Julia Kulla-Mader. Graphs via Ink: Understanding How the Amount of Non-data Ink in a Graph Affects Perception and Learning. A Master's Paper for the M.S. in I.S degree. April, 2007. 64 pages. Advisor: Deborah Barreau

There is much debate in the design community concerning how to make an easy-to-understand graph. While expert designers recommend including as little non-data ink as possible, there is little empirical evidence to support their arguments. Non-data ink refers to any ink on a graph that is not required to display the graph's data. As a result of the lack of strong evidence concerning how to design graphs, there is widespread confusion when it comes to best practices. This paper describes a preliminary study of graph perception and learning using an eye-tracking system at UNC's School of Information and Library Science.

Headings:

eye tracking

graph perception

data-ink ratio

graph design

GRAPHS VIA INK: UNDERSTANDING HOW THE AMOUNT OF NON-DATA INK IN A GRAPH AFFECTS PERCEPTION AND LEARNING

by
Julia Kulla-Mader

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

April 2007

Approved by

____

Deborah Barreau

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION

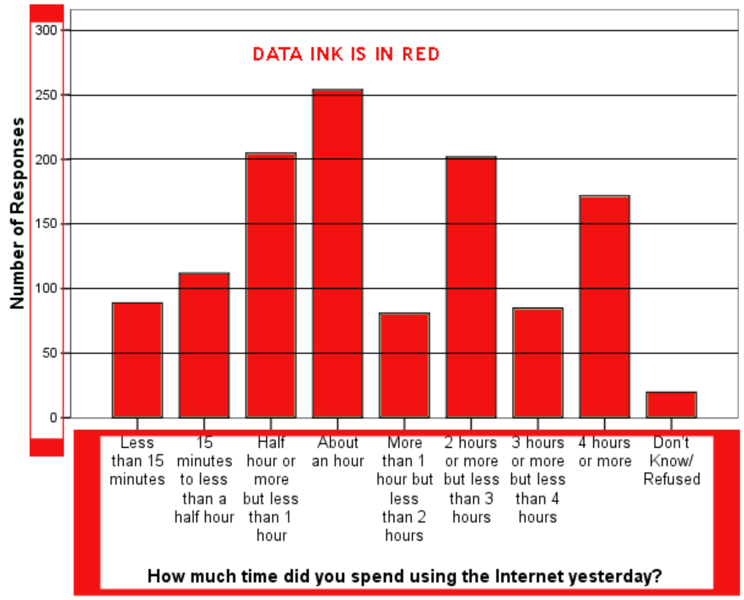
One of the unique problems of designing for the web is determining how to best integrate graphs. It has been established that users are more likely to scan text on a web page than read it – 79 percent of web users scan rather than read (Nielsen, 1997). There are some guidelines in place for writing web text that lends itself to reading scanning – highlighting keywords, having meaningful subheadings, using particular fonts, using bullets, and having one idea per paragraph (Nielsen, 1997). However, there are no guidelines in place for creating graphs that lend themselves to reading scanning.

This study aims to test how well one potential method of making online graphs more scannable – having a low data-ink ratio — affects graphical comprehension online. In other words, how does changing the data-ink level of a graph influence user perceptions of online graphs?

Tufte, a noted design theorist, defines data ink as “the non-erasable core of a graphic.” The data ink is all of the ink used to make the graphic that cannot be removed without changing the meaning of the graph itself. The data values included in bar graphs and the bars themselves are examples of data-ink in a bar graph. If someone erased the data values on a graph associated with different bar heights, the graph would no longer convey the necessary quantitative information. As a result, these numbers can be classified as data-ink. In contrast, good examples of non-data ink elements include tick marks, unnecessary use of color, background graphics and grid lines. All of these design features can be erased without changing the meaning of the quantitative information contained in the graph. Graphs generated by PowerPoint and Microsoft often contain a good deal of non-data ink.

Tufte states that when an equal amount of ink is used to draw the data and to depict the entire graph, the graph's quality increases (Tufte, 1983). In other words, an ideal, high quality graph contains only essential data – it is impossible to erase any part of the graph without removing actual data.

Figure 1



$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{total ink used to print the graphic}}$$

The purpose of this study is to test Tufte's data-ink ratio theory using eye-tracking technology. Although Tufte's concept of the data-ink ratio is widely accepted, a 1989 study found evidence that the data-ink ratio may not have any effect on the accuracy of graph understanding (Kelly, 1989). However, a later study found that there were perceptive differences in accuracy when subjects interpreted graphs with lower data-ink ratios in comparison to graphs with higher ratios (Gillian, 1994). There is a gap in the literature concerning online graph comprehension and a lack of consensus surrounding Tufte's data-ink theory.

Although there is not a wide body of literature focused on how the data-ink ratio influences graphical comprehension, there have been many studies investigating how people process information. The roots of current research in graphical perception can be found in the field of Human-Computer Interaction. Specifically, current research uses the Model Human Processor as a basis for asking questions about graphical perception. The Model Human Processor is a simplified model of human cognition developed to help researchers predict human behavior in response to stimuli. It is a high-level, oversimplified view of human information-processing (Card, 23).

Card's Model Human Processor presents three systems that affect how one interacts with a graph: the perceptual system, the motor system and the cognitive system. The perceptual system includes sensors and "buffer memories" (Card, 24). Buffer memories relate to physical representatives of images such as the brightness of the colors used. Buffer memories are the representations of images in the viewer's mind without regard for meaning. For instance, when someone first glances at a graph, she uses her perceptual system to figure out that it is indeed a graph that they are viewing. At the perceptual stage, someone may be discerning the length, color, shape, and position of the objects that make up a graph and storing this information as buffer memories.

Using the cognitive system, a person looks at a picture, stores it in her working memory, and then makes connections between what she is looking at and what she has seen before (Card 24). Working memory contains the ideas and concepts temporarily stored in one's mind while analyzing information. One's previous experiences in the form of long-term memory are stored in the cognitive system. The motor system is where someone carries out a response. So, an artist may look at a picture and perceive each line using her perceptual system. Then, she could make connections between what she is viewing and images that she has seen before by accessing her cognitive system. Finally, she might say out loud, “That's a Picasso!” using her motor system.

In addition to the Model Human Processor, much of the theory surrounding what makes a usable graph comes from design experts whose decisions are based on their instincts and experiences rather than experimental results. Kosslyn evaluated five landmark graph design books, including Tufte's, and looked at how well their recommendations match up with research findings on graphical perception in a research review (Kosslyn, 1985). Although he did not study graphical perception empirically, his review is still relevant as a basis for understanding how to evaluate graphs according to the Model Human Processor. Kosslyn presents concrete ways to evaluate how well graphs use what is known about the perceptual system, the cognitive system, and long-term memory.

Kosslyn states that when looking at a graph the perceptual system is affected by adequate discriminatibility, visual properties, processing priorities, and perceptual distortion. Discriminability refers to the fact that variations in lines on a graph must be noticeable for the viewer to discern meaning. Visual properties include the size of elements on a graph, darkness, texture, color, and related aspects of visual presentation. Processing priorities are the order in which one looks at elements on a graph. For instance, someone is more likely to first look at dark lines than at light ones. Perceptual errors refer to people's tendency to misjudge how they see patterns. Subjects often do not perceive equal sized objects on a graph as equal. Kosslyn also gives specific examples of factors that influence the cognitive system and short-term memory. He points out that short-term memory is organized into units and that only a few units can be in short-term memory at a time. As a result, he has found that grouping objects together on a graph helps them stay in short-term memory. Yet, the design theory books evaluated lacked sufficient discussion of grouping.

Kosslyn also presents criteria for overcoming the memory-capacity limitations of graphs. Too much information should not be put inside of graph keys or inside of graph displays themselves. He recommends against putting too much information inside of displays or graph keys because this will place an unnecessary burden on working memory. Although he recommends not putting too much information inside of graph displays, he does not shed any light on how the data-ink ratio affects short-term memory-capacity limitations.

1.1 Research Question

My research question relates directly to the data-ink ratio. As more non-data ink is added to an online graph, what effect does it have on how people read graphs and where their attention is drawn?

The independent variable in this study will be the amount of non-data ink included in an online graph. The dependent variables will be the accuracy with which subjects read an online graph, the attention subjects pay to elements on a graph, and their enjoyment level of interacting with online graphs with different levels of non-data ink.

My hypothesis is that Tufte's theory is partially incorrect. I predict that subjects will prefer graphs with a moderate data-ink ratio and that they will process information more accurately on graphs with a moderate data-ink ratio than a low or high data-ink ratio.

2 LITERATURE REVIEW

This study is concerned with the overall topic of how people draw perceptual inferences and thus learn from graphs. The focus is on what causes errors (incorrect inferences) in graphical perception and whether changing the data-ink ratio affects learning and perception. When someone makes a graph, they encode information via design decisions such as whether to use a pie chart or box plot. Graphical perception refers to the process of decoding and understanding the qualitative and quantitative information represented in graphs (Cleveland, 1985).

In order to better understand how individuals interact with data on a graph, it is necessary to look at the overall cognitive model of graphic interaction with a particular focus on what causes breakdowns in understanding. Using these cognitive models, we can then look at how people's previous experience with graphs and their cognitive understanding of the data being displayed affect their understanding. Lastly, it is relevant to look at existing research and design principles related to how graph design affects memory, cognitive overload, and graph processing speed and accuracy.

2.1 Cognitive Models

Research into graph perception divides interactions with visual information into three distinct systems. Using their perceptual system, individuals glance at a graph (Card, 1983). At the perceptual stage, as mentioned previously, someone may be discerning the length, color, shape, and position of the objects that make up a graph. For instance, using the perceptual system, someone could perceive that lines on a graph are positioned on a common scale (Cleveland, 1985). The perceptual system depends on people being able to identify visual features (Shah, 2002). For instance, if someone was unaware that a continuous line segment is a representation of a line, then their understanding of graphs would be limited by their perceptual system. Additionally, if someone has severely-impaired vision and cannot see a graph, their understanding would be limited by the perceptual system.

Although breakdowns in graph perception are possible at the perceptual level, most breakdowns occur at the motor system level. Individuals use their motor system to understand that they are viewing a graphical representation of quantitative information (Card, 1983; Kosslyn, 1989). Using the motor system, a person looks at a picture, stores it in their short-term memory, and then makes connections between what they are looking at and what they have seen before (Card, 1983). The sum of an individual's previous experience and knowledge is stored in long-term memory. Chunks of long-term memory are activated in the motor stage and become part of working memory (Card, 1983). Working memory is a bottleneck in the processing of graphical information because more chunks of information often go into working memory than can be effectively processed (Lohse, 1997). As a result, this second level is the most important when it comes to understanding how individuals interact with graphs and what information processing errors they make.

2.2 Graph Perception Breakdowns

There are a number of factors that influence whether information is able to effectively get from the perceptual system to short-term memory. There are also a number of factors that influence how short-term memory interacts with long-term memory. With respect to understanding graphs, breakdowns at either side of this relationship can lead to misinterpretation. An additional element that effects both sides of this relationship is capacity. Since people can only hold so much information in their short-term memory at a time, how much information they are required to hold onto can influence how effectively they interact with graphs (Kosslyn, 1989).

2.2.1 Breakdowns Between the Perceptual and Motor Systems

Kosslyn expanded on four types of breakdowns in the transfer on information from the perceptual system from graphs to short-term memory that were initially identified by Spoehr and Lehmkukle (Kosslyn, 1989). The four areas of breakdown relate to discriminability of elements on a graph, size distortions, organizational groupings, and attention focus.

Whether someone is able to discriminate between marks in a graph influences whether information is able to make it to short-term memory. If text is too small or the marks on a graph denoting differences in values are unclear, it is impossible to process the graphical information (Kosslyn, 1989).

Since humans naturally experience perception distortions when evaluating similarly sized objects, graphs must account for this by not requiring people to make visual judgments that they are likely to make incorrectly. The most commonly referenced example of perceptual distortion relates to the size of objects. People tend to distort the size of equally sized objects. For instance, if three equally sized squares are displayed in a row, subjects often do not perceive them to be of equal size. In fact, some researchers recommend making objects used in size comparisons an unequal size appropriate to how they are likely to be perceived (Teghtsoonian, 1965). Related to this, when size judgments are made, it is better to convey information horizontally than vertically because people have difficulty detecting vertical spacing (Cleveland, 1985).

To account for perception distortions, Cleveland and McGill developed a hierarchy of graphical perception tasks from most to least accurate based on experimental findings and existing theories. In their classification, the most accurate graphical perceptual task is perceiving object position along a common scale. The second most accurate task is perceiving position along identical but non-aligned scales. Accounting for the middle rankings, in order of decreasing accuracy, were perceiving length, angle, slope, area, volume, and density. Humans are least able to accurately perceive color hue (Cleveland, 1985).

People naturally organize information on a graph as they perceive it. More specifically, individuals will make natural assumptions about the implications of a graph based on where marks are placed. Marks that are close together will be grouped together in a subject's mind. In addition, objects on a graph that look the same will be grouped together in one's mind, as will any marks on a graph that look like they could comprise a continuous line (Kosslyn, 1989).

Wickens' proximity compatibility principle states that two objects will be perceived as similar when they are spatially closer together, are connected or enclosed by line segments, or have the same color or physical dimensions such as length (Wickens, 1995). When parts of a graph are perceived as similar due to spatial proximity, color, size similarity, or another factor, it is easier for the viewer to make comparisons. This is because there is less of a burden on the viewer to keep track of comparisons in their working memory (Wickens, 1995).

When studying graphical perception, it is important to recognize that people naturally pay attention to some things more than others. When someone looks at a graph, they are more likely to pay attention to bright colors and darker marks (Kosslyn, 1989). If a graph is designed to give attention to unimportant details through the use of color or mark thickness, a viewer might have a hard time interpreting the information it contains. Young children are especially prone to incorrectly interpreting graph results when the graphs contain unnecessary attention focusing elements (Shah, 1992).

2.2.2 Breakdowns Between the Motor and Cognitive Systems

Breakdowns between the motor and cognitive systems are commonly referred to as encoding errors. The previously discussed problems — discriminability of elements on a graph, perceptual distortions, organizational groupings, attention focus — can carry over when information is encoded into long-term memory. For instance, if a graph has illogical groupings, then someone will be more likely to draw incorrect inferences that are stored in long-term memory (Shah, 1992). However, difficulties in graph perception are not only related to the characteristics of the graph. It is also relevant to evaluate who is viewing a graph and what role their previous experiences have in determining how they interact with graphical information. For instance, an elementary school student and a college economics professor will likely have very different reactions when looking at a graph of the United States GNP in the last decade. Three distinct factors influence how an individual interacts with a graph: previous experiences with graphs, conceptual understanding of the topic of the graph, and cognitive overload level.

2.2.3 Previous Experiences With Graphs

One area of breakdown during encoding relates to how people are able to connect the information contained in a graph to what they already know about how to process graphical information (Shah, 1992). Someone's previous experience with graphs influences their interpretations and how they formulate their “graph schema” (Shah, 1992; Pinker, 1990). For instance, if someone sees that two lines intersect on a line graph, they will assume equality because of previous graph-interpreting experiences (Shah, 1995). In addition, one's previous experience with a particular type of graph will influence how they process the information it contains. In a study of the speed and accuracy of graphical reasoning depending on the type of graph and task, researchers found that subjects expect different types of information based on the type of graph presented (Simkin, 1987).

Furthermore, when looking at a graph does not result in an automatic interpretation, it is harder for an individual to make inferences. When explaining the importance of being able to make inferences, Larkin presents the example of a chess board. Although a novice chess player and a chess master both visually see the same board, they are able to make very different inferences about the relationship between the chess pieces and squares (Larkin, 1995). Yet, being able to see a relationships on a graph accurately does not ensure comprehension. For instance, in one study, student subjects were asked to draw the relationship between lines on a graph after looking at a line graph. Although they were able to very closely represent the actual graph, they were at a loss when asked to explain the meaning of what they had just drawn (Shah, 1995).

Even when subjects have an idea about how to interpret graphs, they may use the incorrect inference for the given situation. As an example, people are also naturally drawn to any data points included on graphs and assume importance (Shah, 1995). As a result, when someone sees a line graph, they assume that they are looking for a relationship between the X and Y variables and are likely to make judgments about the steepness of the lines (Shah, 1995).

2.3 Conceptual Understanding of a Graph's Content

Another area for confusion during encoding relates to whether someone can conceptualize what is being quantified (Shah, 1992). Early math education research found that children cannot understand abstract math concepts until later in their education. However, more recent research has found that understanding abstract math concepts is often related to context. If someone does not understand the context surrounding a graph, they will be unable to process its information. For example, a study found that Brazilian street children were able to perform high-level math operations when working as street vendors but were unable to apply the same conceptual understanding to solving similar problems in a classroom (Nunes, 1993).

In addition, those who are able to conceptualize information often make incorrect inferences given their previous knowledge of the topic. In other words, just because someone is familiar with a topic that is discussed in a graph, it does not mean that they will accurately interpret the graph. In a landmark study, Lord found that people are more likely to rate research as convincing that supports their prior attitudes and beliefs (Lord, 1979). This finding is called the biased assimilation effect. In a later study, researchers had low and high-prejudice subjects read two separate studies concerning homosexuality and gender roles. One of the studies that the subjects read concluded that “homosexuality was associated with cross-gender behavior/psychopathology” while the other argued the opposite. The researchers found that subjects were likely to let their preexisting attitudes about homosexuality influence their evaluations of the research (Munro, 1997). The biased assimilation effect provides evidence that even when someone understands a graph, they are likely to judge its quality based on their previously held knowledge and perspective.

2.4 Cognitive Overload

The concept of cognitive overload is related to both subjects' conceptual understanding and previous knowledge when interacting with graphs. In a study of how people interact with different types of line graphs, Shah and Carpenter found that the more effort needed to process graph information, the less the chance that the interpretation is accurate (Shah & Carpenter, 1995). In their experiment, they found that people will be more likely to process information represented on an x-y plane accurately than information on a y-z plane in part because doing so requires less cognitive overhead. Since most people have more experience interpreting graphs with x and y axis than with y and z axis, it is easier for them to understand quantitative information displayed on an x-y plane.

An additional discussion of cognitive overload relates to the question of when it is appropriate to use a graph for display as opposed to a table. One primary assumption in the literature concerning how people perceive information from graphs is that graphs are preferable to tables for complex learning. The decision whether to use a graph or table depends on the amount of information that is required to be stored in working memory. For quick data retrieval tasks, it easier for someone to glean information from a table than from a graph because they can store relevant points in short-term memory. However, for more complex tasks, a table is not ideal because storing all of the gathered information in short-term memory is no longer possible (Lohse, 1991). This inability to store large amounts of information in working memory when looking at a table is an example of cognitive overload.

There is much discussion in the literature concerning when cognitive overload starts affecting how people process information from graphs. In other words, at what points does the amount of information in a graph become overwhelming and how do these individual working memory overload levels affect how one interacts with a graph? Lohse used a reading span test to measure working memory overload levels and then had his subjects who had differing levels of working memory capacity complete tasks of varying complexity. He found that while graphs helped subjects with limited working memory complete tasks, the graphs alone were not enough to help his subjects overcome working memory limitations. In other words, although those with limited working memory capacity were helped by graphs, they were not able to perform as well as subjects without the same limitations (Lohse, 1997).

Lohse's study of working memory capacity and cognitive overload relates to an area of research concerning graph perception and task analysis. The idea is that there is a relationship between the appropriate graph and the task at hand. Since different tasks require different forms of comparison, the type of graph employed will make a difference. For instance, bar charts lead subjects to make comparisons related to the lengths of the bars, while pie charts lead subjects to make proportional assessments. Furthermore, Simkin and Hastie found that subjects are able to make more accurate comparison judgments with bar charts (in comparison to pie charts) and that they are able to make more accurate proportional judgments by using pie charts instead of bar charts (Simkin, 1987).

2.5 The Role of the Data-Ink Ratio

Although there is a great deal of research concerning how people process graphical information and what breakdowns might occur when doing so, there is a lack of research into how graph design affects graphical understanding. In addition, there is a lack of discussion concerning the differing graph presentation needs of the general public and other specialized groups. Wainer(1981) points out that graphs used by the Census and by newspapers have two purposes: to inform and to attract viewer attention. The audience of newspaper graphs is the general public. In contrast, graphs included in scientific journals and other scholarly work assume a captive audience (Wainer, 1981). Despite the fact that scholarly and newspaper graphs have different purposes, the same design guidelines have been applied to both genres.

Graph design is one area where applying the same design criteria may not make sense. While the general public is likely to need more explanation of key principles, scientists are less likely to need clarification. Despite this, the prevailing wisdom among graph designers is that “less is more” for all audiences when it comes to adding information to a display (Wainer, 1984; Tufte, 1983).

Guidelines for how to draw graphs are often based on extrapolated theories from previous research or on advice from graphic design theories. In many of these graphing guideline articles, Tufte's data-ink ratio is referenced. As mentioned previously, Tufte argues that you can tell the quality of a graph via its data-ink ratio, which is the amount of ink used to draw the data divided by the total amount of ink in a graph. Tufte states that when an equal amount of ink is used to draw the data and to depict the entire graph, the graph's quality increases (Tufte, 1983). In other words, an ideal graph contains no inessential data – it is impossible to erase any part of the graph without removing actual data. Although Tufte's concept of the data-ink ratio is widely accepted, a 1989 study found evidence that the data-ink ratio may not have any effect on the accuracy of graph understanding (Kelly, 1989). However, a later study found that there were perceptive differences in accuracy when subjects interpreted graphs with lower data-ink ratios in comparison to graphs with higher ratios. Yet, the same study also found that additional ink can increase the accuracy of interpretation if it is used to illustrate meaning (Gillian, 1994).

Tufte uses the data-ink ratio as evidence when arguing that excessive descriptions on a graph fall within the spectrum of “chartjunk” and get in the way of processing graphical information (Tufte, 1983). Specifically, he warns against “redundant representations” of the simplest data. However, at the same time, he points out that descriptions on a graph can be used to avoid distortions and ambiguity (Tufte, 1983).

Throughout the research into how people process graphs, there has been some discussion of what types of graphs people prefer and how these graphs vary depending on their level of detail. The decision to use a three dimensional or two dimensional graph involves choosing which level of detail to use. This level of detail is synonymous with Levy's concept of gratuitous graphics and Tufte's chartjunk (Levy, 1996). As a result, the decision to use a 2D or 3D graph lends itself to thinking about how people interact with extra graph information. To evaluate preference levels for graphs with different levels of “gratuitous graphics,” one study required its subjects to choose between nine different types of graphs to present in front of a fictitious Board of Directors depending on the scenario. The researchers found evidence that while people prefer simple 2D graphs for analyzing data on their own, they prefer to present more complicated 3D graphs to others, particularly in situations where the task required memory. This research lends credence to the idea that the “less is more” principle may not be true for graphs, although more research is needed to fully answer this question.

2.6 Summary

Although there is a lack of conclusive research into how the data-ink ratio affects graphical perception, there are many opportunities for future research. More research is needed to evaluate whether adding non-data ink to a graph adds to comprehension. Any analysis of this question should take into account the model of graph perception, and breakdowns in transitioning information from the perceptual system to the motor system and also from the motor system to the cognitive system.

3 METHODOLOGY

The goal of this experiment was to test what happens as more non-data ink is added to a graph. What effect does changing the data-ink level have on how people read graphs and where their attention is drawn? To test this question, I observed subjects looking at graphs with varying data-ink levels using an eye-tracking system and evaluated their responses to associated survey questions.

Twelve volunteers participated in the study, which was conducted in the Interaction Design Lab at the School of Information and Library Science at the University of North Carolina at Chapel Hill. The study was conducted in one sitting. Subjects came into the lab, were hooked up to the eye-tracking system, and answered questions via an online questionnaire. The study is described in more details below.

3.1 Sample

Subjects were solicited through the sils-students listserv. Participants included master's and Ph.D. students at UNC's School of Information and Library Science. The actual sample was made up of thirteen students who responded to an e-mail request for subjects, but equipment malfunctions resulted in data collection for twelve of the thirteen subjects.

3.2 Setting

The study took place in the Interaction Design Laboratory on the 4th floor of the library stacks at the School of Information and Library Science. The eye-tracking station was used for this study.

The following equipment is part of the eye-tracking system:

- Command station: Dell Optiplex GX240 Pentium 4 PC running at 1.5GHz, 1GB RAM, ATI RAGE Pro AGP Graphics adapter, 40GB hard disk, 10/100Mb Ethernet, 15" flat panel display, Windows XP. All of the software for the eye-tracking is on this computer. Including:
  - The E500 EYEPOS software tracker
  - EYENAL data analysis software
  - Gazetracker
- Subject station: Dell Pentium II 400MHz PC, 256MB RAM, 9GB hard disk, CDROM drive, internal zip drive, SCSI adapter, ATI 3D Rage Pro 8MB AGP graphics card, 10/100Mb ethernet card, 19" monitor, multimedia sound and speakers. Windows 2000, SILS Lab Software Setup.
- ASL Model 5000 Eye Tracker Control Unit
- Flock of Birds Head tracker electronics unit, magnetic transmitter, head mounted sensor & halo
- Remote scene camera (Costas)
- Pan/Tilt optics module (Sony video camera)
- A pair of black and white video monitors for (Eye monitor and scene monitor)
- TView Gold Video Scan Converter (resolution up to 1024x768 in 32 million colors, replaces remote scene camera)

3.3 Eye-Tracking Environment

The eye-tracker was used in this study to measure pupil dilation, points of fixation, and eye movement. During the study, subjects sat on a wooden chair at the subject computer station and answered questions on the subject computer via an online survey while viewing graphs. All subjects wore a headband with a magnetic head-tracking sensor.

3.4 Procedure

Introduction: When subjects first arrived, they were given an overview of the experiment and asked to sign a consent form. After signing the consent form, they participated in an individual evaluation session. The purpose of the evaluation session was to screen students for use of the eye tracker.

Eye-tracker calibration: After signing the consent form, each subject was then fitted with a small magnetic head tracker, which was attached to an elastic headband. The subject was then seated in front of the eye tracking camera and the system was calibrated to their specific eye size and retina shape (Appendix F). Once the pan/tilt camera has their eye in focus, subjects were asked to look at nine points on different points on the subject computer monitor so that their left eye could be calibrated to the eye-tracker system. In order for the eye-tracker to work, the eye-tracker needs to be calibrated to measure both pupil dilation and corneal refraction. The process of calibrating the eye-tracker for both pupil dilation and corneal refraction took up to a half hour for each subject.

Some subject's eye types are incompatible with the eye tracker. In addition, the eye-tracking system has difficulty calibrating the eyes of subjects wearing heavy eyeliner, or other eye-makeup. The subject whose eyes could not be calibrated via the eye-tracker due to these issues was still compensated for her time with a gift certificate. However, no data was gathered in this case.

A pre-session questionnaire: After signing the consent form and successfully having their eyes calibrated, participants began an online demographic questionnaire. All of the questions asked were part of a online survey hosted via the Odum Institute's Qualtrics online survey system (Appendix A). The purpose of the questionnaire is to gather basic demographic information.

Graph Questions: After finishing the questionnaire, subjects clicked the next button to go to a new screen. On the first screen, subjects saw a graph and three accompanying questions designed to evaluate graph comprehension (Appendix C). Answering this set of questions took between one and three minutes. After answering the three questions, the subject clicked the next button again and saw a new graph and three new associated questions. Answering this set of questions took between one and three minutes. Upon completion, the subject clicked the next button again view the final graph with its associated three questions. Answering the last set of questions also took subjects between one and three minutes.

The accompanying questions for each graph remained the same for all subjects throughout the experiment. Participants were assigned the same three graphs, but in a random order. Each subject spent between three and 10 minutes viewing the graphs and answering the associated questions. The purpose of this exercise was to test graph comprehension.

A follow-up questionnaire: After viewing three graphs and answering the associated questions, subjects clicked the next button and answered a follow-up questionnaire. The follow-up questionnaire asked subjects to identify their preferred level of data-ink and also asked them to account for any other factors that may be influencing their results (Appendix D).

Debriefing: Last, subjects were given the opportunity to hear more about the rationale for the study and to ask questions.

3.5 Graphs and Questions

Three separate graphs were used in the study. The source data for these graphs is from the Pew Internet & American Life Project's Feb-March 2006 dataset. Each graph was accompanied by a set of three questions. The three graphs had varying levels of data-ink. Graph 1, the graph with the lowest data-ink ratio, was modeled after Tufte's example graph in *The visual display of quantitative information* (Tufte, 2001). Graph 1 has no outside border and the gridlines are in white. Graph 2, which has a medium data-ink ratio, has black gridlines but no outside border. Graph 3, which has the most data-ink, has both black gridlines and a black border (Appendix C).

4 RESULTS

The following results describe data from 12 subjects. I am analyzing data from multiple sources, including: a demographic questionnaire, a survey, a follow-up questionnaire, and eye-tracking data.

4.1 Demographics

General Information

All subjects were age 34 or younger. Eight of the 12 subjects were between 26 and 34 and the remaining four subjects were between 18 and 25. Out of the 12 subjects, nine were female. Nine of the 12 subjects identified “4-Year College Degree” as the highest degree they have completed. The remaining three subjects all indicated that they have completed master's degrees. The subjects had a wide array of undergraduate degrees. Nine of the twelve students majored in humanities subjects and the remaining students were social science or science majors. The majority of subjects, 7, were working towards their master's degree in library science. Three students were master's of information science students and one students was pursuing a Ph.D. An additional student chose not to indicate his or her intended degree.

Online Behavior

All of the subjects indicated that they go online daily. Responses varied widely to the question of, "How often do you look at graphs online?" Two subjects said that they look at graphs online every two months, three subjects said that they look at graphs online every two weeks, four subjects said that they look at graphs online once a week, two subjects said that they look at graphs online a few times a week, and two subjects said that they look at graphs online daily.

The majority of subjects identified themselves as having expertise with computers. However, when it came to expertise with graphs, nine out of twelve graphs rated themselves as having neutral or below average experience.

4.2 Responses

Graph Questions

The subjects answered almost all of the questions associated with each graph correctly. The only question that was not answered with 100 percent accuracy was connected to the graph with the lowest data-ink ratio. One person answered the following question incorrectly, " True or False: More people access the Internet every few weeks than 3-5 days a week."

Post-Test Survey

At the end of the survey, subjects were asked to compare the three graphs they had looked at and to also provide feedback on the survey itself.

In their responses, all of the subjects stated that they “completely” understood the questions associated with each graph. Eleven out of the twelve subjects indicated they found the graphs “somewhat” different. One subject said he found the graphs “not at all” different.

In addition, the majority of subjects, 8, stated that they found the graphs “somewhat” similar to graphs they have previously seen online.

When asked to describe how different they found the three graphs, seven of the subjects described how the graphs visually looked similar. One subject noted, “Not very different, considering the layout, color scheme, use of bars.” Subjects did note a difference that was unanticipated. All three graphs were ordered by decreasing frequency on the x-axis. However, while two of the three graphs showed the bars in decreasing size order, the third graph showed bars that were out of order by size. Four of the twelve subjects commented on the out-of-order graph. When asked, “what did you dislike about each graph?” one subject wrote, “in the second graph, I did not like that the bar were not arranged from highest to lowest, or vice versa. It was a little difficult because of the uneven nature of the graph.” Only one student pointed out gridlines as a point of difference on the three graphs. “I noticed that one didn't have black lines at the key value points,” the subject stated.

When asked to describe how similar they found the three graphs, the majority of subjects noted that they found the graphs easy to interpret. One subject noted, “They were all bar graphs, so each was very simple to read. The questions didn't require a very specific answer (the exact number of responses), so the graphs were helpful for answering the questions."

When asked what they liked about the graphs, two subjects mentioned the gridlines, " I liked that there were horizontal lines running across the graph where the main number categories on the y axis were located" said one subject. "Horizontal lines made reading the column heights easier," said another subject. When responding to the question of what graphs they liked best, some subjects again noted the order of the bar heights, "Visually, I liked the graphs that had the items ordered by highest to lowest so the bars descended in a stepwise fashion," one subject responded.

Figure 2

On a scale of 1 to 5, with 5 being the highest, please rate your preference for each graph.						
	1	**2**	**3**	**4**	**5**	**Total**
No Ink		3	2	6	1	12
Medium Ink			3	6	3	12
Lots of Ink		1	1	7	3	12

Note: Graphs were identified by number and not description on the survey

When the subjects rated the graphs on a preference scale of 1 to 5 with 5 being the highest rating, there was a perceived preference for graphs with medium ink and lots of ink. Five subjects assigned a two or three rating to the graph with no ink, including three ratings of two. In contrast, only two subjects assigned a two or three rating to graphs with lots of ink and no subjects assigned a rating below 3 for graphs with medium ink.

Efficiently

There were no major differences in the speed that subjects completed the graph exercise when the order of the graphs was switched.

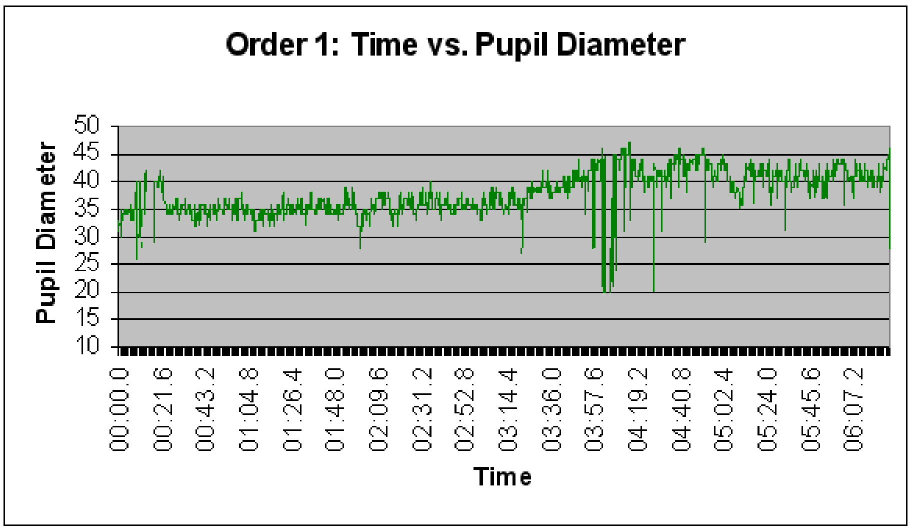
4.3 Eye-Tracking Data

Pupil Dilation

Research has shown that pupil size may increase when people look at content that they view as interesting and decrease when people look at visual input that they view as uninteresting. In addition, there is some evidence that pupil dilation increases are correlated with increases in mental activity (Hess, 1975). A previous study also drew attention to the pattern that pupil dilation temporarily increases and then decreases when the eye-tracking system first starts working (Disabato, 2006).

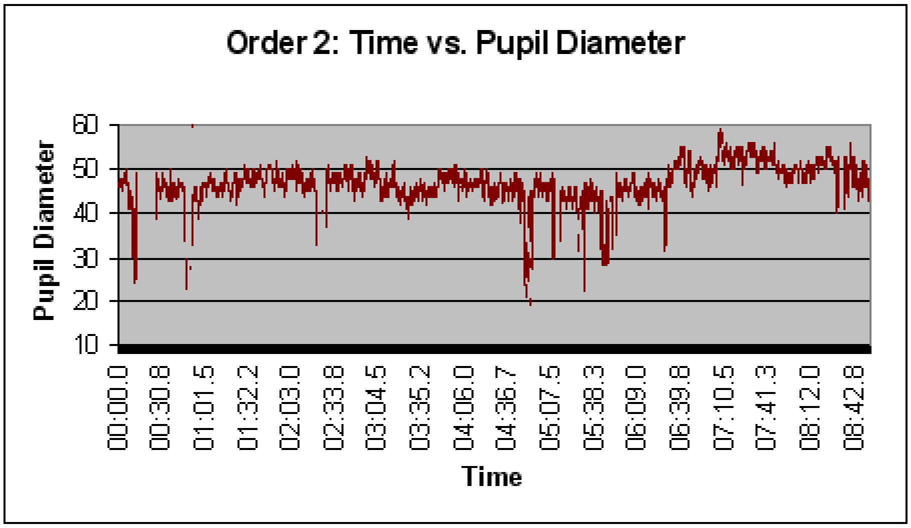
In two of the following three graphs, there is an observable increase in pupil dilation when the recording first starts. The following three sets of pupil dilation levels over time were chosen because they had the fewest missing pupil dilation data points.

Figure 3



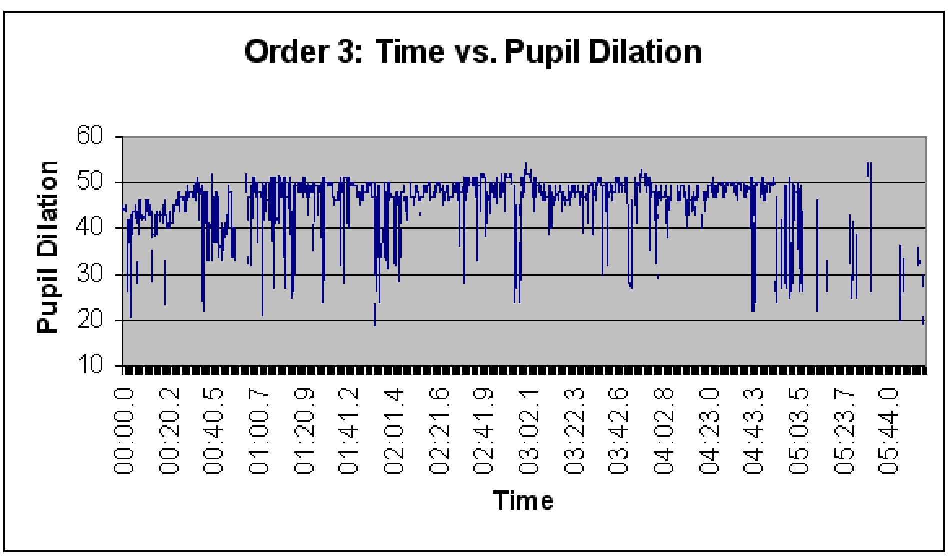
In this graph the order of the graphs presented to the subject was (1)No Ink, (2) Medium Ink, and (3) Lots of Ink, pupil dilation follows the pattern on increasing when the subject first begins the eye-tracking session. There is also an increase in pupil dilation when the subject looks at the third graph, which has the highest data-ink ratio. There appears to be a pattern that pupil dilation increases at the start of an eye-tracking session, then stays relatively constant, and increases for the final third of the experiment.

Figure 4



In the second graph with the graph order (Some Ink, No Ink, Lots of Ink), the same pattern of an pupil diameter increase during the final third of the experiment is observed. Due to missing data points, it is difficult to see if there is an increase in pupil dilation when the subject first puts on the eye-tracking system.

Figure 5

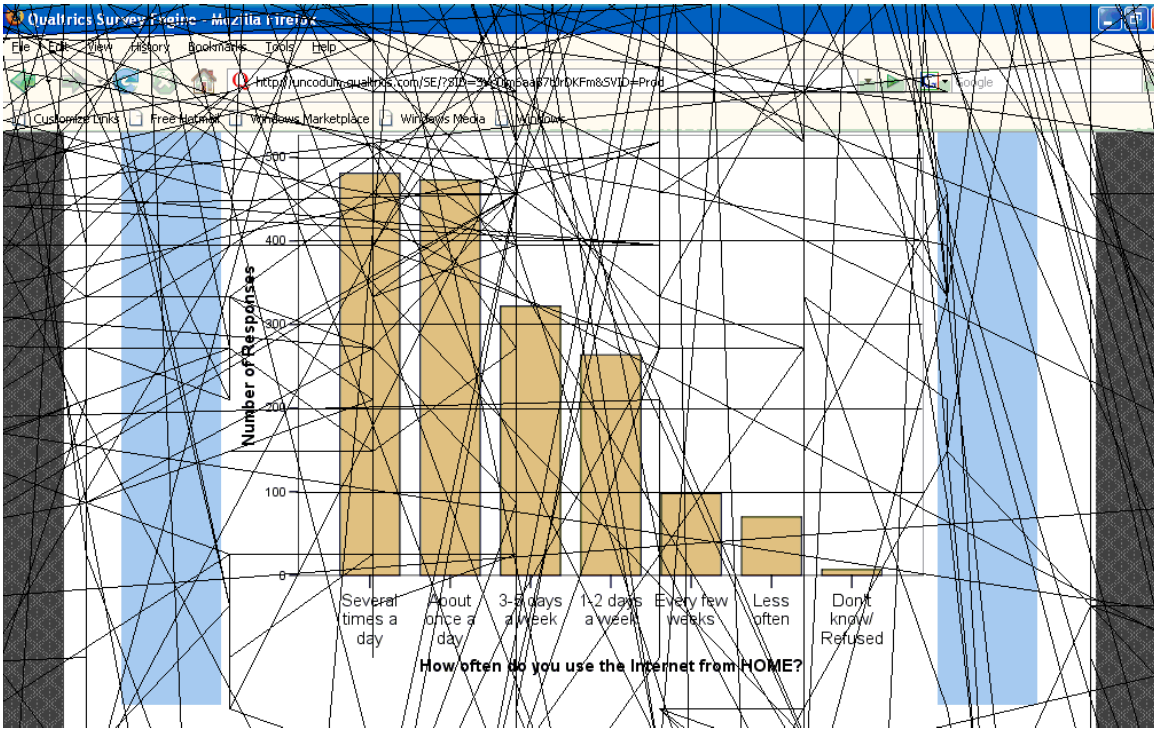


In the third graph, with graph order (1) Medium Ink, Lots of Ink, (3) No Ink, there is again an increase in pupil dilation at the start of the experiment. Unfortunately, due to missing data points, we cannot see if there is also an increase in dilation at the end of the experiment. However, as mentioned previously, pupil dilation does not appear to increase over the final third of the experiment if we just look at the existing data points.

Graphs orders 1 and 2 both included the graph with the least data-ink last. This suggests that there may be a pupil increase associated with having higher data-ink ratios. However, further research is required to fully investigate this question.

Eye Scan Paths

Figure 6



There was not a noticeable difference in how subjects looked at the three types of graphs. However, there appears to be some patterns in how subjects look at graphs in general. In the example scan path above, the subject focused on the Y-axis values of the graph and on the heights of the first two bars. The up and down lines on the graph are a result of the subject scrolling up and down the screen. There appears to be a pattern of looking at a question and then looking again at the graph. It also appears that the subjects focused their eyes on the parts of the graph with content and not on the white, blank parts of the graph.

5 DISCUSSION

The results of this study offer little evidence to support the data-ink ratio hypothesis. However, there are some questions raised that warrant further study. First, the fact that some differences were observed suggests that further testing should be done to investigate people's preference for graphs as determined by the data-ink ratio.

It is interesting that the subjects in this study appeared to show an overall dislike of the graph with the lowest data-ink ratio. In addition, when asked what they liked about the graphs, a number of students pointed to the grid lines. Both these results suggest that subjects might actually prefer graphs with a higher data-ink ratio.

Although there were some indications that subject prefer graphs with a higher data-ink ratio, we have no clear idea of how subject preference relates to accuracy in answering questions. In a timed test with more subjects, it would be useful to again ask for graph preference and then compare that to the accuracy of responses. Unfortunately, because the subjects answered almost all of the questions correctly, we are not able to effectively compare graph preference and question accuracy in this study.

In a future study, it would be useful to further test whether there is a relationship between pupil dilation and the amount of data-ink on a graph. In this very preliminary study, it appears that pupil dilation may increase as data-ink is added to a graph.

5.1 Eye Tracker System Areas for Improvement

In a future study, there are a number of improvements that should be made to both how the eye-tracking system is used and the testing procedure.

One problem throughout the study was that the magnetic head tracker did not work correctly on the “auto” setting. As a result, subjects had to remain still throughout the study because the eye-tracker could not accurately account for body movement. One side effect of this is that subjects read graphs during the experiment while having to pay a lot of attention to keeping their body still. This may have altered how they read the graphs, and, as a result, their behavior may not accurately reflect how they read graphs in a natural setting.

There were also a number of problems related to calibrating subjects' eye to the eye-tracking system. First, it appears that heavy eye-makeup, and especially mascara, leads to difficulty in calibrating the eye-tracker to subjects' pupils. Second, hard contact lenses and thick soft contact lenses make it difficult for the eye-tracker to measure corneal refraction. Additionally, as mentioned in the E5000 EYEPOS software tracker documentation guide, it is difficult to calibrate the eye of subjects with glasses. However, during the experiment, I found that if you first calibrated subjects' eye without glasses and then had them put their glasses back on, it was possible to gather eye-tracking data.

An additional problem using the eye-tracker related to subject blinking. Each time a subject blinked, the eye-tracker would not record the pupil dilation or corneal refraction. Although this was not a problem when a subject blinked briefly, blinking for a longer stretch of time appeared to throw off the calibration of the eye tracker. This effect was seen both when a subject closed their eye without opening them for a few seconds and also when a subject blinked quickly multiple times in a row.

5.2 Experimental Design Areas for Improvement

In a future study, there are a number of changes that should be made to the experimental design to more accurately test whether how the amount of data ink on a graph affects graph perception and learning.

First, one of the issues that subjects raised in their comments is that the graphs used in the study were not necessarily representative of graphs seen online. In a future study, it would be useful to have graphs that look like they could have appeared online at USA Today or CNN's website. Or, as an alternative, the study could include both academic and popular media graphs.

An additional problem that subjects pointed out with the study was that the bars of the graphs were not always organized in ascending or descending order. Subjects noted that having the bars out of order in some cases made the graphs harder to read. In a future study, all of the graphs should either have bars increasing or decreasing order or out-of-order. Although the order of the bars does not always reflect the way the data appears, this factor can be controlled when selecting data sets to use in the experiment.

A third improvement that could be made is that the graphs should have varying levels of difficulty and there should be at least two graphs of each difficulty level and type. For example, there could be eighteen graphs total with six graphs for each level of difficulty and six graphs for each data-ink ratio. Related to the level of difficulty, it would also be helpful to have some open-ended questions that required subjects to look more closely at the graphs and that did not have a specific “correct answer.”

In addition, one area for improvement in the study relates to the type of graph tested. It would be interesting to test whether subjects have different data-ink level preferences depending on the type of graph used. A number of subjects pointed out that bar graphs are “very simple to read.” In a future study, it would be interesting to also how graph perception changes with varying data-ink levels in test scatter plots, linear regressions and graphs with non-linear relationships.

In addition to varying the type of graph used, in a future study it would be useful to vary the Y-axis intervals for the grid lines. In this study, all of the grid lines were at rounded values such as 200. However, in a future study, it would be useful to test whether having grid lines becomes more important to subjects when they are asked to find Y-axis values such as 213.

A final area for improvement relates to accounting for scrolling during the experiment. One problem with this study was that it was difficult to isolate eye-movement patterns while subjects looked at each graph because subjects were always scrolling down to see the questions and then up again to look at the graph. To accurately measure eye-movement, future studies should have subjects answer the questions orally or answer questions on a subsequent screen.

6 CONCLUSION

The intent of this study was to gain a preliminary understanding of whether Tufte's notion of graph quality would hold up in a lab setting. For the purpose of the study, quality was defined by response accuracy, attention focus, and enjoyment level.

My initial hypothesis was that subjects would prefer graphs with a moderate data-ink ratio and that they would process information more accurately on graphs with a moderate data-ink ratio than a low or high data-ink ratio. The data from this preliminary study suggests that the first part of this hypothesis may have validity. Subjects indicated that they preferred the graphs with a medium and high data-ink ratio to the graph with the lowest data-ink ratio. Although preference between high data-ink and medium data-ink is unclear, there does appear to be a preference for both options over graphs with low data-ink.

Unfortunately, the results do not provide insight into the second part of the hypothesis. Due to the fact that only one question was answered incorrectly throughout the study, we cannot make any conclusions about the relationship between the type of graph and the accuracy of responses.

Questions for Future Research

Through this preliminary study, a number of additional questions emerged concerning how to measure graph quality and the data-ink ratio. Previous studies testing the data-ink ratio have focused on comparing data-ink levels and question response accuracy (Kelly, 1989; Gillian, 1994). However, the eye-tracking presents the opportunity for finding new ways to measure how people interpret graph quality visually.

One potential future research topic is to further study how people define graph quality and to compare patterns in eye-movement for graphs that are perceived as high-quality and graphs that are perceived as low-quality. In the subject responses, a number of subjects focused on the content of the graph itself and not its design when asked what they liked and disliked about the graphs. In a future study, it would be useful to explore how much perceived graph quality is connected to the content of a graph and how much it is connected to the graph's visual design.

Additionally, it would be useful to study whether people with different educational backgrounds and working memory capacities define graph “quality” in distinct ways (Shah & Carpenter, 1995). In this homogeneous group of well-educated subjects, common indicators of quality included the easiness of interpreting the data, and the clarity of the data labels. However further testing should be done to determine whether people from less-educated backgrounds identify the same items as graph quality indicators.

A final useful topic of study would be to gain a better understanding of the importance of the data-ink ratio for non-designers. When the data-ink ratio is compared to other graph design features such as graph color and graph title location, which feature do subjects rate as being the most important? Although the design community has identified the data-ink ratio as a key indicator of graph quality, further research should be done to determine whether non-designers agree with this design standard.

References

Card, S. K. *The psychology of human-computer interaction / Stuart K. Card, Thomas P. Moran, Allen Newell*: Hillsdale, N.J. : L. Erlbaum Associates, 1983.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology-Applied, 4*(2), 75-100.

Cleveland, W. S., & McGill, R. (1985). Graphical Perception and Graphical Methods for Analyzing Scientific-Data. *Science, 229*(4716), 828-833.

Disabato , N. J. (2006). The effects of Ajax web technologies on user expectations: a workflow approach. A Master's Paper for the M.S. in I.S degree. April, 2006. 51 pages

Gillan, D. J., & Richman, E. H. (1994). Minimalism and the Syntax of Graphs. *Human Factors, 36*(4), 619-644.

Hess, E.H. *The Tell-Tale Eye: How Your Eyes Reveal Hidden Thoughts and Emotions*. Van Nostrand Reinhold Co, Chicago, 1975.

Kelly, J. D. (1989). The Data-Ink Ratio and Accuracy of Newspaper Graphs. *Journalism Quarterly, 66*(3), 632-639.

Kosslyn, S. M. (1989). Understanding Charts and Graphs. *Applied Cognitive Psychology, 3*(3), 185-226.

Larkin, J. H., & Simon, H. A. (1987). Why a Diagram Is (Sometimes) Worth 10000 Words. *Cognitive Science, 11*(1), 65-99.

Levy, E., J. Zacks, et al. (1996). Gratuitous graphics? Putting preferences in perspective. *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*. Vancouver, British Columbia, Canada, ACM Press.

Lohse, G. L. (1997). The role of working memory on graphical information processing. *Behaviour & Information Technology, 16*(6), 297-308.

Lohse, J. (1991). A cognitive model for the perception and understanding of graphs. *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*. New Orleans, Louisiana, United States, ACM Press.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*. 37, 2098-2109.

Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin, 23*(6), 636-653.

Jakob Nielsen (1997). *How Users Read on the Web*. Retrieved December 1, 2006 from Alertbox.com: http://www.useit.com/alertbox/9710a.html

Nunes, T. *Street mathematics and school mathematics / Terezhina Nunes, Analucia Dias Schliemann, David William Carraher*: Cambridge [England] ; New York, NY, USA : Cambridge University Press, 1993.

Pinker, S. (1990). A Theory of Graph Comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing*: Hillsdale, N.J. : L. Erlbaum Associates, c1990.

Shah, P., & Carpenter, P. A. (1995). Conceptual Limitations in Comprehending Line Graphs. *Journal of Experimental Psychology-General, 124*(1), 43-61.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14*(1), 47-69.

Simkin, D., & Hastie, R. (1987). An Information-Processing Analysis of Graph Perception. *Journal of the American Statistical Association, 82*(398), 454-465.

Teghtsoonian, M. (1965). The Judgment of Size. *American Journal of Psychology, 78*(3), 10.

Tufte, E. R. *The visual display of quantitative information / Edward R. Tufte* (2nd ed. ed.): Cheshire, Conn. : Graphics Press, c2001.

Wainer, H. (1984). How to Display Data Badly. *American Statistician, 38*(2), 137-147.

Wainer, H., & Thissen, D. (1981). Graphical Data-Analysis. *Annual Review of Psychology, 32*, 191-241.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: its psychological foundation and relevance to display design. *v37*(n3), p473(422).

Appendix A

University of North Carolina-Chapel Hill
Consent to Participate in a Research Study
Adult Participants
Social Behavioral Form

IRB Study #_____________________
Consent Form Version Date: November 14, 2006

Title of Study: Graphs via Ink

Principal Investigator: Julia Kulla-Mader
UNC-Chapel Hill Department: School of Information and Library Science
UNC-Chapel Hill Phone number: (919) 962-3701
Email Address: jkullama@email.unc.edu
Funding Source: School of Information and Library Science

Study Contact telephone number: (919) 620-1442
Study Contact email: jkullama@email.unc.edu

What are some general things you should know about research studies?
You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study. There also may be risks to being in research studies.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. You will be given a copy of this consent form. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?
The purpose of this research study is to learn about how people interact with online graphs.

Are there any reasons you should not be in this study?

You should not be in this study if you have extensive experience with graphs, are a Ph.D. Student, or have a science or social science background.

How many people will take part in this study?
If you decide to be in this study, you will be one of approximately twelve people participating.

How long will your part in this study last?
The study will take between forty five minutes and one hour. No follow-up contact will be required.

What will happen if you take part in the study?
If you choose to participate in the study, you will first be fitted with an eye-tracking system headset. You will then complete a questionnaire detailing your experiences interacting with graphs online. Next, you will view three separate graphs and answer questions about each. Finally, you will complete a questionnaire comparing the graphs. Your use of the graphs during the study will be logged with a statistics and key logging software package.

What are the possible benefits from being in this study?
Research is designed to benefit society by gaining new knowledge.

What are the possible risks or discomforts involved from being in this study?
There may be uncommon or previously unknown risks. You should report any problems to the researcher.

Participants are not at risk from the eye-tracking equipment. The eye-tracking equipment uses a camera positioned more than one foot from the subject to track eye movements and operates at a level at least an order of magnitude lower than what is considered to be safe for reflected light. The safe chronic ocular exposure estimates for near infrared exposure is 10 mW/sq.cm (David Sliney of US Army Environmental Hygiene Agency and Myron Worbarsht of Duke University Medical Center) and our ASL device uses between 0.1 and 0.3 mW/sq. cm on average. Note that near infrared is technically referred to as IR-A, light between the 760 and 1400 nanometer wavelength. The ASL system operates at the 850 nanometer wavelength. The head tracker receiver is positioned 12-24 inches from the transmitter that produces a magnetic field between one-eighth and one-quarter (depending on head position) of the normal earth field of 0.6 gauss (manufacturer specifications). This configuration of eye and head tracker supplied by ASL Laboratories has found wide applications in government, university, and corporate laboratories.

How will your privacy be protected?
Participants will not be identified in any report or publication about this study. Although every effort will be made to keep research records private, there may be times when

federal or state law requires the disclosure of such records, including personal information. This is very unlikely, but if disclosure is ever required, UNC-Chapel Hill will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies for purposes such as quality control or safety.

Study participants will be assigned ID numbers. The names of the participants will not be recorded. The study records will be stored in a read/write/execute protected folder on the School of Information and Library Science server.

What will happen if you are injured by this research?

All research involves a chance that something bad might happen to you. This may include the risk of personal injury. In spite of all safety measures, you might develop a reaction or injury from being in this study. If such problems occur, the researchers will help you get medical care, but any costs for the medical care will be billed to you and/or your insurance company. The University of North Carolina at Chapel Hill has not set aside funds to pay you for any such reactions or injuries, or for the related medical care. However, by signing this form, you do not give up any of your legal rights.

Will you receive anything for being in this study?

Participants will receive a $5 gift certificate to the UNC bookstore. Participants who withdraw from the study before completing it will not be eligible for the gift certificate.

Will it cost you anything to be in this study?

There will be no costs for being in the study

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

What if you are a UNC employee?

Taking part in this research is not a part of your University duties, and refusing will not affect your job. You will not be offered or receive any special job-related consideration if you take part in this research.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

- -

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

_____________________________________ ________________
Signature of Research Participant Date

____
Printed Name of Research Participant

_____________________________________ ________________
Signature of Person Obtaining Consent Date

____
Printed Name of Person Obtaining Consent

Appendix B

Pre-Study Questionnaire

Please answer the following demographic questions.

How old are you?

18-25
26-34
35-54
55-64
65 and over

What is your gender?

female male

What is the highest level of education you have completed?

Less than High School
High School / GED
Some College
2-year College Degree
4-year College Degree
Master's Degree
Doctoral Degree
Professional Degree (JD, MD)

In what discipline did you receive your undergraduate degree?

What graduate degree will you receive when you graduate?

MSLS
MSIS
PhD

How often do you go
online?

- Never
- Every other Month
- Once a month
- Once every 2 weeks
- Once a week
- A Few Times a week
- Daily

How often do you look at graphs online?

- Never
- Every other Month
- Once a month
- Once every 2 weeks
- Once a week
- A Few Times a week
- Daily

On a scale from 1 to 5 (5 being most expert) please rate

	1	2	3	4	5
Your expertise with computers					
Your expertise with graphs					

Appendix C

Graphs and Associated Questions

Graph 1

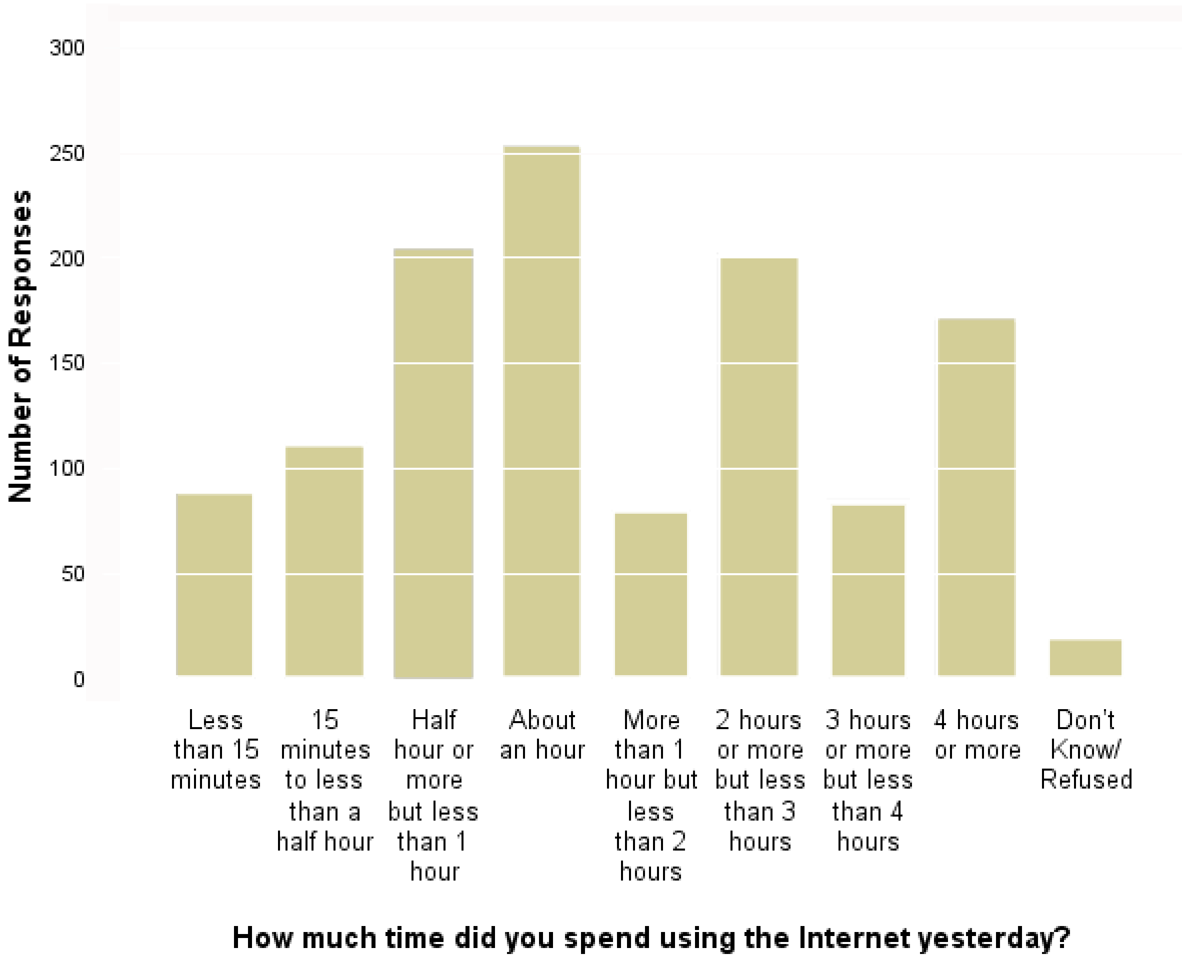
Graph 1 Questions

1. The highest number of people said that, yesterday, they spent:
   - Less than 15 minutes online
   - Half hour or more but less than 1 hours online
   - About an hour online
   - 2 hours or more but less than 3 hours online
   - 4 hours or more online

2. Approximately how many people said that they spent a half hour or more but less than 1 hour online yesterday:
   - 50
   - 100
   - 150
   - 200
   - 250
   - 300

3. True or False: More people spend 4 hours or more online than spent less than 15 minutes online
   - True
   - False

Graph 1: Low Data-Ink Ratio



Graph 2

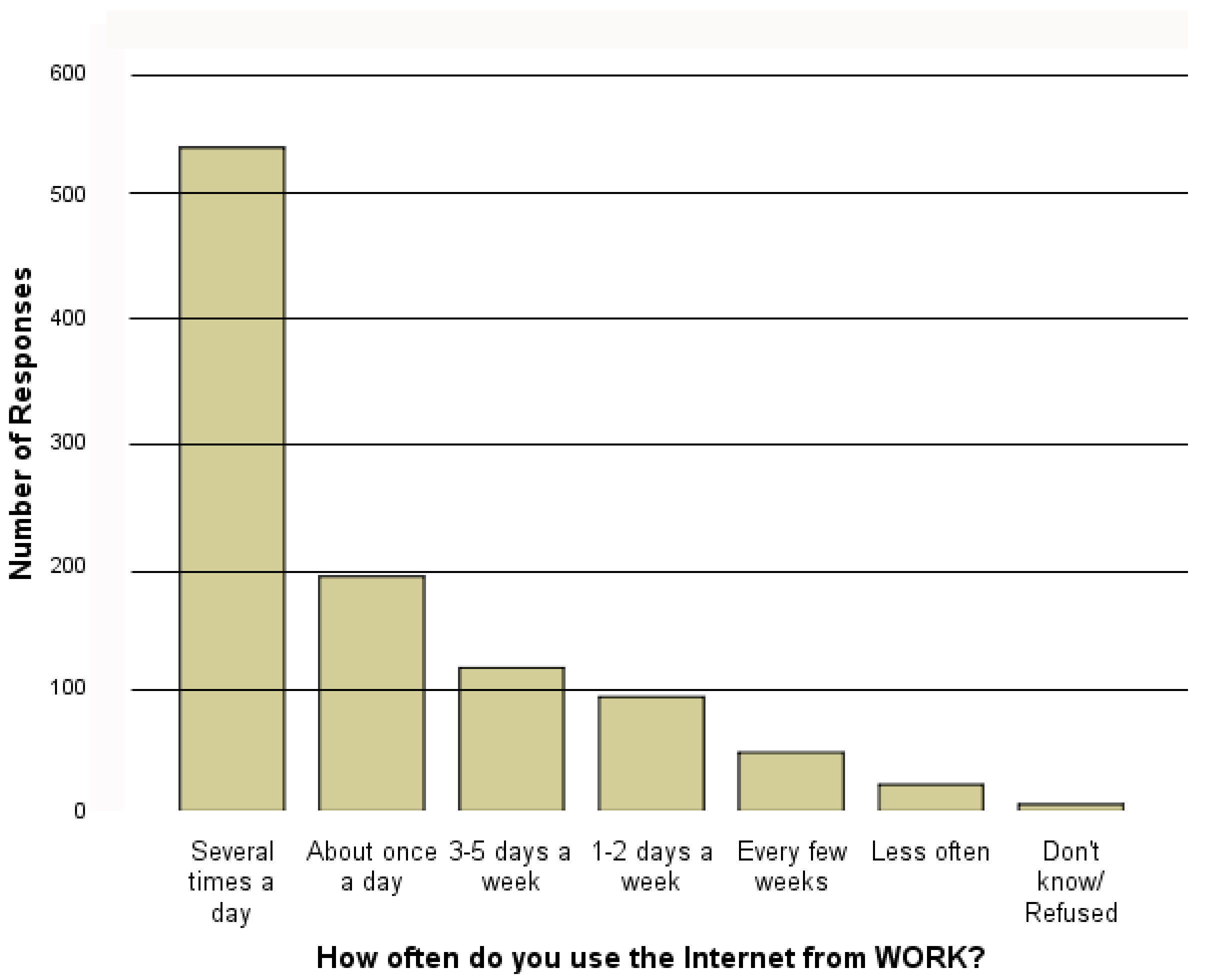
Graph 2 Questions

1. The highest number of people said that they use the Internet from work :
   - Every few weeks
   - 1-2 days a week
   - 3-5 days a week
   - About once a day
   - Several times a day

2. Approximately how many people said that they used the Internet 1-2 days a week from work:
   - 100

- 200
- 300
- 400
- 500
- 600

3. True or False: More people access the Internet every few weeks than 3-5 days a week
- True
- False



Graph 2: Medium Data-Ink Ratio

Graph 3

Graph 3 Questions

1. The highest number of survey takes said that they use the Internet from home

- Every few weeks
- 1-2 days a week
- 3-5 days a week
- About once a day
- Several times a day

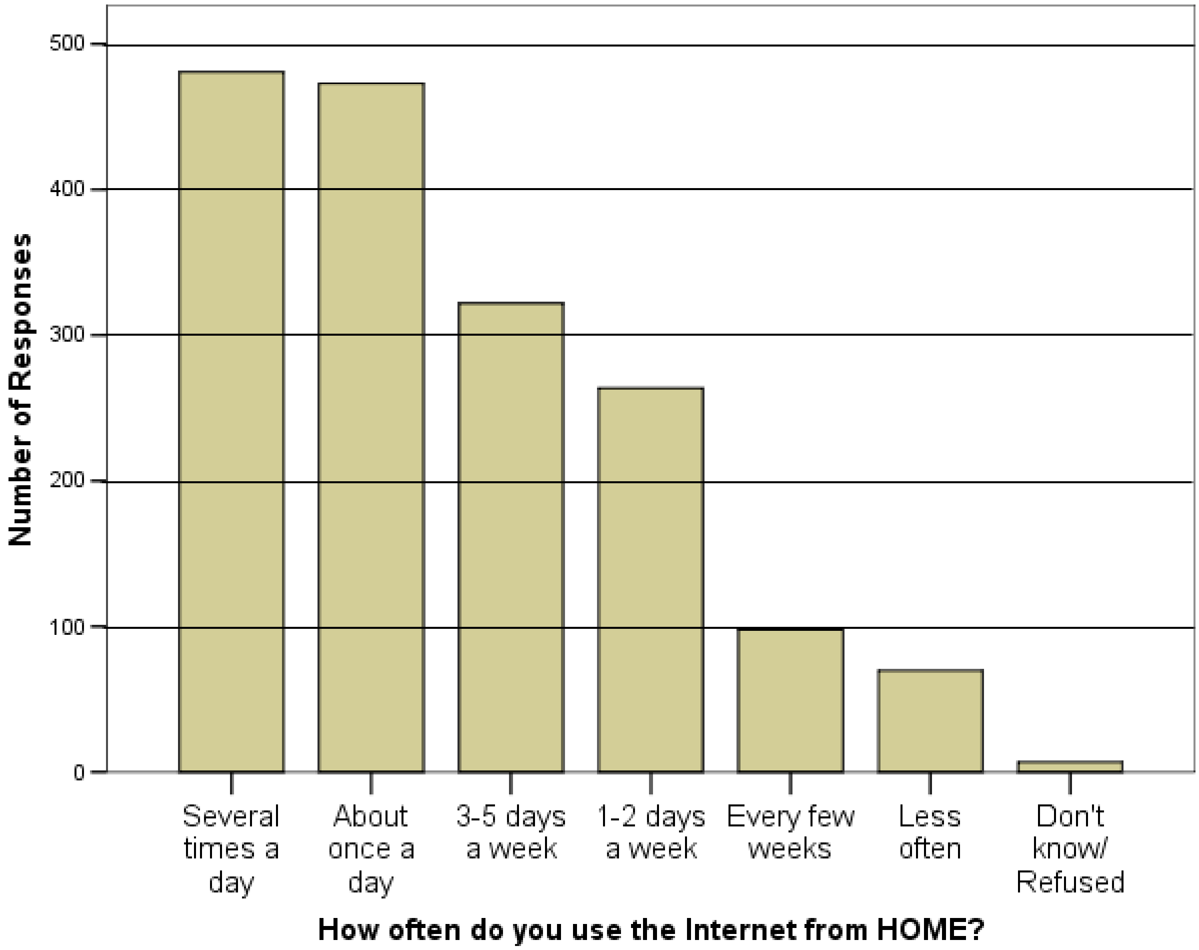
2. Approximately how many people said that they used the every few weeks from home

- 100
- 200
- 300
- 400
- 500
- 600

3. True or False: More people access the Internet 3-5 days a week than every few weeks from home

- True
- False

Graph 3: High Data-Ink Ratio


500
400
300
200
100
0
Number of Responses
Several times a day
About once a day
3-5 days a week
1-2 days a week
Every few weeks
Less often
Don't know/ Refused
How often do you use the Internet from HOME?

Appendix D

Post-Study Survey

To what extent did you understand the nature of the sets of questions following each graph?	Not at all	Somewhat	Completely
To what extent did you find the sets of questions following each graph similar to previous interactions you have had with online graphs?	Not at all	Somewhat	Completely
How different did you find the three graphs?	Not at all	Somewhat	Completely

Which graph did you most enjoy using?

What did you like about each of the graphs?

What did you dislike about each of the graphs?

On a scale of 1 to 5, with 5 being the highest, please rate your preference each graph.

	1	2	3	4	5
Graph 1					
Graph 2					
Graph 3					

Appendix E

E-mail Solicitation

Subject: Request to participate in research study on online graphs

Hello SILS students and community members,

I am a Master of Science in Information Science student at the School of Information & Library Science at the University of North Carolina, Chapel Hill. I would like to invite you to participate in an eye-tracking study to collect data on how people interact with online graphs. The study will only require one hour of your time and is completely voluntary. If you decide to participate, you will wear an eye-tracking headset designed to measure your eye movement patterns and pupil dilation as you look at online graphs. The eye-tracking system is designed for user's comfort.

The study is anonymous and no information that might personally identify you will be collected during the data gathering process. You may refuse to join, or you may withdraw your consent to participate at any time.

Please contact me at jkullama@email.unc.edu if you are interested in participating or if you have any questions. The study will take place in the Interaction Design Lab at the School of Information and Library Science. This research has been approved by the Behavioral Sciences Institutional Review Board at UNC-CH.

Sincerely,
Julia Kulla-Mader

MSIS Candidate
School of Information and Library Science
UNC Chapel Hill

Appendix F

Eye-Tracker Procedures

1. Turn on the equipment in the following order: command station, subject station, Flock of Birds Head tracker electronic unit, the pan/tilt camera, the eye monitor and scene monitors, the ASL Model 5000 Eye Tracker Control Unit.
2. Log-in at the command station computer.
3. Log-in at the subject station computer.
4. Open a command prompt and navigate to the c:\ASL\EYEPOSV5.34.
5. When in the EYEPOSV5.34 folder, type “load_nt.bat”. When the file completes loading (approximately 2 minutes), type e5win.exe. A window with the E5000 Control Program software will pop-up.
6. Open Start Menu>Programs>ERICA>GazeTracker on the subject station.
7. Click on “app” for application mode and then select “Create a new Analysis Configuration.”
8. To calibrate the eye via the E5000 Control Unit:
    1. Check the “Illuminator” checkbox under Power. The Illumination bar under Discrimination should be at more than 60 percent.
    2. Ask the subject to pick one point on the screen to fix their eye on. To find the subject's eye in the camera, click Pan/Tilt>Setup. Move the camera to the subject's eye using the up and down errors. Then, adjust the Shutter, Zoom, Iris, and Gain on the Pan Tilt Setup pop-up window and the Pupil, CR, and Illumination bars on the main screen until the subject's eye is located and the Pupil and CR lights are green on the main screen.

3. Click the MHT button in the right of the main screen. Next, in the Pan/Tilt menu, select “MHT sensor calibration.” Click “ok” on the pop-up window.
4. If the magnetic head tracker is working, click the Auto radio button under Pan/Tilt tracking on the main screen.
5. To calibrate the subject's eye, go to Calibrate > Eye Calibration. have the subject look at points 1 – 9. When while they look at each point and the Pupil and CR lights are green, click “Store Data for Current Point.”
6. When all 9 points are correctly calibrated, crosshairs will appear in the POG box on the E5000 Control Program window and in the scene monitor.
7. When the crosshairs appear and the subject is ready, you can start the experiment. First, you will need to create a new file in the E5000 control program. Go to File>New and create a name for your .eyd file. Press “Save”.
8. To start recording on both the GazeTracker and the E5000 system, press the red record circle on the GazeTracker screen and and the record arrow on the E5000 control screen.
9. To stop the experiment, press the stop button on the E5000 control program and on the GazeTracker window.