there were differences in user eye movements as they viewed the different visualizations. Further explorations of these findings have the potential to supplement the other results in detecting usability issues. Many user comments were concerned the axis and eye tracking results which indicated there were significant eye movements in those areas of the visualizations.

The case study results, while limited, are still helpful in evaluating these visualizations. The case study was done with domain experts, not novices as in the empirical studies or focus groups. The data collected from the experts can be used to refine this methodology framework, in particular the tasks. The usability survey item "value as a healthcare tool" scored unusually low with the novices; however, the domain experts saw more value in these tools for healthcare than the novices. This is not to say that the feedback from novices is not helpful. From the focus groups, the nurses agreed that the use of information visualization techniques for emergency department information systems is beneficial because they allow a user "to see a lot of records at once" which would reduce search time. However, they also noted these images would be best if they were add-ons to the current reports they already have. In usability engineering, the input of the end-user, novice or expert, is valuable during all stages of the design cycle.

5.2 Guidelines for Developers

Information visualization is an effective way to view and understand large amounts of emergency medical patient data. However, the designers of emergency department information systems must be aware of the ways in which their system may be used to its fullest potential and estimate its impact on the end-user. Certain visualization techniques are appropriate for specific tasks and situations. An EDIS should be designed to integrate the usage of an appropriate technique for a health care decision-making task to mitigate potential usability issues. This section will present five main categories of guidelines. To identify and assist with the design process, several guidelines based on this research were developed for designers of such systems.

Respect Real Estate. Critical to the design of any information system that utilizes information visualization techniques is to design with concern for the space. Information visualization is used to deal with large amounts of data; however, even in a compressed state the data representation may take up a large amount of space. The designer must allow the users to set their own parameters for the data filtering beyond the system defaults. An issue raised in this research was in regards to the frequency and the font size with the density charts. Some comments from the focus group about the density chart, network diagram, and tree map implied frustration on the part of the users because these visualization techniques had "words hiding other words". Users wanted to change the font sizes on words with respect to their own frequency filter. An example would be allowing users to filter out variables that occurred less than some frequency (say a symptom that occurred less than 100 times) to de-clutter the screen. This negative aspect of the aforementioned visualizations is particular to this research on EDIS. However "cluttering" could be generalized to visualizations for other complex information systems and this should be given attention in the design phase. With respect to space, the axes and legends should also be clear, large enough to see but not covering data points, in obvious locations, intuitive, and should not cover one another. Issues with axes and legends were mentioned in almost every study of this research.

Get Their Attention. Information visualizations require human attention for the data to have meaning. It is necessary for designers to investigate and determine the best way to support human information processing with regard to stages of attention and cognition. Think about the implications for working memory and automatic attention filters users may already have. An example of this principle would be the use of flashing messages to get someone's attention inadvertently (i.e. instant messaging systems). A designer may want to integrate a feature into such a system where when the visualization changes, the screen flashes. Designer should not be afraid to direct users to something on the screen. In the focus groups, users mentioned that "highlighting" would be a useful feature to add to the network diagram. Caution should be exercised when using color coding to get a user's attention because the colors need to be highly distinct from one another to make it past someone's filter. One of the negative comments about the tree map was that the colors were not distinguishable from one another making it difficult to identify the most frequently occurring entity. In addition, in the emergency medical domain there are always going to be environmental distractions so it is necessary to design the system to get attention even when a distraction may occur.

Educate the User. Training is typically something that happens after a system is developed and installed on site. Designers of EDIS that employ information visualization techniques can still assist in this endeavor. Users should be trained to understand the system and how to use it as a decision support tool. Both can be assisted by embedding a user manual into the system, a task a designer can accomplish. Once a user understands a visualization well enough to customize it for their role/responsibilities it becomes easier to adopt the new technology. In many other human-computer interaction studies and throughout this research, a low correlation between performance and usability score exists. However the nurses in the empirical studies did not show a low correlation between performance and usability. The nurses of the focus groups made positive comments about the density chart and network diagram being "easy to learn" and "quick understanding". These comments supported the performance and usability results of those nurses. If designers can help the user understand how to use the visualization technique, this can improve their performance and user opinion.

Know the User. An EDIS that use information visualization to display data needs to be flexible and adaptable. Within an emergency department a user's role and tasks will change. During the heuristic evaluation, the evaluators brought up the fact that the parallel coordinate visualization would give the same user different perspectives on the same data set if the order of the axes were changed. The evaluator meant that comment as a negative aspect. However, focus group comments indicated that giving a user a different perspective on the same data set shows flexibility of the visualization and supports multiple tasks of the user. Designers need to be familiar with users' characteristics, level of experience, domain/task characteristics, cultural background, etc. in order to build a system that will support a wider range of roles. It has been shown that acceptability is higher when multiple users can use a system. A user requirements analysis on site can provide insight into the needs of users (not to be confused with the wants of users). Designers should support the use of the best visualization technique or a combination of visualizations for tasks in an emergency department.

Know the Task. This category is an extension of "Know The User". To optimize the user performance and experience an EDIS should be built with concern as to how it supports and integrate into the emergency medical staff's decision-making models. Designers should take into account various decision-making models in healthcare from the broad range of literature available. If a designer knows a user needs more details to solve a problem, then he should provide those details where they can quickly access them. This issue surfaced during the heuristic evaluation and again in the focus groups where users wanted more opportunity to zoom in and get specific details about data points, in particular for the tree map and network diagram. However, remember the principle "Respect Real Estate". Developers must also keep in mind that an information visualization technique can only be as good as the data it is built from. Emergency departments can be chaotic and often reports are left unfinished until a later time. A visualization technique for emergency medical data should be able to accommodate for poor (or missing) data input.

The principles presented here represent the initial stages of developing a set of detailed guidelines for the development of emergency department information systems (EDIS) that employ the use of visualization techniques for large data sets. In conjunction with the assessment framework, these guidelines provide significant contribution to the domain of emergency medical data visualization. Both components are unique in that they were specifically designed for emergency medical patient data visualizations and are the result of a comprehensive evaluation methodology. Future work is needed to fully capitalize on the results of this research and these guidelines established in this research.

5.3 Limitations of Research

Several limitations were mentioned previously: fidelity and functionality of the graphical user interface, the heuristics used in the heuristic evaluation, small sample size for empirical study part two, small number of domain experts, and the limited framework scope. First, the graphical user interface (GUI) that was developed as the test bed for this framework had limited interaction features for the visualizations. One crucial aspect of visualizations is users' interaction with data sets. The 3D scatter plot only allowed users to zoom, overview, and to slide left or right. However, in its real context users are able to rotate the scatter plot on the axes. Additionally in a field test, the environment would be far more distracting and complex than the setting of a usability laboratory.

The heuristics used by the usability experts may not be appropriate because information visualization techniques need their own heuristic list and that has yet to be developed. However, Bastien and Scapin's list has been used in previous visualization heuristic evaluations. Second, the tasks and user scenario that were provided to the evaluators were developed by subject matter experts; however, they may have limited the evaluators' scope when inspecting a visualization technique. Lastly, only three evaluators were used in the heuristic evaluation and additional evaluators may have provided further in-depth and insightful results.

The sample size for the empirical study part two was small and made it difficult to assess the inferential statistics. Increasing the sample size could increase reliability, sensitivity, and statistical power of the analysis. Additionally, more domain experts were needed. Unfortunately, due to the nature of their work domain experts are difficult to obtain for these types of studies.

The assessment framework developed for this research is limited to the application of EDIS. The application of the framework or generalizing the results to other domains should be done with caution and only to other complex decision-making environments. Further, this assessment while comprehensive, only used a few selected usability testing methods to assess the visualizations.

CHAPTER 6

CONCLUSION

Information visualization is a method of effectively communicating large amounts of data. User interaction gives this data “meaning”. If this meaningful data, or information, is used in complex, critical systems such as an emergency department information system (EDIS), usability issues should be addressed. Many of today’s visualization techniques have not been developed with these issues in consideration. To assist with identifying these issues, this dissertation proposed a framework for evaluating information visualization techniques with specific regards to usability issues for emergency medical data and provided guidelines for developers to create systems that enhance the user’s cognition and support their decision-making. This chapter will present a summary of this research, its contributions, and suggestions for future work.

6.1 Summary of Dissertation

This dissertation demonstrated the complexity of evaluating information visualization for EDIS. Within this domain, decision-making is multi-dimensional, time-dependent, dense, and can often result in fatality. For these reasons, clinical decision support systems have been developed and used to improve the quality of patient care. However, these systems are not without their associated issues with performance, user opinion, and additional burden on mental workload. Effective evaluation methods are necessary to improve the technology and decision-making of healthcare professionals.

The purpose of this research was to construct a theoretical framework for a multi-layered evaluation approach to measure the efficiency of several information visualization techniques for an emergency medical patient dataset. In support of the theoretical framework, a test bed interface was designed to be used in the evaluation. For multiple reasons, a heuristic evaluation was completed with usability experts for the visualization test bed and to identify usability issues. Empirical testing in a controlled lab setting with novices was completed to collect objective and subjective user data. Qualitative data was gathered from focus group sessions with those same novices. Lastly, feedback from domain experts was assembled from a case study with medical professionals. The composition of all data collected was used to create a set of guidelines for designers of EDIS.

Results indicated the visualization type was found to have an impact on performance, usability, mental workload, and eye tracking metrics. Further, qualitative data visualizations elicited faster performance times, better accuracy, higher usability, and lower mental workload ratings among novices. Although no significant differences were found among the dependent variables for difficulty level, that aspect of the research may be influenced by a small sample size. These results yield a deep understanding of factors to be considered when evaluating the effectiveness of visualization techniques.

6.2 Contribution of Research

The core contributions of this dissertation can be divided in two areas: a framework and a set of guidelines. Using a comprehensive assessment method of information visualization has provided significant insight to usability issues with EDIS from the results gathered in the studies. The methodology framework developed for this dissertation was comprised of several types of traditional usability tests and measured multiple dependent variables simultaneously. The variables were a mixture of direct performance measurements, user opinion on usability, user perception of mental workload, eye tracking metrics, and qualitative feedback. Designing this comprehensive framework made it possible to investigate the effects of visualization type and difficulty level on multiple variables, as well as proposing a set of guidelines for developers to design efficient EDIS. Since there has been a lack of effort on evaluating information visualization techniques that support emergency medical decision-making, this dissertation filled a gap in previous research studies. The approach developed herein was designed specifically for this domain and this type of information display, and has contributed to the knowledge bank of human factors in information visualization.

6.3 Future Work

Regardless of the contributions of this research, there are several areas for improvement using this framework for assessment. The heuristic evaluation was completed with three usability experts with no background in the domain. The results of such an evaluation could be significantly enhanced if the evaluators were increased to five and they had backgrounds in both emergency department decision-making and usability. It has already been mentioned that the GUI used for this research was limited. A GUI designed with more capabilities would provide additional insight to user interaction. The experimental design was a within-subject design; however, a mixed between-subject/ within-subject design would be beneficial in revealing more insight into particular groups of users. In this study four classes of metrics were gathered (performance, usability, workload, and eye tracking); however, more metrics may be included in the future. In addition, it is not known if one variable is more useful than another in determining the effectiveness of a visualization technique (all variables were weighted evenly). Future work should definitely include the addition of more domain experts for the case study and possibly field testing to provide more validity to the results obtained from this framework. Chapter 5 noted that the tasks may have been over-simplified, and if this methodology were to be utilized, the tasks should reflect the complexity from the user requirements analysis. Perhaps the addition of ranking tasks would suffice. Accounting for all aspects of evaluating a technology from a human factors perspective can be quite complex (as is human performance), and extremely difficult, but there is always a need to deal with shortfalls of other assessment methods.

REFERENCES

Au, P., Carey, M., Sewraz, S., Guo, Y., & Rüger, S. M. (2000). New Paradigms in Information Visualization. In Belkin, N. J., Ingwersen, P., Leong, M. (Eds.) *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, July 24-28, 2000, Athens, Greece, 307-309.

Ball, M. J. (2003). Hospital Information Systems: Perspectives on Problems and Prospects, 1979 and 2002. *International Journal of Medical Informatics, 69 (*2-3), 83-89.

Balogun, Y., Chenou, J., Jenkins, Q., & Park, G. (2009) *Project Report: Vygotsky Health Center Emergency Department Information System.* Unpublished report, North Carolina A&T State University, Greensboro, NC.

Bastien, J., & Scapin, D. L. (1993). *Ergonomic Criteria for the Evaluation of Human-Computer* (Technical Report 156). France: Institut National de Recherche en Informatique et en Automatique (INRIA).

Bertram, D. A., Opila, D. A., Brown, J. L., Gallagher, S. J., Schifeling, R. W., Snow, I. S., & Hershey, C. O. (1992). Measuring Physician Mental Workload: Reliability and Validity Assessment of a Brief Instrument. *Medical Care*, 95-104.

Bourzac, K. (2007) *Mapping Complex Diseases*. In Technology Review: Biomedicine. Retrieved on July 1, 2011 from http://www.technologyreview.com/biomedicine/19023/

Budnitz, D. S., Shehab, N., Kegler, S. R., & Richards, C. L. (2007). Medication Use Leading to Emergency Department Visits for Adverse Drug Events in Older Adults. *Annals of Internal Medicine, 147*(11), 755.

Buehler, J. W., Sonricker, A., Paladini, M., Soper, P., & Mostashari, F. (2008). Syndromic Surveillance Practice in the United States: Findings from a Survey of State, Territorial, and Selected Local Health Departments. *Advances in Disease Surveillance, 6*(3), 1–20.

Bui, A. A. T., Aberle, D. R., & Kangarloo, H. (2007). TimeLine: Visualizing Integrated Patient Records. *IEEE Transactions on Information Technology in Biomedicine, 11*(4), 462-473.

Camoes, J. (2008, March 10, 2008). Fibonacci, Working Memory and Information Overload. Retrieved on July 1, 2011 from http://www.excelcharts.com/blog/fibonacci-working-memory-and-information-overload/

Canossa, A., & Drachen, A. (2009). Patterns of Play: Play-Personas in User-Centered Game Development. In *Breaking New Ground: Innovation in Games, Play, Practice and Theory, Proceedings of Digital Games Research Association*, September 1–4, 2009. Brunel University, West London, United Kingdom.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufmann.

Carpendale, S. (2008). Evaluating Information Visualizations. In Kerren, A., Stasko, J. T. Fekete, J.D. and North C., (Eds.) *Information Visualization: Human-Centered*

Issues and Perspectives, 19-45. Verlag Berlin Heidelberg: Springer.

Cawthon, N., & Moere, A. V. (2007). The Effect of Aesthetic on the Usability of Data Visualization. In *Proceedings of IEEE International Conference on Information Visualization*, July 2- 6, 2007. Zurich, Switzerland, 637–648.

Center for Disease Control and Prevention (2010). *Emergency Department Visits*. Retrieved on April 21, 2011 from http://www.cdc.gov/nchs/fastats/ervisits.htm

Center for Disease Control and Prevention-National Center for Health Statistics. (2008). National Hospital Ambulatory Medical Care Survey. *United States Department of Health and Human Services. Centers for Disease Control and Prevention.* Hyattsville, Maryland: Public Health Service.

Center for Disease Control and Prevention –National Center for Health Statistics. (2009). National Hospital Ambulatory Medical Care Survey 2008 User Guide. *United States Department of Health and Human Services. Centers for Disease Control and Prevention.* Hyattsville, Maryland: Public Health Service.

Chapman, W. W., Dowling, J. N., Ivanov, O., Gesteland, P., Olszewski, R., Espino, J., & Wagner, M. M. (2004). *Evaluating Natural Language Processing Applications Applied to Outbreak and Disease Surveillance*. Paper presented at 36th Symposium on the Interface: Computing Science and Statistics; May 26, 2004. Baltimore, Maryland.

Chen, C. (2004). *Information Visualization: Beyond the Horizon.* New York: Springer-Verlag.

Cheng, C., Shahar, Y., Puerta, A., & Stites, D. (1997). Navigation and Visualization of

Abstractions of Time-Oriented Clinical Data. *Section on Medical Informatics Technical Report 97.* Stanford University, CA.

Chittaro, L. (2006). Visualizing Information on Mobile Devices. *In IEEE Computer Society Press,* 39 (3), 40–45.

Compeau, D. R., & Higgins, C. A. (1995). Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly*, 189-211.

Cooper, A. (1999). *The Inmates are Running the Asylum: Why High Tech Products Drive Us Crazy and how to Restore the Sanity*. Indianapolis, Indiana: Pearson Higher Education.

Duke, J., Faiola, A., & Kharrazi, H. (2009). A Novel Visualization Tool for Evaluating Medication Side-Effects in Multi-Drug Regimens. *Human-Computer Interaction: Interacting in Various Application Domains*, 478-487.

Eppler, M., Mengis, J. (2003). A Framework for Information Overload Research in Organizations. *Insights from Organization Science, Accounting, Marketing, MIS, and Related Disciplines*, Lugano, Switzerland: Institute for Communications Management.

Etzioni, A. (1989). Humble Decision-making. *Harvard Business Review on Decision Making*, 67(4), 122-126. Boston, Massachusetts: Harvard Business School Publishing.

Faiola, A., & Matei, S. A. (2010). Enhancing Human–Computer Interaction Design Education: Teaching Affordance Design for Emerging Mobile Devices. *International Journal of Technology and Design Education*, 1-16.

Falkman, G. (2000). Similarity Measures for Structured Representations: A Definitional Approach. *Advances in Case-Based Reasoning*, 533-562.

Favela, J., Rodríguez, M., Preciado, A., & González, V. M. (2004). Integrating Context-Aware Public Displays into a Mobile Hospital Information System. *IEEE Transactions on Information Technology in Biomedicine, 8*(3), 279-286.

Fischer, S., Lowe, R., & Schwan, S. (2008) Effects of Presentation Speed of a Dynamic Visualization on the Understanding of a Mechanical System. *Applied Cognitive Psychology 22*, 1126–1141.

González, V., & Kobsa, A. (2003). A Workplace Study of the Adoption of Information Visualization Systems. In *Procedures of Third International Conference on Knowledge Management (I-KNOW)*, July 2-4, 2003. Graz, Austria, 92-102.

Grisé, M. L., & Gallupe, R. B. (1999). Information Overload: Addressing the Productivity Paradox in Face-To-Face Electronic Meetings. *Journal of Management Information Systems, 16*(3), 157-185.

Haux, R. (2002). Health Care In The Information Society: What Should Be The Role Of Medical Informatics? , *Methods of Information in Medicine, 41*(1), 31-35.

Hibbard, J. H., & Peters, E. (2003). Supporting Informed Consumer Health Care Decisions: Data Presentation Approaches that Facilitate the Use of Information in Choice. *Annual Review of Public Health, 24*(1), 413-433.

Huang, W. (2007). Using Eye Tracking to Investigate Graph Layout Effects. *Asia Pacific Symposium on Visualization,* February 5-7, 2007. Sydney, Australia, 97-100.

Huang, W., Eades, P., & Hong, S. H. (2009). Measuring Effectiveness of Graph

Visualizations: A Cognitive Load Perspective. *Information Visualization*, 8(3): 139-152.

Hudson, D., Cohen, M., & Anderson, M. (1991). Use of Neural Network Techniques in a Medical Expert System. *International Journal of Intelligent Systems, 6*(2), 213-223.

Iselin, E. R. (1988). The Effects of Information Load and Information Diversity on Decision Quality in a Structured Decision Task. *Accounting, Organizations and Society*, *13*(2), 147-164.

Jensen, T. B., & Aanestad, M. (2006). How Healthcare Professionals “Make Sense” of an Electronic Patient Record Adoption. *Information Systems Management, 24*(1), 29-42.

Kaptelinin, V. (1996). Activity Theory: Implications for Human-Computer Interaction. *Context and Consciousness: Activity Theory and Human-Computer Interaction*, 103-116.

Keim, D. A. (2002). Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics, 8(1),1-8.*

Kendler, J. (2004). Visual Artifacts as Decision Analysis and Support Tools. Retrieved April 21, 2012 from http://www.wiklundrd.com/kendler_decisionmaking.pdf

Kosara R. (2007). Visualization Criticism – The Missing Link Between Information Visualization and Art. In *Proceedings of IEEE International Conference on Information Visualization*, July 2- 6, 2007, Zurich, Switzerland, 631–636.

Kroeze, J. H., Matthee, M. C., & Bothma, T. J. D. (2003). *Differentiating Data and Text-*

Mining Terminology. September 2003, Republic of South Africa: South African Institute for Computer Scientists and Information Technologists, 93-101.

Kunitz, S. J. (1974). Professionalism and Social Control in the Progressive Era: The Case of the Flexner Report. *Social Problems, 22*, 16.

Kuutti, K. (1996). Activity Theory as a Potential Framework for Human-Computer Interaction Research. *Context and Consciousness: Activity Theory and Human-Computer Interaction*, 17-44.

Lengler, R., & Eppler, M. J. (2007). Towards a Periodic Table of Visualization Methods for Management. In *Proceedings of the Conference on Graphics and Visualization in Engineering*, January 3 – 5, 2007, Clearwater, Florida, USA, 1-6.

Leontiev, A. (1978). *Activity, Consciousness, and Personality*. Englewood Cliffs, NJ: Prentice Hall.

Li, L., Xu, L., Jeng, H. A., Naik, D., Allen, T., & Frontini, M. (2008). Creation of Environmental Health Information System for Public Health Service: A Pilot Study. *Information Systems Frontiers, 10*(5), 531-542.

Linder, J. A., Ma, J., Bates, D. W., Middleton, B., & Stafford, R. S. (2007). Electronic Health Record Use and the Quality of Ambulatory Care in the United States. *Archives of Internal Medicine, 167*(13), 1400.

Lucas, N., Crémilleux, B., & Turmel, L. (2003). Signalling Well-Written Academic Articles in an English Corpus by Text Mining Techniques. *UCREL Technical Papers, 16 (Special Issue Proceedings Corpus Linguistics*), 465-474.

MacEachren, A. M., & Stryker, M. I. (2008) Health GeoJunction: Geovisualization of

News and Scientific Publications to Support Situation Awareness. Paper presented at Geospatial Visual Analytics Workshop, GIScience September 23, 2008 at Park City, UT.

Maciejewski, R., Hafen, R., Rudolph, S., Larew, S. G., Mitchell, M. A., Cleveland, W.S., & Ebert, D. S. (2010). Forecasting Hotspots- A Predictive Analytics Approach. *IEEE Transactions on Visualization and Computer Graphics*, 440-453.

Mackay, J. M., & Elam, J. J. (1992). A Comparative Study of How Experts and Novices Use a Decision Aid to Solve Problems in Complex Knowledge Domains. *Information Systems Research, 3*(2), 150.

Mirel, B. (1998). Minimalism for Complex Tasks. In *J. M. Carroll (Ed.), Minimalism Beyond the Nurnberg Funnel*. Cambridge, MA: MIT Press. 179-218.

Mirel, B. (2003). Dynamic Usability: Designing Usefulness into Systems for Complex Tasks. *Content and Complexity: Information Design in Technical Communication*, 233-261.

Montgomery, D. (2005) *Design and Analysis of Experiments*. New Jersey: John Wiley and Sons, Inc.

Moore, D. A., & Carpenter, T. E. (1999). Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology. *Epidemiologic Reviews*, *21*(2), 143-161.

Morse, E., Lewis, M., & Olsen, K. A. (2000). Evaluating Visualizations: Using a Taxonomic Guide. *International Journal of Human-Computer Studies, 53*(5), 637-662.

Musen, M. A., Shahar, Y., & Shortliffe, E. H. (2006). Clinical Decision-Support Systems. *Biomedical Informatics*, 698-736.

Nardi, B. A. (Ed.) (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction*: The MIT Press.

Nielsen, J. (1994). *Usability Inspection Methods*. New York: Wiley.

O'Reilly, C. A., & Caldwell, D. F. (1980). Job Choice: The Impact of Intrinsic and Extrinsic Factors on Subsequent Satisfaction and Commitment. *Journal of Applied Psychology, 65*(5), 559.

Patel, V. L., Zhang, J., Yoskowitz, N. A., Green, R., & Sayan, O. R. (2008). Translational Cognition for Decision Support in Critical Care Environments: A Review. *Journal of Biomedical Informatics*, *41*(3), 413-431.

Pillat, R. M., Valiati, E. R. A., & Freitas, C. M. D. S. (2005). Experimental Study on Evaluation of Multidimensional Information Visualization Techniques. In *Latin American Conference on Human-Computer Interaction*, October 23-26, 2005, Cuernavaca, México, 20–30.

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2000). The Effect of Information Scent on Searching Information: Visualizations of Large Tree Structures. In Di Gesu, V. Levialdi, S. Arantino, L. (Eds.) *Proceedings of the Working Conference on Advanced Visual Interfaces*, May 23-26, 2000, Palermo, Italy, 161-172.

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2001). Visual Information Foraging in a Focus+ Context Visualization. Presented at Conference on Human Factors in Computing Systems, March 31- April 5, 2001, Seattle, Washington.

Plaisant, C. (2004). The Challenge of Information Visualization Evaluation. *Proceedings of Conference on Advanced Visual Interfaces*, May 25- 28, 2004, Gallipoli, Italy, 109-116.

Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). *LifeLines: Visualizing Personal Histories*. *Proceedings of Conference on Human Factors in Computing Systems*, April 13- 18, 1996, Vancouver, British Columbia, Canada, 221-227.

Plaisant, C., & Schneiderman, B. (2004). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. New Jersey: Pearson Higher Education

Pohl, M., Rester, M., & Wiltner, S. (2007). Usability and Transferability of a Visualization Methodology for Medical Data. In *Proceedings of the 3rd Symposium of HCI and Usability for Medicine and Health Care*, November 27, 2007, Graz, Austria, 171–184.

Ratib, O., Swiernik, M., & McCoy, J. M. (2003). From PACS to Integrated EMR. *Computerized Medical Imaging and Graphics, 27*(2-3), 207-215.

Recarte, M. A., & Nunes, L. M. (2003). Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making. *Journal of Experimental Psychology: Applied, 9*(2), 119.

Reiser, S. J. (1991). The Clinical Record in Medicine Part 1: Learning from Cases. *Annals of Internal Medicine*, *114*(10), 902.

Rester, M., Wiltner, S. P., Hinum, K. (2007). Mixing Evaluation Methods for Assessing the Utility of an Interactive InfoVis Technique. In Jacko, J. (Ed.) *Proceedings of* *the 12th International Human-Computer Interaction Conference*, July 22-27, 2007, Beijing, China, 604–613.

Reyna, V. F. (2004). How People Make Decisions that Involve Risk. *Current Directions in Psychological Science*, *13*(2), 60.

Rothenhaus, T., Kamens, D., Keaton,B., Nathanson, L., Nielson, J., McClay, J., Taylor, T., Villarin, A. (2009). Emergency Department Information Systems: Primer for Emergency Physicians, Nurses, and IT Professionals. American College of Emergency Physicians. Retrieved on April 21, 2011 from http://apps.acep.org/WorkArea/DownloadAsset.aspx?id =45756

Rothenhaus, T., Kamens, D., McClay, J., & Coonan, K. (2007). Emergency Care Special Interest Group- HL7 Emergency Department Information System Functional Profile. Retrieved on April 21, 2011 from http://www.hl7.org/Special/committees/emergencycare/index.cfm

Rzhetsky A, Wajngurt D, Park N, Zheng T. (2007) Probing Genetic Overlap Among Complex Human Phenotypes. *Procedures of National Academy of Science USA*,104, 11694–11699.

Salogarvi, J., Kojo, I., Jaana, S., & Kaski, S. (2003). Can Relevance be Inferred from Eye Movements in Information Retrieval? In *Proceedings of the Workshop on Self-Organizing Maps*, September 11-14, 2003, Hibikino, Kitakyushu, Japan, 261-266.

Senders, M. G. J. (2009). Human Error in Road Accidents. Retrieved on January 20, 2011 from http://www.visualexpert.com/Resources/roadaccidents.html

Shneiderman, B. (2002). Inventing Discovery Tools: Combining Information Visualization with Data Mining. *Information Visualization, 1*(1), 5.

Soukup, T., & Davidson, I. (2002). *Visual Data Mining: Techniques and Tools for Data Visualization and Mining*. New York: John Wiley & Sons.

Speier, C., & Morris, M. G. (2003). The Influence of Query Interface Design on Decision-Making Performance. *MIS Quarterly*, 397-423.

Spence, R. (2001). *Information Visualization*. Harlow, United Kingdom: Addison Wesley/ACM Press.

Stepanovich, P. L., & Uhrig, J. D. (1999). Decision Making in High-Velocity Environments: Implications for Healthcare. *Journal of Healthcare Management/American College of Healthcare Executives*, *44*(3), 197.

Svensson, E., Angelborg-Thanderz, M., Sjoberg, L., & Olsson, S. (1997). Information Complexity-Mental Workload and Performance in Combat Aircraft. *Ergonomics, 40*(3), 362-380.

Swayne, D. A., Yang, W., Voinov, A., Rizzoli, A., & Filatova, T. (2010) Choosing the Right Data Mining Technique: Classification of Methods and Intelligent Recommendation. *Proceedings of International Environmental Modelling and Software Society*, July 5-8 2010, Ottawa, Ontario, Canada.

Tang, P. C., & McDonald, C. J. (2001). Computer-Based Patient-Record Systems. *Medical Informatics: Computer Applications in Health Care and Biomedicine.* New York: Springer-Verlag, 327-358.

The American College of Emergency Physicians (2003). *Emergency Department*

Technology. Retrieved on April 21, 2011 from http://www.acep.org/News-Media-top-banner/Emergency-Department-Technology/

Thorndike, A. N., Regan, S., & Rigotti, N. A. (2007). The Treatment of Smoking by US Physicians During Ambulatory Visits: 1994 2003. *American Journal of Public Health*, *97*(10), 1878.

Thyvalikakath, T. P., Monaco, V., Thambuganipalle, H. B., & Schleyer, T. (2008). A Usability Evaluation of Four Commercial Dental Computer-Based Patient Record Systems. *The Journal of the American Dental Association*, *139*(12), 1632.

Tory, M., & Möller, T. (2004). Rethinking Visualization: A High-Level Taxonomy. In Ward, M. O., & Munzer, T. (Eds.) *Proceedings of 10th IEEE Symposium on Information Visualization*, October 10-12, 2004, Austin, TX, 151–158.

Tory, M., & Torsten, M. (2004). Human Factors in Visualization Research. *IEEE Transactions on Visualization and Computer Graphics*, 72-84.

Tufte, E. R. (1991). Envisioning Information. *Optometry & Vision Science*, *68*(4), 322.

Van Raaij, J., Peek, M., Vermaat-Miedema, S. H., Schonk, C. M., & Hautvast, J. (1988). New Equations for Estimating Body Fat Mass in Pregnancy from Body Density or Total Body Water. *The American Journal of Clinical Nutrition*, *48*(1), 24.

Wehrend, S., & Lewis, C. (1990). A Problem-Oriented Classification of Visualization Techniques. Proceedings of IEEE Visualization, October 23–26, 1990, San Francisco, CA, 139–143.

Wei, F., Liu, S., Song, Y., Pan, S., Zhou, M. X., Qian, W., Shi, L., Tan, L., & Zhang Q. (2010). TIARA: A Visual Exploratory Text Analytic System. In *16th ACM*

SIGKDD International Conference on Knowledge Discovery and Data Mining, July 25–28, 2010, Washington, DC, USA, 153–162.

Wei, F., Shi, L., Tan, L., Sun, X., Lian, X., Liu, S., & Zhou, M. X. (2010) Visual Content Correlation Analysis. *Proceedings of the First International Workshop on Intelligent Visual Interfaces for Text Analysis*, February 7-10, 2010, Hong Kong, China, 25-28.

Whitney, H. (2010). *Beyond the Medical Chart: Information Visualization For Improving Personal And Public Health*. Retrieved on June 13, 2011 from http://www.uxmag.com/design/beyond- the-medical-chart

Wickens, C. D. (1992). *Engineering Psychology and Human Performance*. New York: Harper Collins.

Wickens, C. D., Pringle, H. L., & Merlo, J. (1999). Integration of Information Sources of Varying Weights: The Effect of Display Features and Attention Cueing. Army Research Lab Technical Report 99-2. Savoy, IL: Institute of Aviation- Aviation Research Lab.

Wickens, C. D., Lee, J., Liu, Y., & Becker, S. (2003). *Introduction to Human Factors Engineering*. Boston: Pearson Prentice-Hall.

Wright, M., Taekman, J., & Endsley, M. (2004). Objective Measures of Situation Awareness in a Simulated Medical Environment. *Quality and Safety in Health Care*, *13*(suppl 1), 65.

Wu, C., & Liu, Y. (2006). Queuing Network Modeling of Driver Workload and Performance. In *50th Annual Conference of the Human Factors and Ergonomics*

Society. October 16-20, 2006, San Francisco, CA.

Yost, B. and North, C. (2006).The Perceptual Scalability of Visualization. *IEEE Transactions on Visualization and Computer Graphics 12*(5), 837–844.

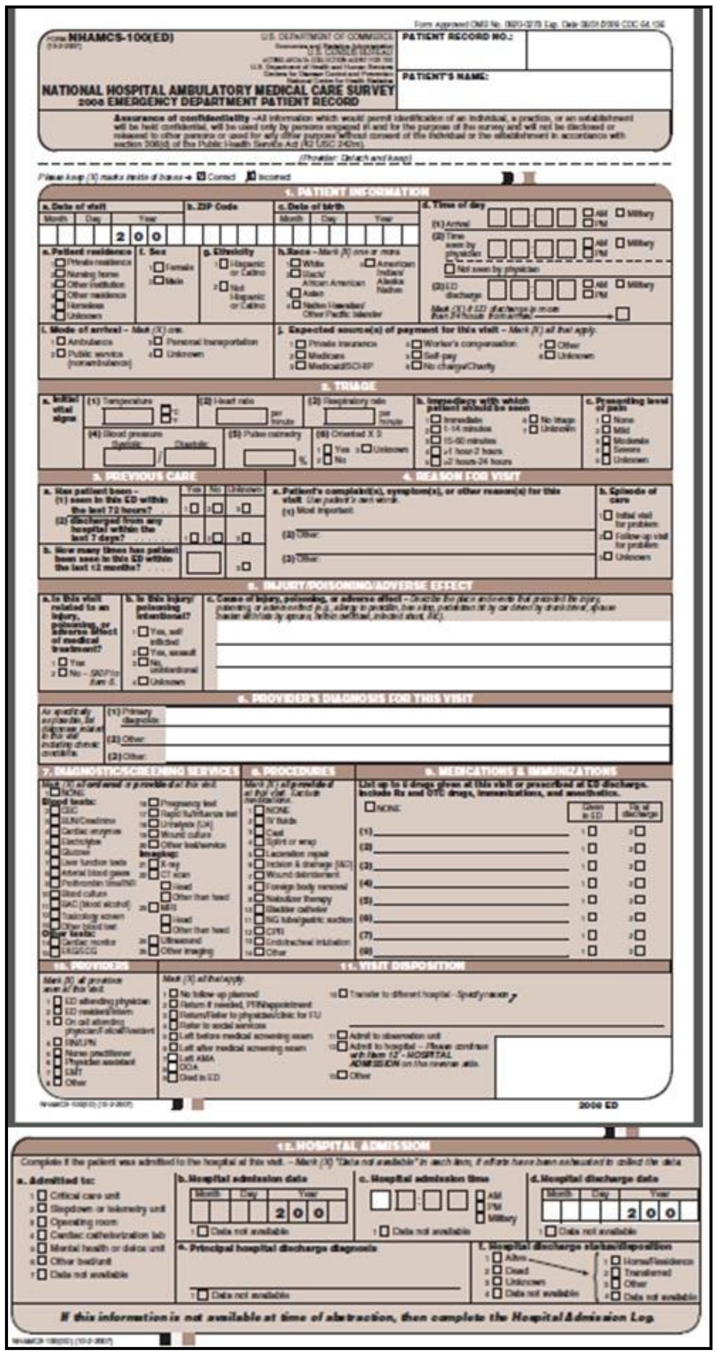
Zhang, H., & Salvendy, G. (2001). The Implications of Visualization Ability and Structure Preview Design for Web Information Search Tasks. *International Journal of Human-Computer Interaction*, *13*(1), 75-95.

Zhou, M. X., & Feiner, S. K. (1998). Visual Task Characterization for Automated Visual Discourse Synthesis. *Proceedings for Conference on Human Computer Interaction*, April 18–23, Los Angeles, CA, 392–399.

Ziegler, J., Kunz, C., & (2002). Matrix Browser: Visualizing and Exploring Large Networked Information Spaces. In *ACM Conference on Human Factors in Computing Systems: Extended Abstracts of the International Conference on Computer Human Interaction*, Minneapolis, MN, 602-603.

APPENDIX A

NHAMCS PATIENT DATA FORM


NHAMCS-100(ED)

NATIONAL HOSPITAL AMBULATORY MEDICAL CARE SURVEY

2008 EMERGENCY DEPARTMENT PATIENT RECORD

PATIENT RECORD NO.:

PATIENT'S NAME:

1. PATIENT INFORMATION

2. TRIAGE

3. PREVIOUS CARE

4. REASON FOR VISIT

5. INJURY/POISONING/ADVERSE EFFECT

6. PROVIDER'S DIAGNOSIS FOR THIS VISIT

7. DIAGNOSTIC/SCREENING SERVICES

8. PROCEDURES

9. MEDICATIONS & IMMUNIZATIONS

10. PROVIDERS

11. VISIT DISPOSITION

2008 ED

12. HOSPITAL ADMISSION

a. Admitted to:

b. Hospital admission date

c. Hospital admission time

d. Hospital discharge date

e. Principal hospital discharge diagnosis

f. Hospital discharge status/disposition

If this information is not available at time of abstraction, then complete the Hospital Admission Log.

APPENDIX B

DEMOGRAPHIC SURVEY

Title: Performance Evaluation of Information Visualizations in Healthcare Information System Interface
Principal Investigator: Quaneisha Jenkins-Penha

Pre-Test Questionnaire

1. What is your age group?
   - ☐ 18-25
   - ☐ 26-35
   - ☐ 36-45
   - ☐ > 45
2. What is your gender? ☐ Female ☐ Male
3. What is your primary hand? ☐ Right hand ☐ Left hand ☐ Ambidextrous
4. Are you color blind? ☐ Yes ☐ No
5. What best describes your expertise area?
   - ☐ Business
   - ☐ Engineering
   - ☐ Education
   - ☐ Health
   - ☐ Technology
   - ☐ Other
6. Do you use email? ☐ Yes ☐ No
7. How frequent do you use a computer and/or internet devices? (5 being most frequent)
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
8. What is your attitude toward learning new technologies? (5 being most favorable)
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
9. Have you ever worked in a hospital? ☐ Yes ☐ No
10. If yes to 9, have you ever worked in the emergency department? ☐ Yes ☐ No
11. If yes to 9, did you work with a hospital information system? ☐ Yes ☐ No
12. If yes to 11, how often did you work with the system? (5 being most frequent)
    ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
13. If yes to 11, what type of tasks did you perform with the system?
    - ☐ Admission
    - ☐ Billing
    - ☐ Communication
    - ☐ Patient Care
    - ☐ Scheduling
    - ☐ Other

Participant #______
Date____________
VA ____/R_____/L

APPENDIX C

USABILITY SURVEY

Title: Performance Evaluation of Information Visualizations in Healthcare Information System Interface
Principal Investigator: Quaneisha Jenkins-Penha

Post-Test Questionnaire

1. Please give the visualization a usability score from 1 to 5 (1 least favorable to 5 most favorable)

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

2. Was it easy using the visualization to answer the question? (1 not easy at all to 5 very easy)

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

3. Was it easy to view the variables in the visualization? (1 not easy at all to 5 very easy)

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

4. Would you consider this visualization a valuable tool for healthcare data? (1 not valuable at all to 5 very valuable)

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Participant #______
Date____________
VA ____/R_____/L

APPENDIX D

NASA-TLX MENTAL WORKLOAD SURVEY

Title: Performance Evaluation of Information Visualizations in Healthcare Information System Interface
Principal Investigator: Quaneisha Jenkins-Penha

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Mental Demand
How mentally demanding was the visualization?

Low High

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Physical Demand
How physically demanding was the visualization?

Low High

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Temporal Demand
How much time pressure did you feel due to the rate or pace?

Low High

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Performance
How successful were you in accomplishing what you were asked to do?

Failure Perfect

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Effort
How hard did you have to work to accomplish your level of performance?

Low High

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Frustration
How insecure, discouraged, irritated, stressed, and annoyed were you?

Low High

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐6 ☐ 7

Participant #_______
Date____________
Visualization______

APPENDIX E

RESULTS OF EMPIRICAL STUDY (PART 1)

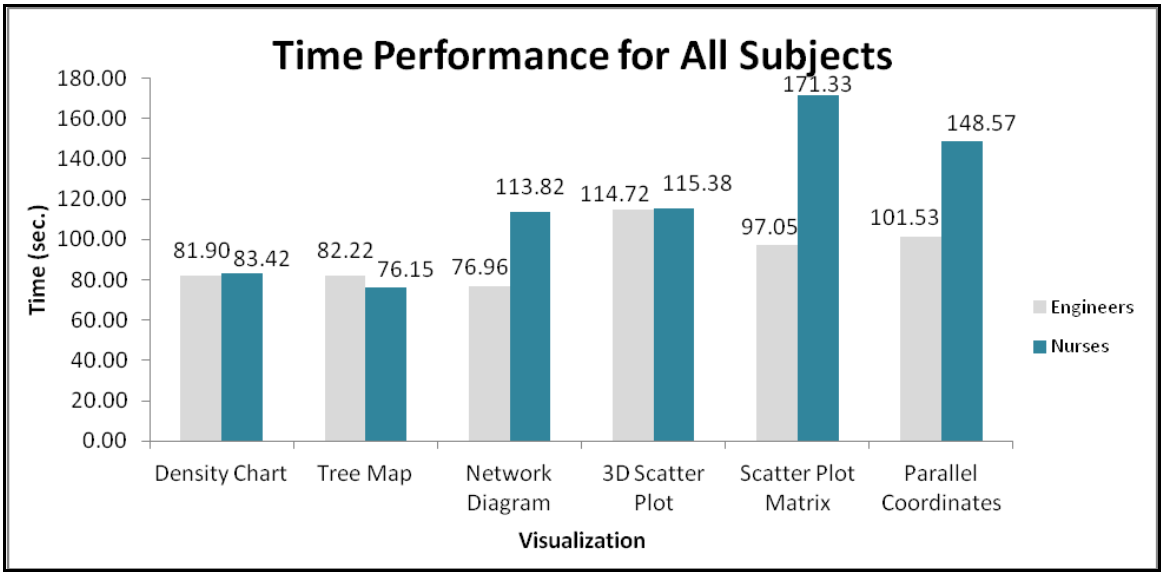


Figure E.1: Completion Time for Nurses and Engineers

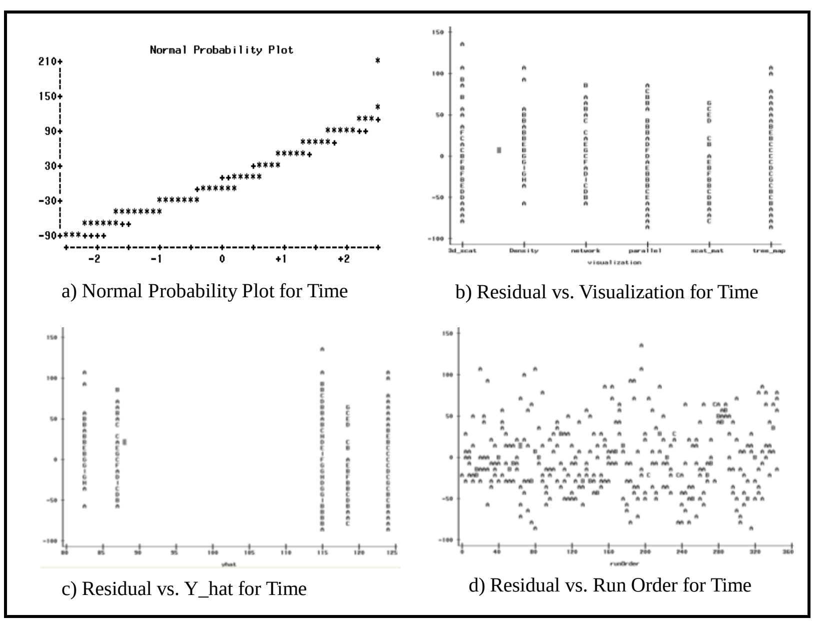


Figure E.2: Normal and Residual Plots for Time, Empirical Study Part 1

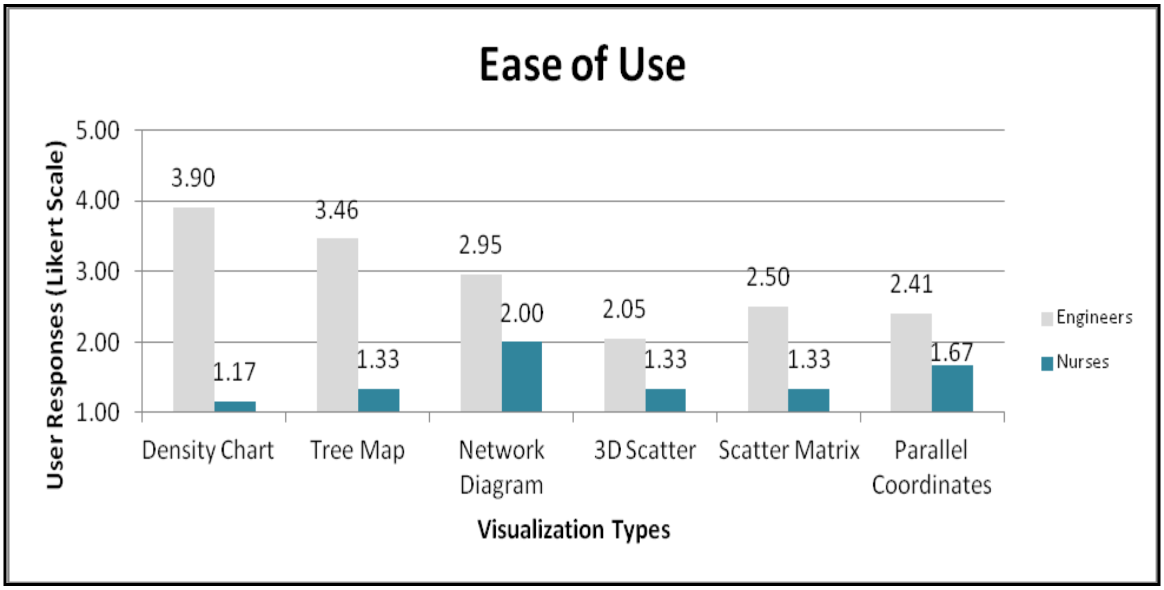


Figure E.3: Means for Ease of Use

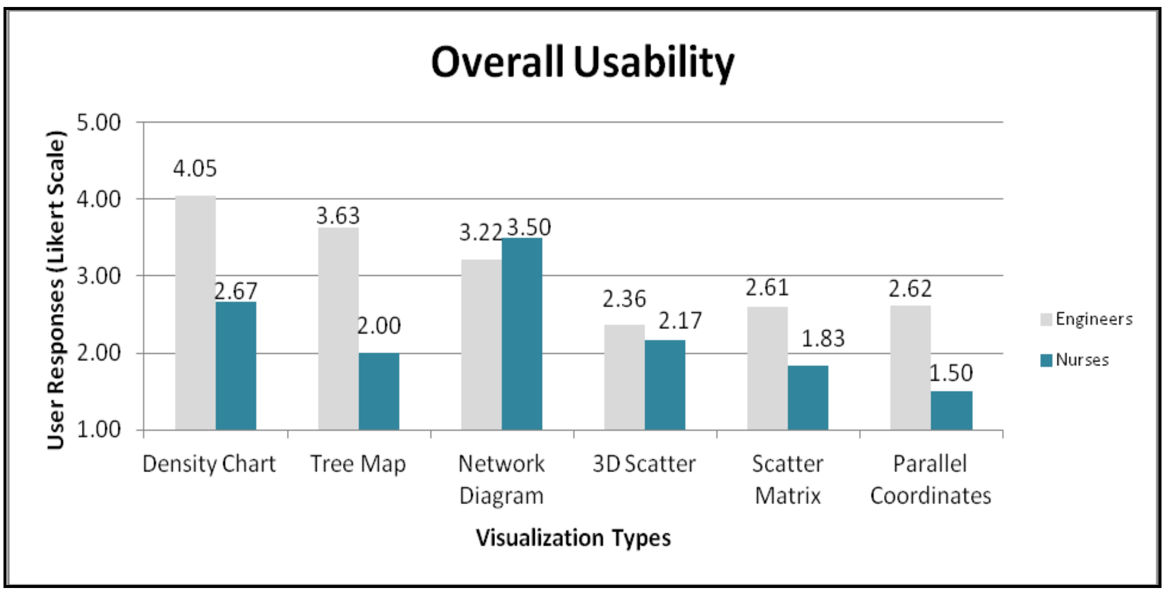


Figure E.4: Means for Overall Usability

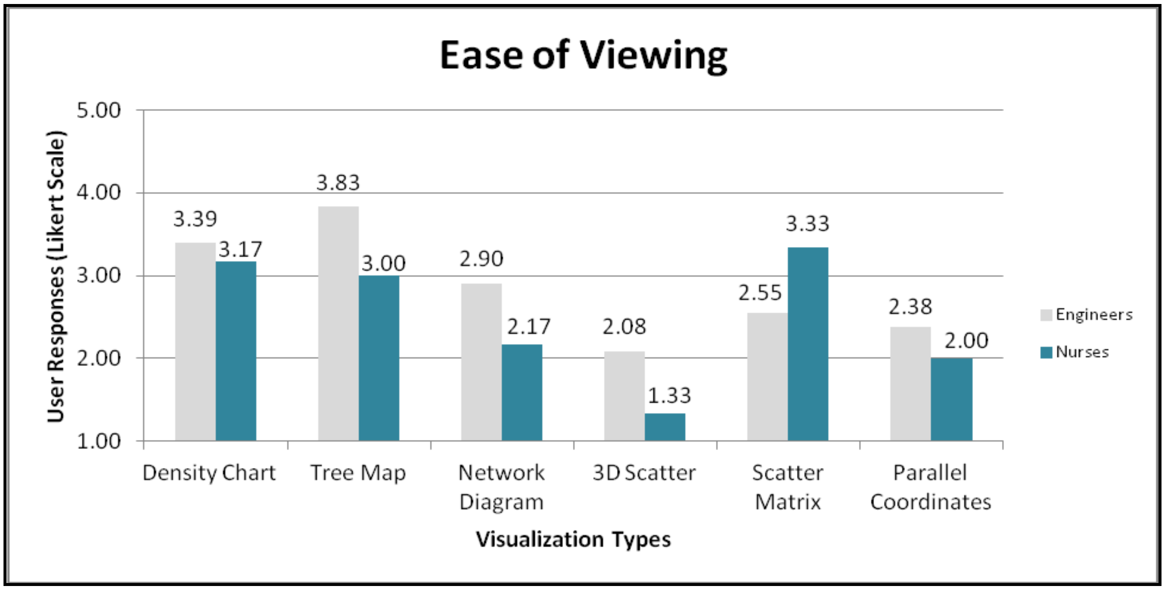


Figure E.5: Means for Ease of Viewing

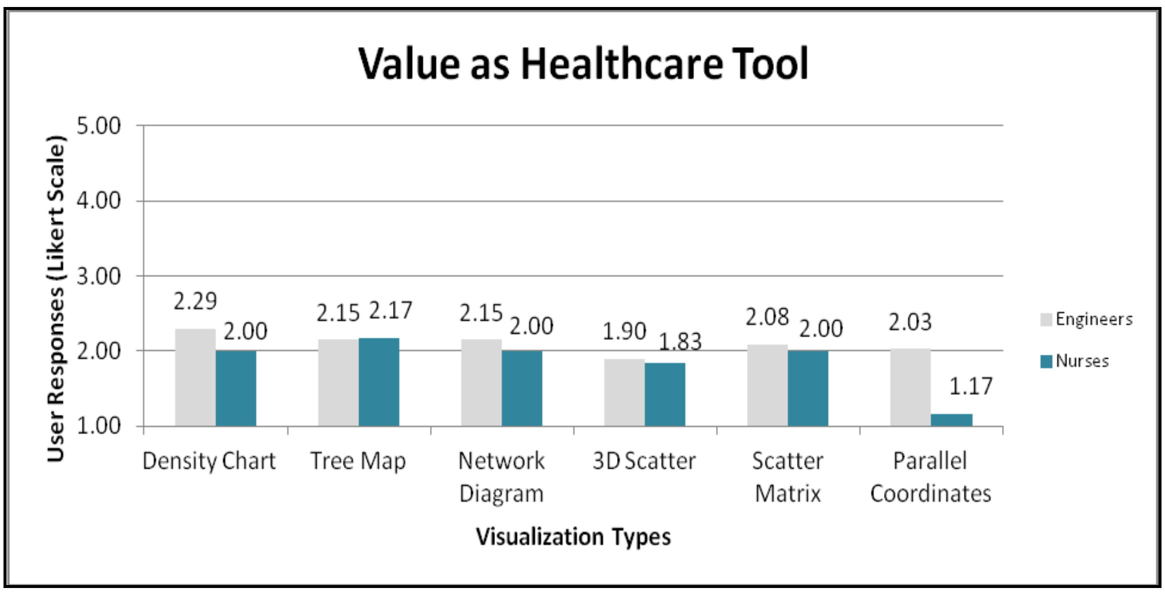


Figure E.6: Means for Value as Healthcare Tool

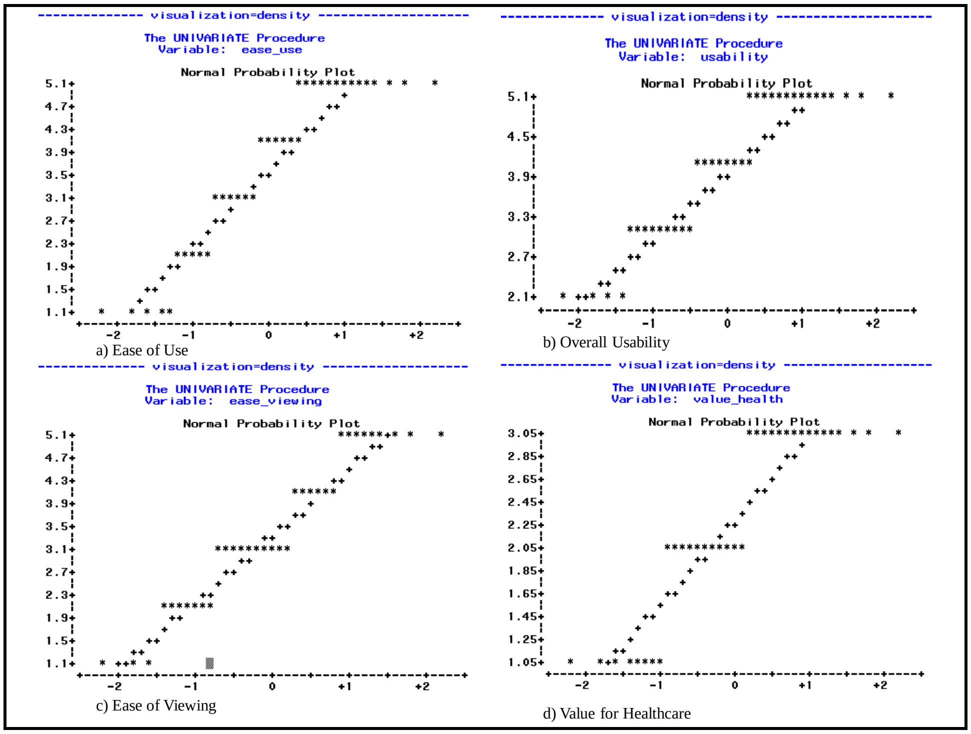


Figure E.7: Normality Plots for Overall Usability

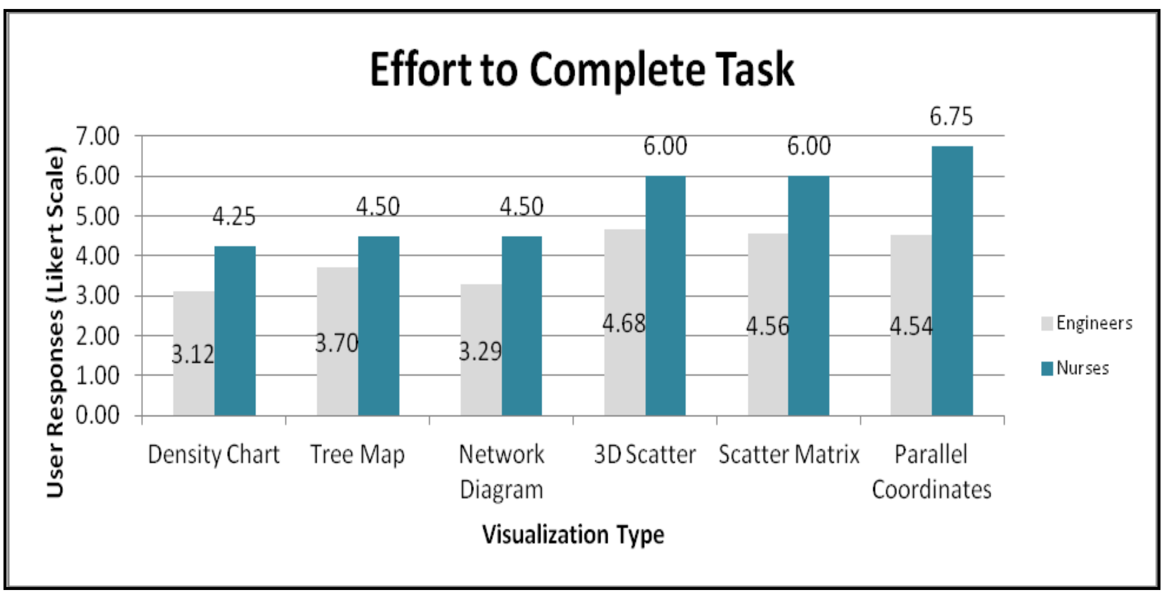


Figure E.8: NASA TLX Sub-Scale for Effort

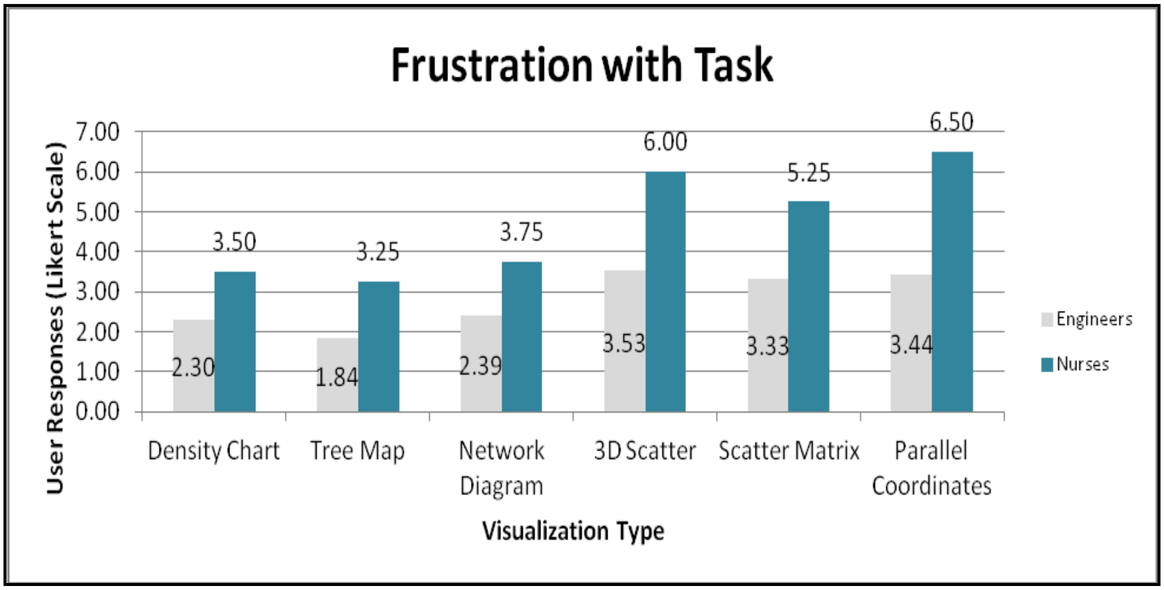


Figure E.9: NASA TLX Sub-Scale for Frustration

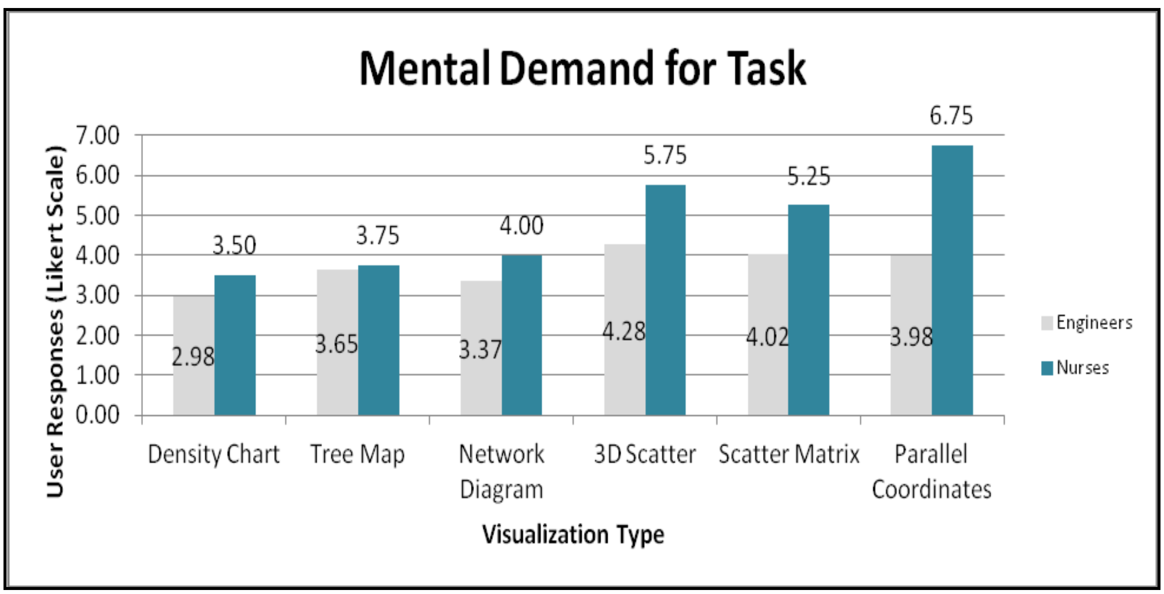


Figure E.10: NASA TLX Sub-Scale for Mental Demand

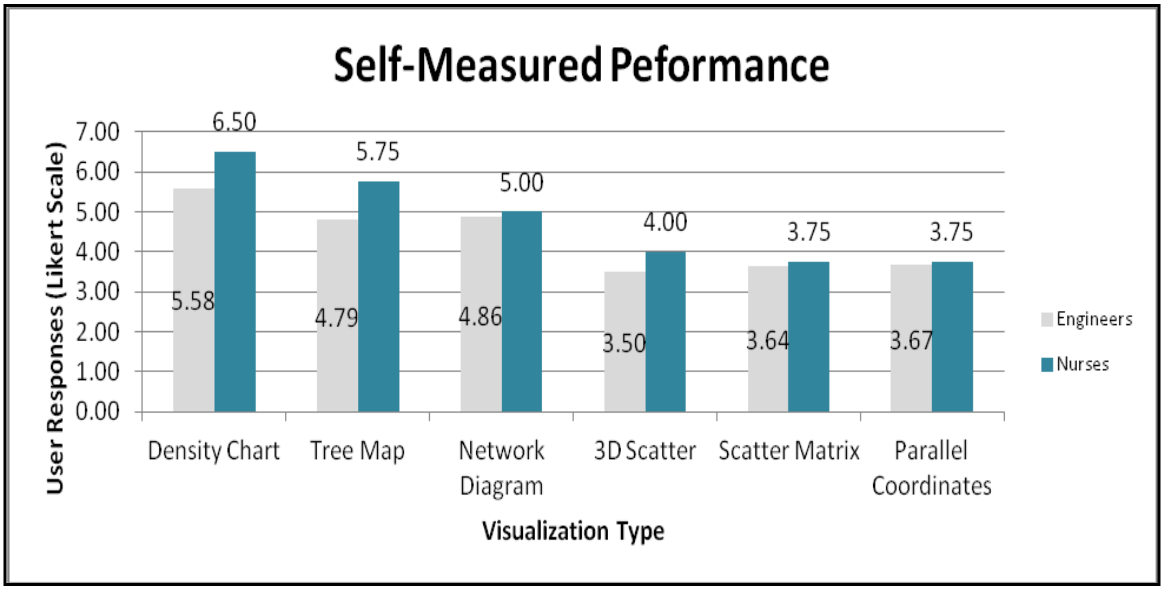


Figure E.11: NASA TLX Sub-Scale for Performance

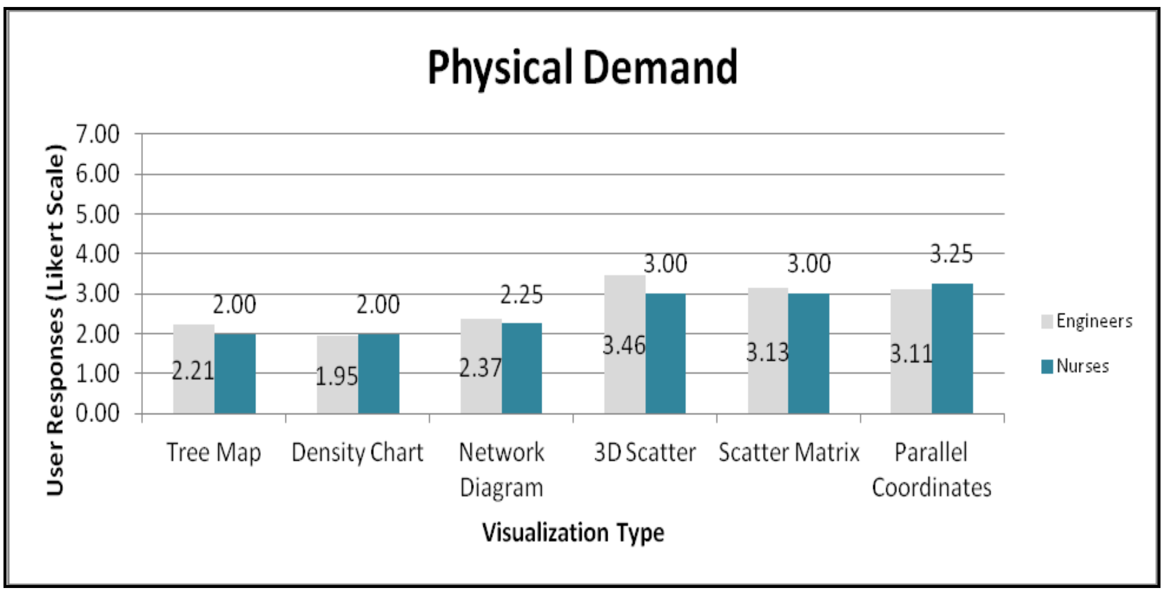


Figure E.12: NASA TLX Sub-Scale for Physical Demand

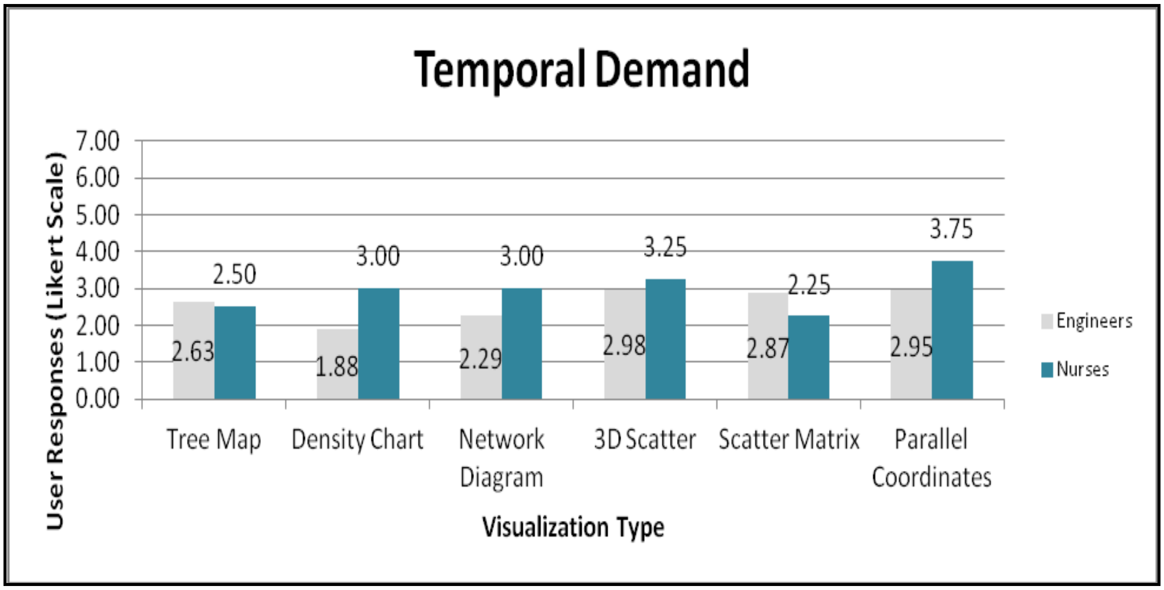


Figure E.13: NASA TLX Sub-Scale for Temporal Demand

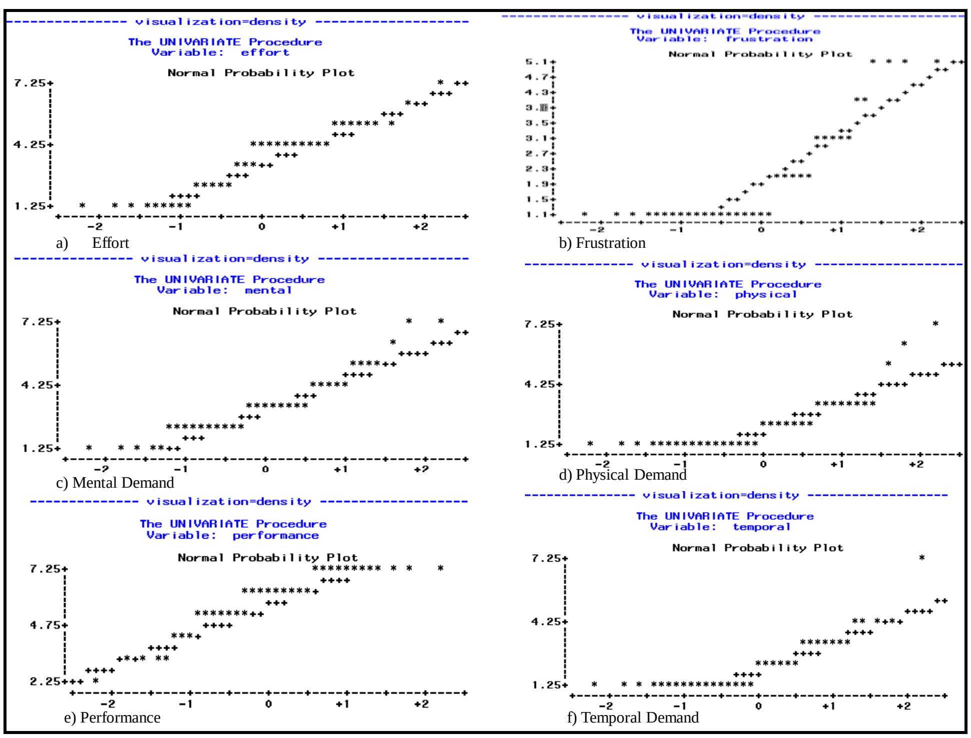
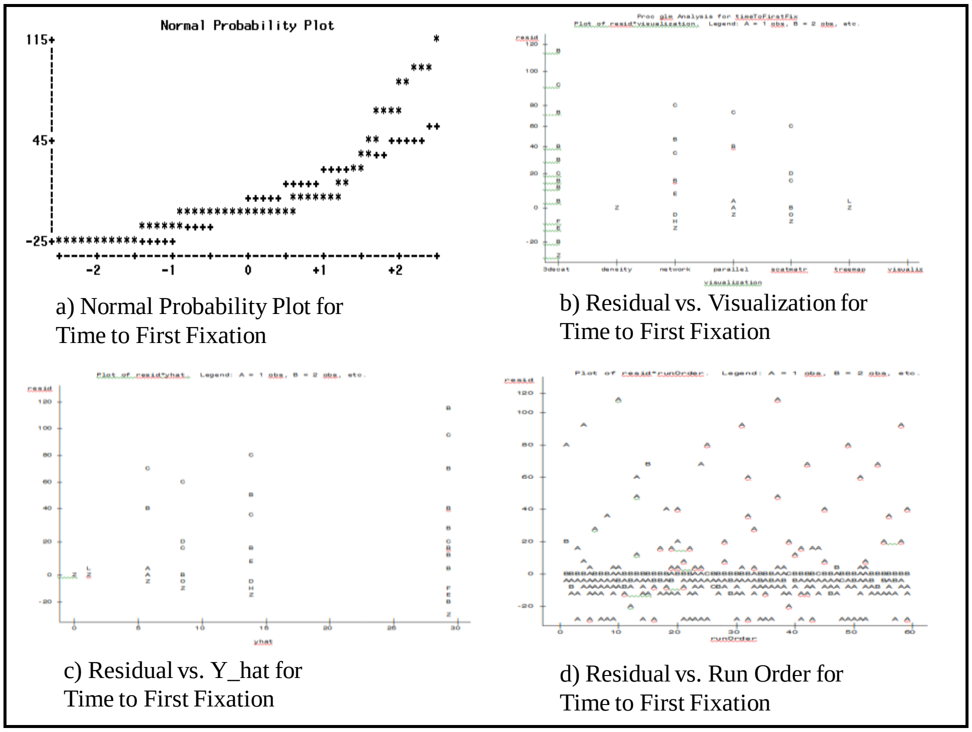


Figure E.14: Normality Plots for Workload Survey Items



a) Normal Probability Plot for Time to First Fixation

b) Residual vs. Visualization for Time to First Fixation

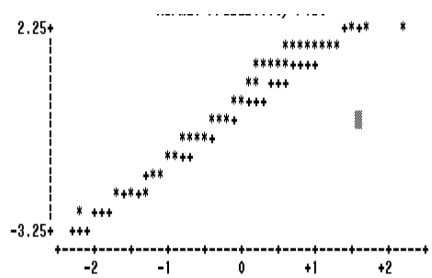
c) Residual vs. Y_hat for Time to First Fixation

d) Residual vs. Run Order for Time to First Fixation

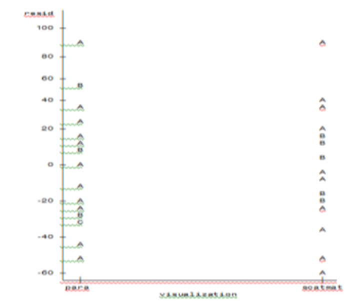
Figure E.15: Normal and Residual Plots for Time to First Fixation

APPENDIX F

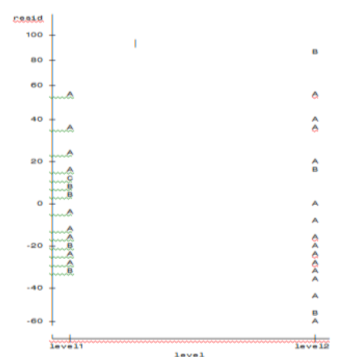
NORMAL AND RESIDUAL PLOTS FOR TIME IN EMPIRICAL STUDY (PART 2)



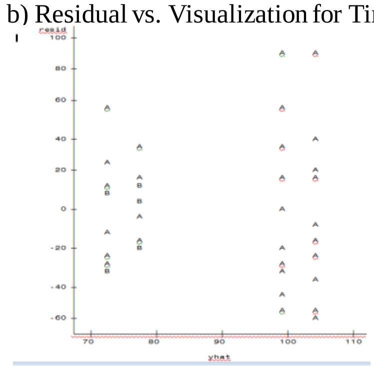
a) Normal Probability Plot for Time



b) Residual vs. Visualization for Time



c) Residual vs. Level for Time



d) Residual vs. y_hat

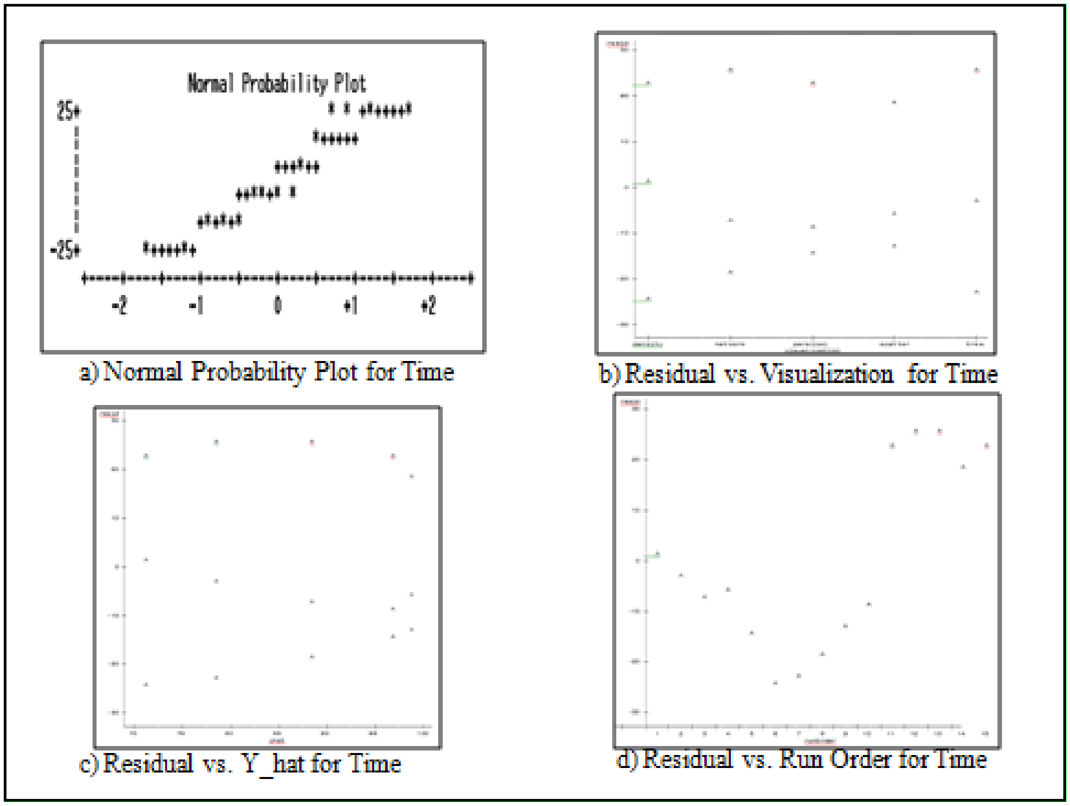
APPENDIX G

FOCUS GROUP TRANSCRIPT SUMMARY

No.	Visualization	Comment Favor	Theme	Comments
1	Density Chart	positive	visual comprehension	quick understanding, easy to learn
2	Tree Map	positive	compatibility, visual comprehension	difference in box sizes
3	Network Diagram	positive	visual comprehension	easy understanding
4	Scatter Plot Matrix	positive	visual comprehension	quick overarching glimpse
5	ALL	positive	visual comprehension	can see a lot of records at once, decrease search time
6	Tree Map	neutral	compatibility	indistinguishable colors
7	Tree Map	neutral	interaction	needs a pop up or way to see details of each boxes
8	Network Diagram	neutral	tech support	add legend, color code groups of symptoms
9	Network Diagram	neutral	interaction	add highlighting
10	ALL	neutral	other	Images best as add-ons to reports
11	Scatter Plot Matrix	neutral	visual comprehension	prefer to see ALL boxes, not just the bottom triangle
12	Parallel Coordinates	neutral	interaction	changing order of axes, changes how image looks
13	Tree Map	negative	compatibility, visual comprehension	difference in box sizes
14	Network Diagram	negative	visual comprehension	some words block out other words
15	Scatter Plot Matrix	negative	visual comprehension	axes crowd each other
16	Parallel Coordinates	negative	tech support	confusing axes, locations are poor
17	Parallel Coordinates	negative	compatibility	cluttered, the coloring makes it worse
18	3D scatter	negative	tech support	axes are in bad locations, towards the edge of the screen
19	3D scatter	negative	visual comprehension	cluttered
20	Density Chart	negative	visual comprehension	cluttered, words block other words

APPENDIX H

NORMALITY AND RESIDUAL PLOTS FOR TIME IN CASE STUDY



a) Normal Probability Plot for Time

b) Residual vs. Visualization for Time

c) Residual vs. Y_hat for Time

d) Residual vs. Run Order for Time