2013

Reducing Cognitive Load Using Adaptive Uncertainty Visualization

Gregory Block
Nova Southeastern University, greg.block@outlook.com

Reducing Cognitive Load Using Adaptive Uncertainty Visualization

by

Gregory Block

A dissertation submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy
in
Computer Information Systems

Graduate School of Computer and Information Sciences
Nova Southeastern University

2013

We hereby certify that this dissertation, submitted by Gregory Block, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.

Maxine S. Cohen, Ph.D. Date
Chairperson of Dissertation Committee

Sumitra Mukherjee, Ph.D. Date
Dissertation Committee Member

William M. Garrabrants, Ph.D. Date
Dissertation Committee Member

Approved:

Eric S. Ackerman, Ph.D. Date
Dean

Graduate School of Computer and Information Sciences
Nova Southeastern University
2013

An Abstract of a Dissertation Submitted to Nova Southeastern University
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Reducing Cognitive Load Using Adaptive Uncertainty Visualization

by
Gregory Block

March 2013

Uncertainty is inherent in many real-world settings; for example, in a combat situation, darkness may prevent a soldier from classifying approaching troops as friendly or hostile. In an environment plagued with uncertainty, decision-support systems, such as sensor-based networks, may make faulty assumptions about field conditions, especially when information is incomplete, or sensor operations are disrupted. Displaying the factors that contribute to uncertainty informs the decision-making process for a human operator, but at the expense of limited cognitive resources, such as attention, memory, and workload.

This research applied principles of perceptual cognition to human-computer interface design to introduce uncertainty visualizations in an adaptive approach that improved the operator's decision-making process, without unduly burdening the operator's cognitive load. An adaptive approach to uncertainty visualization considers the cognitive burden of all visualizations, and reduces the visualizations according to relevancy as the user's cognitive load increases. Experiments were performed using 24 volunteer participants using a simulated environment that featured both intrinsic load, and characteristics of uncertainty. The experiments conclusively demonstrated that adaptive uncertainty visualization reduced the cognitive burden on the operator's attention, memory, and workload, resulting in increased accuracy rates, faster response times, and a higher degree of user satisfaction.

This research adds to the body of knowledge regarding the use of uncertainty visualization in the context of cognitive load. Existing research has not identified techniques to support uncertainty visualization, without further burdening cognitive load. This research identified principles, such as goal-oriented visualization, and salience, which promote the use of uncertainty visualization for improved decision-making without increasing cognitive load. This research has extensive significance in fields where both uncertainty and cognitive load factors can reduce the effectiveness of decision-makers, such as sensor-based systems used in the military, or in first-responder situations.

Acknowledgements

I am deeply grateful to Dr. Maxine Cohen, whose patience, thoughtfulness, and sagacity served both as a model and inspiration to me during the long nights and weekends I devoted to this research. Through her research, teaching, and dissertation work with students, Dr. Cohen has left an indelible mark on the field of Human Computer Interaction.

I am also grateful to the members of my dissertation committee, Dr. William Garrabrants and Dr. Sumitra Mukherjee, for their time, insights, and feedback. Dr. Garrabrants paved the way with the excellent research in his dissertation, and Dr. Mukherjee opened my eyes to the marvels of machine learning and artificial intelligence through his unforgettable classes.

My parents and children never expressed any doubts that I would complete the doctoral coursework and the dissertation. Their constant words of support were a source of strength and motivation.

I am forever indebted to my wife, Cheryl, who shares my passion for incessant learning and self-enrichment; my fervent hope is to return the favor for all the sacrifice, devotion, and encouragement she has shown me in my doctoral studies when she undertakes hers.

My journey through the dissertation process has taught me one unmistakable truth, that the more I learn, the more I see there is yet to learn. I am in awe of the men and women whose brilliance and hard work shine in the books, journals, and proceedings I have spent so many hours studying. Our knowledge, it seems, overflows the bounds of the ocean; let us hope we can match that abundance with wisdom.

Table of Contents

Abstract iv
Acknowledgements v
List of Tables vii
List of Figures viii

Chapters

1. **Introduction 1**
   Background 1
   Problem Statement 3
   Goal 4
   Hypothesis and Research Questions 6
   Relevance and Significance 9
   Barriers and Issues 10
   Assumptions, Limitations and Delimitations 12
   Definition of Terms 14
   Summary 15

2. **Brief Review of the Literature 17**
   Introduction 17
   Early Studies in Uncertainty Visualization 17
   Techniques for Visualizing Uncertainty 20
   Visualization and Perception 20
   Visual Search 26
   Evaluating the Effectiveness of Uncertainty Visualization 28
   Defining Uncertainty 30
   Visualization, Uncertainty, and Problem Solving 33
   Cognitive Load Theory 37
   Summary 40

3. **Methodology 42**
   Research Methods Employed 42
   Specific research method(s) to be employed 49
   Instrument development and validation 59
   Formats for presenting results 60
   Resource requirements 62
   Summary 63

4. **Results 65**
   Findings 65
   Summary of Results 81

5. Conclusions, Implications, Recommendations, and Summary 83
Conclusions 83
Implications 86
Recommendations 89
Summary 92

Appendices

A. Participant Instructions 97

B. User Survey 102

C. IRB Approval Memorandum 102

D. Simulation Results 104

E. Survey Results and Participant Comments 107

F. Minitab Output 110

References 125

List of Tables

Tables

1. List of Independent, Dependent, and Controlled Variables 44
2. Sample Size in Previous Studies 48
3. Probability (*p*) Values for Memory Factors 69
4. Mean Responses for Survey Questions for Memory 72
5. Probability (*p*) Values for Attention Factors 74
6. Mean Responses for Survey Questions for Attention 77
7. Probability (*p*) Values for Workload Factors 78
8. Mean Responses for Survey Questions for Workload 81

List of Figures

Figures

1. Example Graphic of the Five-Day Track Forecast Cone (Definition of the NHC Track Forecast Cone, 2011) with permission. National Weather Service 7

2. The monitor window displays an overhead, photographic view of a scene under surveillance (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission. 51

3. The computer simulation monitor window displays a prompt for the dispatcher to respond to a threat (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission. 52

4. The Activities List displays activities the dispatcher must approve or report (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission. 55

5. Primary and secondary tasks displayed in split panels (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission. 56

6. Uncertainty visualization using bounded polygons Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission. 57

7. Histogram. Age of Participants 67

8. Histogram. Education of Subjects 68

9. Histogram. Primary Task Accuracy. Standard Interface 70

10. Histogram. Primary Task Accuracy. Uncertainty Interface 70

11. Bar Chart. Primary Task Mean Accuracy. Both Interfaces 71

12. Histogram. Primary Task Response Time. Standard Interface 75

13. Histogram. Primary Task Response Time. Uncertainty Interface 75

14. Bar Chart. Primary Task Mean Response Time. Both Interfaces 76

15. Histogram. Secondary Task Accuracy. Standard Interface 79

16. Histogram. Secondary Task Accuracy. Uncertainty Interface 79

Chapter 1

Introduction

Background

As the cost of collecting and storing data decreases, and the demand increases for up-to-the-second information to aid in decision-making and analysis, user interface designers will be challenged to develop interfaces that allow end-users to access large volumes of data from a plenitude of sources without overwhelming the end-user or diminishing the user's ability to interact with, analyze, and make decisions on the flow of information. As data collection moves out of the office or the factory floor, and into the field, the reliability of the data declines, due to faulty sensors, hostile or unanticipated environmental conditions, or technological limitations (Estrin, Govindan, Heidemann, & Kumar, 1999; Chong & Kumar, 2003).

Vehicle tracking systems provide a good example of the hazards that affect data reliability. A GPS device in the vehicle receives signals from satellites that can help determine the vehicle's latitude and longitude. The GPS coordinates are queued on a storage device in the vehicle, and periodically transmitted to a centralized server. In bad weather, or when travelling through an urban area, the GPS device may not be able to receive satellite signals; the signals may be distorted, resulting in inaccurate or imprecise GPS calculations; the operator may disable the GPS device, preventing data collection; some readings may be lost due to a faulty storage device or a network transmission failure. To a dispatcher analyzing the GPS data, detecting out-of-band or unreliable data may be more critical than tracking the in-band, nominal data (for example, it is more critical to the dispatcher to know if the vehicle operator had disabled the GPS device). To provide a decision maker with complete context, a system must provide information about unreliable data, including enough state information for the system or the decision maker to determine the likely source of the uncertainty (Lim & Dey, 2009).

The data sources for a vehicle tracking system are static (road and expected route information), profiled (vehicle type, vehicle operator information) sensed (GPS data, weather movements) and derived (actual route and speed calculations), each of which can be a source of error, or uncertainty in a context-aware system (Henricksen & Indulska, 2004). Data fusion techniques involve the integration of data from multiple sources, for example, a vehicle tracking system that integrates GPS readings with camera readings for line-of-sight perspectives. Computational techniques can be used to resolve conflicts between multiple data sources (Zhao, Fang, & Jiang, 2007). For decision-support systems, however, integrating data from multiple, and potentially conflicting sources, introduces an additional challenge when displaying uncertainty. One way of viewing uncertainty is through a probability distribution function (Thomson, Hetzler, MacEachren, Gahegan, & Pavel, 2005). A decision maker can evaluate data sources that convey conflicting information according to their differing probability distribution functions, but only if the software designer conveys the degree of uncertainty in visualizations so the decision maker can interpret the inputs.

Software visualizations that convey uncertainty information can provide a richer context for decision-making. However, visual elements must compete for limited end-user resources, such as attention and working memory (Wickens, 2002), factors that contribute to the user's cognitive load. Cognitive load reduces a decision maker's performance and ability to complete many decision-making tasks. This research discusses the effects of uncertainty and uncertainty visualization on the observer's cognitive load.

Problem Statement

Real-world environments are plagued with uncertainties, from faulty sensors, unreliable location readings, sporadic network connectivity (Girardin & Nova, 2005) to environmental factors, such as bad weather, darkness, and unplanned intrusions. These conditions lead to uncertainty, which adversely affects the decision-making process, and can even add to the user's cognitive load, further diminishing the user's capability to interact with visualizations in an augmented reality system (Zuk & Carpendale, 2006). A number of researchers have proposed various techniques for displaying environmental information when there is uncertainty about the information's reliability (Skeels, Lee, Smith, & Robertson, 2010). These approaches are inspired by the insight that humans are accustomed to dealing with uncertainty in their daily lives, and are well-equipped to make decisions in that context. However, adding uncertainty visualizations to a crowded visualization canvas can also adversely affect a user's cognitive load. As argued by Antifakos et al. (2004) displaying uncertainty information can increase cognitive load while providing some improvements that can offset or reduce cognitive load, and more research is needed to evaluate the trade-offs between the two approaches.

By understanding the trade-offs between uncertainty visualization and cognitive load, technologists can more effectively represent the physical and cognitive aspects of an environment, especially in situations of high uncertainty and increased cognitive load (van der Kleij, de Jong, te Brake, & de Greef, 2009). The need to represent uncertainty effectively without increasing cognitive load is especially acute in emerging technologies, such as context-aware systems. In context-aware systems, for example, the system may present unreliable conclusions due to the probabilistic nature of data sources (e.g. faulty sensors). Consequently, the user can lose trust in the system when faulty presentation leads to erroneous outcomes. Conveying data quality can improve the user's level of trust in the system, but the presentation of data quality must be simplified in order to minimize the burden on the user's cognitive load (Mühlhäuser & Hartmann, 2009).

However, the literature has not provided a proven theory for effective uncertainty visualization (Lapinski, 2009) and the effects of uncertainty visualization on reasoning (Zuk & Carpendale, 2007). Uncertainty visualization remains a persistent challenge, and consequently, implementation of uncertainty visualization has not met with widespread use (Zuk & Carpendale, 2006).

Goal

Techniques for conveying uncertainty can compete and conflict with conventional visualizations; for example, increasing the number of visual elements the user must track (called clutter) may burden the user's working memory, and interlacing visual elements with orthogonal characteristics may degrade the user's scanning strategies, affecting the user's attention (Bunch & Lloyd, 2006). Consequently, adding uncertainty visualization to a system may increase a user's cognitive load. However, uncertainty itself can contribute to cognitive load (Back & Oppenheim, 2001). The challenge is to convey sufficient degrees of uncertainty to the decision maker to reduce uncertainty-induced cognitive load (intrinsic complexity), without increasing the decision maker's cognitive load due to visual (extrinsic) complexity.

Zuk and Carpendale (2007) acknowledged that uncertainty visualization may increase the user's cognitive load, and described two methods for integrating uncertainty visualization without interfering with the user's task performance: first, by reducing the role of uncertainty visualization to after-the-fact analysis, and second, by supporting dual visualization systems so the user can choose which view is best suited for the situation. For example, a dual visualization system would reserve complex interfaces for more sophisticated users, but display simpler interfaces for unsophisticated users in order to reduce distractions and lighten the user's cognitive burden. The user can then adjust the complexity of the interface to match the user's skill level (Shneiderman, Plaisant, Cohen, & Jacobs, 2010). This research builds on this research by proposing an adaptive visualization system. The adaptive visualization system affects the display of uncertainty visualization during conditions where cognitive load is high. The adaptive approach seeks to reduce cognitive load by examining the cognitive costs of all visualizations, and disabling the visualizations with the lowest degree of saliency. This approach is taken from the insight that uncertainty artifacts may play a more crucial role in the decision-making process than ordinary data-driven artifacts in the decision making process.

The goal of this research was to identify how adaptive uncertainty visualization can decrease cognitive load arising from uncertainty more than visualization increases cognitive load arising from complex user interfaces. By mitigating the increase in cognitive load due to uncertainty, uncertainty visualization techniques can reduce the user's overall cognitive load.

Hypothesis and Research Questions

Uncertainty visualization is a term that describes techniques to represent uncertainty or ambiguity in information, to support the subject's cognitive processes in decision-making (Zuk & Carpendale, 2007). Uncertainty affects decision-making by presenting an incomplete specification of the problem, reduced relevance of inputs, and lacks clear stopping criteria. Solving problems in a context of uncertainty requires non-linear thinking, fragmented solutions, and step-wise refinements. Uncertainty visualization integrates the representation of data and uncertainty to augment the subject's reasoning.

The hurricane "uncertainty cone" is an example of uncertainty visualization. When a hurricane forms, The National Hurricane Service (a part of the National Oceanic and Atmospheric Administration) releases information to the public describing the storm's location and projected path. A cone surrounds the projected path, indicating a forecast error that is averaged over 10 years. The purpose for displaying the cone of uncertainty was to aid the public in their decision-making process (Broad, Leiserowitz, Weinkle, & Steketee, 2007). Figure 1 below illustrates the projected path of Tropical Storm Dean, which originated in the Caribbean, and was projected to make landfall over the island of Puerto Rico. An uncertainty cone is used to depict the forecast error over three and five day periods.

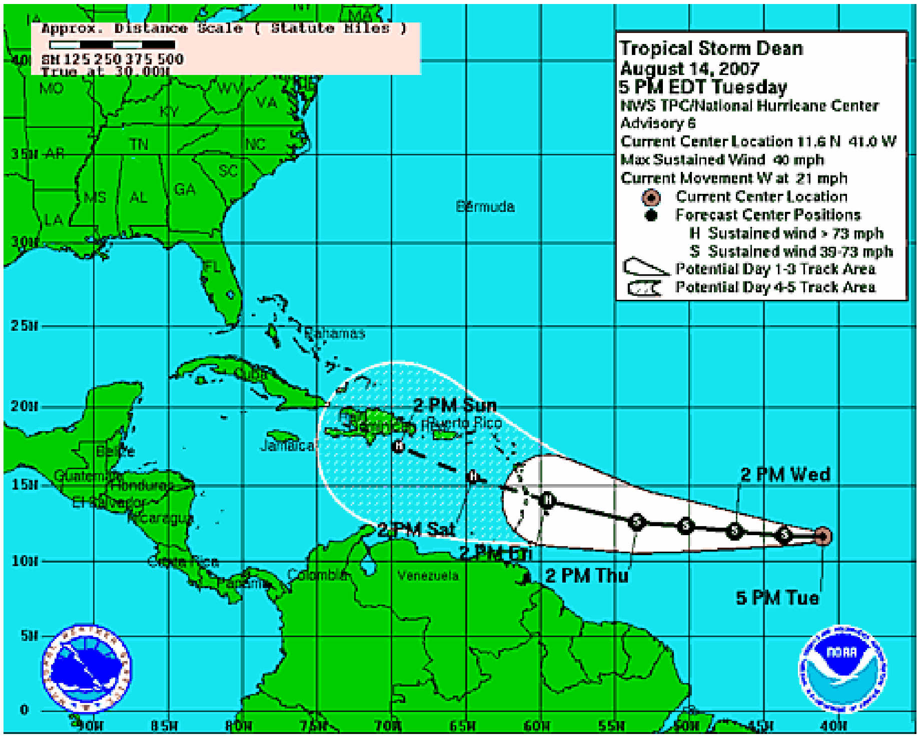


Figure 1. Example Graphic of the Five-Day Track Forecast Cone (Definition of the NHC Track Forecast Cone, 2011) with permission. National Weather Service

This research examined the effects of adaptive uncertainty visualization on the user's cognitive load. Uncertainty can increase the user's cognitive load (Zuk & Carpendale, 2006). Visualization techniques can be used to reveal and explain the source and degree of uncertainty, so that problem solvers can make more informed decisions (Skeels, Lee, Smith, & Robertson, 2010). This approach capitalizes on the fact that humans are generally good problem solvers, although a user's judgment may be biased (Tversky & Kahneman, 1974). Adding visual elements to a display, however, can also increase the user's cognitive load. For example, displaying uncertainty characteristics may increase clutter (Bunch & Lloyd, 2006). Consequently, there may be a trade-off between the reduction in cognitive load by displaying uncertainty factors, and the increase in cognitive load by displaying additional visual elements (Antifakos, Schwaninger, & Schiele, 2004).

Cognitive load can be assessed by measuring mental load, mental effort, and performance using an empirical approach that includes primary and secondary task measurements, as well as subjective rating scales (Paas, Tuovinen, Tabbers, & Van Gerven, 2003).In this study, performance, attention, and workload were assessed using primary and secondary tasks. Measurements for both the primary and secondary tasks included response time and accuracy rate. The secondary task was used to measure the cognitive burden imposed by the primary task. In addition, cognitive load effects were measured qualitatively by surveying participants' subjective impressions using a questionnaire. Responses were rated according to a Likert scale.

The hypothesis of this research is that adaptive uncertainty visualization will significantly reduce a user's cognitive load in an environment where both stress and uncertainty abound. The hypothesis (H) is that knowledge workers will exhibit better performance and improved decision-making using adaptive uncertainty visualization than when a standard interface is employed. The null hypothesis (H_0) is that knowledge workers will exhibit no better performance or improved decision-making using adaptive uncertainty visualization than when a standard interface visualization is employed.

This study addressed the following research questions:

1. Does adaptive uncertainty visualization improve the system operator's level of performance in completing assigned tasks? Performance was measured by the accuracy rate in completing assigned tasks.

2. Does adaptive uncertainty visualization improve the system operator's level of attention in handling multiple activities? Attention was measured by the response time required to complete assigned tasks.
3. Does adaptive uncertainty visualization reduce the burden on the system workload? Workload was measured by the accuracy rate in completing assigned secondary tasks.

In addition to the quantitative measurements described above, a survey was used to provide qualitative assessment of the operator's memory, attention, and workload. Accordingly, the hypothesis was evaluated using both quantitative and qualitative methods.

Relevance and Significance

The proliferation of mobile devices such as smart phones and tablets has raised interest in the development of pervasive computer systems (Baldauf, 2007). The goal of pervasive or ubiquitous systems is to integrate computing devices with a user's everyday experiences. Context-aware systems adapt to the user's environment, including the user's location, by using sensors, user profile information, and decision-making processes. Augmented reality systems interact with the user's environment by using a device to overlay virtual artifacts on top of physical objects in the user's line of sight, so the user sees a combination of virtual cues, and physical objects (Zhou, Duh, & Billinghurst, 2008). Pervasive, context-aware systems using advanced visualization techniques, such as augmented reality, provide tools to a diverse group of users to improve performance and decision-making; for example, for first-responders responding to a disaster scene (Piekarski & Thomas, 2009).

Given the increased demand for context-aware and augmented reality systems in a number of critical areas, such as medical and battlefield environments (Lundström, Ljung, Persson, & Ynnerman, 2007; Sager, Grier, Jackson, Levchuk, & Stelzer, 2007), problem solving under uncertain conditions is likely to become more essential. Factors that increase the problem solver's cognitive load are likely to increase as well. Accordingly, it will become increasingly more important to understand and quantify the trade-offs between exposing uncertainty to the problem solver, and the cognitive load this additional burden places on the problem solver (Antifakos, Schwaninger, & Schiele, 2004).

Because user interface designers do not know the effect that displaying uncertainty has on the user's cognitive load, they are unable to fully exploit features of uncertainty visualization (Mühlhäuser & Hartmann, 2009). It is hoped that this research will encourage user interface designers to take advantage of uncertainty visualization wihout overloading the user in order to improve the user's decision-making and problem solving tasks.

Barriers and Issues

According to de Jong (2009) there are a number of complications to measuring cognitive load. Cognitive load is typically measured using the following techniques: questionnaires given after a research experiment is conducted, so the participant can rate the results; measuring physiological characteristics, including heart rate and breathing variability, or by asking the participant to perform secondary tasks while participating in the experiment. There can be tremendous variability in responses when using questionnaires due to the wording of the questions, as well as the timing and frequency of conducting the survey during the experiment. Further, it has not been proven that research participants are competent at evaluating their own cognitive load.

Physiological measurements are also subject to a great deal of variability. For example, papillary reactions have been considered sufficiently sensitive for cognitive load studies, but the sensitivity diminishes with age. Some studies indicate that heart rate variability may be more sensitive to time pressures rather than cognitive load, although a combination of heart rate and blood pressure may be more sensitive to cognitive load studies. For example, in Haapalainen et al. (2010) six sensors were used to measure effects of cognitive load, including heat flux, ECG, EEG and pupillometry; however, only the heat flux and ECG produced accurate results. Discrepancies may have been due to the nature of the tasks the subjects were asked to perform, or the placement and sensitivity of the sensors. Finally, physiological measurements are intrusive and are likely to diminish the pool of participants (de Jong, 2009).

Secondary tasks are more useful than questionnaires because they are performed concurrently with the primary task. The motivation for using secondary tasks as a measuring proxy is that the speed or accuracy of the secondary task is diminished as cognitive load increases on the primary task. However, this approach is not frequently used; in fact, in research by Paas et al. (2003), only 4 of 27 studies measuring cognitive load used secondary tasks as a measuring technique. A possible explanation is that secondary tasks may distract or impair the subject from completing the primary task.

Overall, cognitive load studies provide measures that are characterized as relative, do not explore the multi-dimensional characteristics of cognitive load theory, and do not account for varying windows of time, such as immediate versus long-term (de Jong, 2009).

Assumptions, Limitations and Delimitations

Two limitations affected this research. The first limitation addresses the use of a simulation, and the second limitation addresses the measurement of cognitive load.

Simulations

First, as mentioned previously, a simulation was used to gather empirical information. Simulations have been characterized as a "third way" of conducting research (Axelrod, 2003) because the researcher starts with a set of assumptions that are designed into the simulation; however, simulations cannot provide deductive proofs, and can only be used to generate observations that support or refute a proposition. The researcher can improve the effectiveness of a simulation using an iterative approach; first formulating a theory, then building a computational model that generates results, and analyzing the results to refine the theory further (Emond & West, 2004).

Simulations are effective when used to observe phenomena that cannot be directly detected; the data generated by simulation is subject to less noise because the influence of external factors can be reduced (Goldspink, 2002). Further, simulations are useful in capturing adaptive, problem-solving behaviors (Axelrod, 2003).

A number of factors limit the benefit of using simulations in research. Because simulations are path-dependent and sensitive to the initial state (Goldspink, 2002; Axelrod, 2003) there is a challenge to repeatability of these experiments. In addition, the number and variety of variables involved in the execution of a simulation limit the ability to compare the outcomes of different simulations. Finally, social systems have multidimensional characteristics that are difficult to measure (Goldspink, 2002); for example, prototypes are more effective at identifying usability errors than efficiency measures (Sauer, Seibel, & Rüttinger, 2010).

Understanding these limitations, the goal of the researcher is to strengthen the experiment's validity through proper design and analysis. Validity is predicated on establishing a causal relationship between variables and observations that can be generalized in different settings (Oulasvirta, 2009).

Measuring Cognitive Load

A second limitation affecting this research was the measurement of cognitive load. First, a subject's cognitive load cannot be measured directly, and instead is induced indirectly by observing other phenomena, such as the subject's error rate or performance (de Jong, 2009). However, assessing a subject’s mental load, mental effort, and performance level can indirectly measure the subject’s degree of cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Accordingly, this research was limited to measuring indirect effects of cognitive load.

Delimitations

Two delimitations were imposed to define the boundaries set for this research. First, the participants in the research were confined to knowledge workers. Knowledge workers are subject to information overload (Karr-Wisniewski & Lu, 2010) and high levels of stress and anxiety that can lead to high degrees of cognitive load (Kirsch, 2000). Further, knowledge workers frequently interact with visualization tools for decision-making (Reinhardt, Schmidt, Sloep, & Drachsler, 2011). However, the findings in this research may not be generalized to other populations or situations, such as a combat setting, or with first-responders.

The second delimitation pertains to modality and simulation fidelity. Distributing cognitive load across multiple modalities, such as auditory and visual channels can lead to increased learning effects when compared to single-modality techniques (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). A subject may experience cognitive overload when an overwhelming amount of information is presented visually, but not experience cognitive overload if both auditory and visual signals are interspersed. In addition, multi-modal techniques are critical to increased fidelity in a simulated environment, and high fidelity leads to improved learning outcomes (Liu, Macchiarella, & Vincenzi, 2009). Nonetheless, this study was focused on the visual modality.

Definition of Terms

The following is a list of key terms and acronyms used in the fields of cognitive load, uncertainty, and visualization. The terms are defined according to commonly accepted usage among researchers and practitioners in these areas of study.

Attention: A set of cognitive processes that enable the detection and classification of stimuli by switching cortical processing and allocating resources (Sarter & Lustig, 2009).

Cognitive Bias: A tendency to favor one perspective over another due to cognitive factors, such as heuristics, rather than based on evidence (Kahneman, Thinking, Fast and Slow, 2011).

Cognition: The act of acquiring, organizing, and using knowledge (Neisser, 1976).

Cognitive load: A measure of the effort an observer expends to perceive and identify stimuli (Back & Oppenheim, 2001).

Cognitive Load Theory: A theory that holds that the resources allocated for cognitive processing are limited, and that learning is impaired when a task exceeds the capacity of the limited resources (de Jong, 2009).

Confidence: The observer's assessment of uncertainty in a system (Barthelmé & Mamassian, 2010).

Fovea: The central part of the retina, the fovea possesses a higher number of photoreceptors, and has more neurons dedicated to visual processing (Eckstein, 2011).

GPS: Global Positioning System; a satellite-based system for providing time and location information.

Heuristic: A problem-solving technique that seeks to answer difficult, time-consuming questions with adequate but incomplete solutions (Kahneman, 2011).

Information overload: An overwhelming increase in the number of decisions a knowledge worker must make in an environment fraught with disruptions (Kirsch, 2000).

Perception: A continuous, cyclical, cognitive process, consisting of anticipation, exploration, and information pickup (Neisser, 1976).

Saccade: Rapid, jerky steps by which the fovea moves toward a target during visual processing, which acts as a form of sampling (Eckstein, 2011).

Simulation: An experimental approach to studying behavior using models (White & Ingalls, 2009).

Uncertainty: A situation where the user has imperfect knowledge about information, a task, or a potential outcome; or lack of knowledge about the presence of error (Thomson, Hetzler, MacEachren, Gahegan, & Pavel, 2005).

Uncertainty Visualization: A technique to augment software visualization with characteristics of uncertainty to promote alternate interpretations (Zuk & Carpendale, 2007).

Summary

Sensor systems frequently operate in environments that are plagued with uncertainty. Bad weather, temperature fluctuations, and hostile intrusions can affect the reliability of sensors that track and monitor these conditions. Faulty sensors can generate unreliable location readings, and signals can suffer from sporadic network connectivity. Compromised source data can lead to uncertainty, which can adversely the decision-making process, and can even add to the user's cognitive load. Because humans are adept at problem solving under uncertain conditions, the presentation of uncertainty can lead to more effective decision-making.

In stressful settings, such as a battlefield or an air traffic control tower, the user's cognitive load is already strained. Visualizing uncertainty elements add to the visual clutter that competes for the operator's limited attention. The operator must invest increased effort to process probabilistic assessments. Consequently, cognitive load increases, degrading the operator's performance and problem solving effectiveness.

Given the importance of uncertainty in effective decision-making, there is a critical need for research that demonstrates how uncertainty visualization can be used without straining the operator's cognitive load. This research proposes to demonstrate how an adaptive visual system can provide relevant visualization of uncertainty to improve decision-making without further straining the operator's cognitive load.

Chapter 2

Review of the Literature

Introduction

The following section contains a review of the literature regarding key aspects of this research: uncertainty visualization, cognitive processing, and cognitive load. Visualization draws heavily on perceptive, and attentive processes (Barrett, 2011), and is intimately linked to problem solving and decision-making (Zuk & Carpendale, 2007). However, cognitive resources are limited, and perception, cognition, and decision-making activities must compete for scarce resources, such as working memory (Wickens, 2002). The purpose of this literature review is to examine research in the areas of uncertainty visualization, cognitive processing, and cognitive load, in order to validate the relevance and significance of an investigation into the reduction of cognitive load through adaptive uncertainty visualization.

Early Studies in Uncertainty Visualization

Andre and Cutler (1998) characterized uncertainty in the context of aviation display, identifying three separate dimensions that could influence uncertainty: accuracy, precision, and time. Time uncertainty may arise when there is a delay in reporting the location of an erratic or fast-moving object due to a slow refresh rate on the display, or a lag in receiving data feeds from a sensor. Andre and Cutler conducted two experiments using a simulated display to test the effectiveness of uncertainty visualization on pilots' bias for risk and situational awareness. In the first experiment, position uncertainty was represented using three separate techniques: a numeric value that represented the degree of uncertainty; a red-yellow-green color scale (the color red indicating the highest degree of uncertainty); and a circle enclosing an object with a radius that increased with the level of position uncertainty. In the baseline condition, however, there were no visual cues of location uncertainty. The second experiment was similar to the first but used three separate techniques for representing heading uncertainty. Heading was a means of classifying whether another aircraft was friendly or hostile. The researchers found that under conditions of moderate uncertainty, pilots performed equally well when uncertainty was displayed or not; but under highly uncertain conditions, pilots performed better in terms of the number of collisions and misclassifications, when uncertainty factors were displayed than when uncertainty factors were not displayed.

Finger and Bisantz (2000) studied the effectiveness of displaying uncertainty using various graphical formats as compared to quantitative indicators to aid in decision-making tasks. In one study, subjects were asked to decide whether an image on a card was friendly or hostile. The researchers used a number of techniques to convey whether an object was friendly or hostile: icons with associative meanings (for example, a mask with a smile was paired to a mask with a frown); abstract shapes with no associative meaning (for example, the shape of an arc was paired with the shape of a triangle); and symbols that were both iconic and abstract (such as by pairing a green symbol to a red symbol). Uncertainty was conveyed quantitatively using a percentage, or by distorting the graphical image.

For example, an object with equal probability of being friendly or hostile was represented numerically with a score of 50%, or by a mask with a circle in place of its mouth. According to the test results, subjects scored equally well when graphical formats were used as compared to numerical formats, and when the graphical format was combined with the numerical format, subject scores were not improved.

Rukzio et al. (2006) researched the effectiveness of displaying system confidence in a form-filling application, and concluded that users did not rely on, and did not find helpful, the visualization of confidence. However, the researchers suggested that in circumstances where the user was more invested in outcome of a task (for example, in an online reservation system) they would be more likely to rely on confidence visualization.

Cohen and Warren (1990) demonstrated that a user's confidence in an expert system is closely tied to the level of confidence the system expresses. The study sampled user confidence in the expert system before and after presenting the system's confidence in its recommendation. The confidence level the system displayed in its recommendation was adjusted to match the confidence level selected by the user; however, for half the participants the confidence level was increased by nine points (the "plus version"), while for the remaining half, the confidence was decreased by the same amount (the "minus version"). After the system displayed the confidence level in its own recommendation, the user's confidence in the system was captured again. Subsequently, 94% of users who revised their confidence level at this stage changed their confidence level in the direction of the system's confidence level; that is, if the system displayed a higher confidence level (the "plus version") most participants also increased the confidence level they used to assess the system. The outcome exemplified the anchoring heuristic identified by Kahneman and Tversky (1982).

Techniques for Visualizing Uncertainty

Sager et al. (2007) evaluated various techniques to visualize uncertainty. Uncertainty is generally represented by an additional piece of data, or by using a visual element that can describe both the physical element and its associated variability. Common visual techniques included the selection of color (for example, coloring an object using the familiar streetlight colors of red, yellow, and green might be used to indicate levels of confidence); by using texture (a cross-hatched texture could be used to indicate uncertainty) or a variety of icons (such as a question mark). These cues were found to be more effective than displaying numerical probabilities. However, significant shortcomings reduced the effectiveness of these techniques; for example, a particular object may have multiple sources of uncertainty (an unknown speed and trajectory for a moving object, for example); the uncertainty may increase or decrease over time; and the actions of other agents, such as enemy combatants, could not be adequately presented. The researchers developed a system for mitigating uncertainty by addressing these three deficiencies.

Visualization and Perception

Petre et al. (1998) asked several probing questions about the purpose and effectiveness of visualizations, and the impact on the user’s cognitive processing. One purpose of visualization is to change the viewer’s perspective so that a large-scale problem or situation can be compressed into a single view. A goal of this approach is to reduce complexity or scale. Visualization can be used as a symbolic interpretation of an external system, rather than a facsimile. Accordingly, visualization can serve as a platform for display-based reasoning by presenting an improved model of the external system, and consequently assist the user in their reasoning about the external system. Visualization is then a tool for reasoning, and can modify the user's tactics for information gathering, inspection, and comparison. Visualization can be used as an extension of the user's cognitive processes by extending short-term memory (since the user can offload concepts or problem explorations that are not of immediate use) in order to make problems more tractable. Visualizations can also recast a problem using a different model, such as by generating associations the user did not originate, or by serving as a foil for the user to cast a problem using different paradigms.

The contemporary understanding of visualization derives from research in neurophysiology and cognitive psychology (Petre, Blackwell, & Green, 1998). For example, Marr (1982) identified the computational aspects of vision as a form of complex information analysis. During visual processing, the mind scans a scene recursively, building abstractions from visual primitives. The first impression is characterized as a raw sketch, in which the mind evaluates attributes of visual primitives (such as position, contrast, and orientation) into abstractions like edges, blobs, and terminations. A critical aspect of visual processing is edge detection. Because physical surface changes are frequently marked by sudden changes in intensity, the brain uses filters to detect intensity changes on different scales. Once the viewer has formulated the primal sketch, the viewer then constructs a 2½-D sketch, which is comprised of depth, orientation, contours, and discontinuities. Finally, the viewer constructs a 3-D image, which is comprised of objects that have volumetric properties in relation to the space they occupy. Marr's work continues to have a profound impact on a number of fields, including cognitive science and neuroscience (Shagrir, 2010).

Neisser described perception as a continuous, cyclical process, consisting of anticipation, exploration, and information pickup (Neisser, 1976). Perception is not a passive process; there is no homunculus inside an observer's head that perceives objects from a retinal image. Instead, during perception, the mind develops a schemata, or plan for acquiring information. As information arrives, the schemata adapts its strategy for acquiring new information. To illustrate the effectiveness of schemata in information processing, experienced chess players can rapidly memorize the positions of chess pieces on a board because the player can associate the board layout with a schemata that rigidly prescribes the location of each piece. In fact, expert players may store thousands of schemata in memory. Neisser's work in cognition fueled considerable research into object structure, memory, and attention through the mid-1990s; however, due to advances in brain imaging, research has been focused more recently on the localization of functions in areas of the brain, and on attention and awareness (Cavanagh, 2011).

Barrett (2011) elaborated on the active role of the senses in perception. Senses are like tentacles that actively seek out and acquire information. The entire perceptual system is involved in perception, not individual organs; the senses work in concert. The perceptual system looks for affordances in the environment. Affordances represent the ways in which an observer can interact with an object. For example, the affordance of a rock to a person crossing the stream might be the opportunity to step on it. Perception is not easily stratified into stimulus and response; rather, perception is a tight-looped process that involves both senses and motor functions.

Csinger (1992) examined long-standing theories in cognitive psychology on how the mind processes information. For example, pre-attentive processes are extremely fast (less than 100 milliseconds) and the brain can perform multiple pre-attentive processes in parallel. Attentive processing requires more time, and is relegated to tasks that are more complex. The brain can rank perceptual tasks on a continuum, from easy to difficult.

Attention comprises a series of processes that focus sensory systems on certain characteristics of external stimuli by switching modalities and allocating resources, in order to optimize detection and classification (Sarter & Lustig, 2009). Attention can be categorized as selective, divided, and sustained. Selective attention describes perceptual processing where the observer focuses attention on one task to the exclusion of others, while divided attention describes perceptual processing where the observer can balance a number of concurrent tasks. If two tasks are similar, the observer is more effective by practicing divided attention; whereas if the two tasks are dissimilar, the observer is more effective by practicing selective attention. Sustained attention describes the degree to which an observer can maintain a state of readiness to perceive external events for an extended period.

Carrasco (2011) categorized visual attention as spatial, feature-based, or object-based. Spatial attention can be overt (that is, eye movements focus on the location of the subject's attention) or covert (that is, the focus of attention is not accompanied by eye movements). Feature-based attention occurs when the subject's attention is triggered by features in the visual field, rather than the location of these features. Features include such characteristics as color, or orientation. Attention to the feature is enhanced globally, to all locations in the visual field, even to ignored locations (White & Carrasco, 2011). Object-based attention occurs when attention is triggered by an object's structure. Because visual attention places tremendous demands on the brain's finite resources, the brain limits the amount of energy devoted to attention processes. Accordingly, visual attention is a selective process; for example, directing a subject's attention to one location in a visual field diminishes the attentional resources allocated to another location (Beck & Kastner, 2009).

Not all tasks require a subject's complete attention (Scerbo, Bliss, Freeman, Mikulka, & Robinson, 2005). In data-limited tasks, such as performing simple computations, the subject cannot improve performance by investing additional attention to task completion. Consequently, the subject has excess cognitive capacity that results in misdirected attention. Scerbo et al. (2005) categorized the subject's thoughts during task execution as task-relevant, task-related, and task-irrelevant. Thoughts that are unrelated to the task may be characterized as a failure of focused attention. Task-unrelated thoughts can interfere with the subject's task performance. Some task-unrelated thoughts, such as daydreaming, require the same spatial processing as complex cognitive activities, and may compete with the same modalities as the primary task.

Perceptual studies also focused on the sensitivity of the brain to visual primitives, and that the amount of light perceived by the eye provides an early vision, similar to pre-attentive processing. Early vision can detect the direction of light in a scene in a three-dimensional orientation. Csinger (1992) proposed a model of the visualization process using a permutation vector that contrasts a data surface with a perceptual surface. Steps for visualization include identifying the dimensions of the data space that should be projected onto the stimulus space; identifying the perceptual properties that will be used as the dimensions of the stimulus space; and mapping from the data space to the stimulus space.

The goal of mapping the data surface to the stimulus surface is echoed by Healy, et al. (1994). Given the challenge of presenting multi-dimensional data to an end-user for analysis, the researchers employed pre-attentive techniques in order to facilitate the user's understanding. In one experiment, the researchers used two visual features – hue and orientation, to represent data characteristics of salmon migration. Subjects were shown displays for 450 milliseconds, and asked to provide a numerical estimate of the data visualization to the nearest 10%. The subjects were able to provide reasonable estimates of the numerical data, suggesting that the visual features of hue and orientation were equally effective. Accordingly, visualization techniques can be used effectively to improve the effectiveness of user comprehension by leveraging the way the user perceives and processes stimuli.

Advances in visualization techniques will be limited by the lack of benchmarks and quantifiable measurements of effectiveness (Chen, 2005). Intrinsic quality metrics must be identified so that visualization techniques can be evaluated without referencing external sources. Chen cites the stress level used in multi-dimensional scaling as an example of an intrinsic metric; multi-dimensional scaling collapses multiple dimensions into two or three dimensions with minimal distortion. Accordingly, visualization techniques must be evaluated using intrinsic metrics to validate the fidelity to the underlying data and the degree to which intrinsic metrics are maintained. Another challenge for researchers is the role visualization plays in causality and forecasting. Visualization can be a powerful method for enabling an analyst to find causality, for example, in such areas as medicine and forecasting. The challenge for visual reasoning is to help the analyst distinguish noise from information, and to reconcile unrelated or conflicting data. Since the analytical process is exploratory, the analyst must interact with the raw data, as well as the representations.

Visual Search

Bertin (1983) identified eight variables that characterized a graphics system: the planar variables denoting a visual element's location in a visualization (x, and y), as well as six retinal variables: size, color, brightness, orientation, shape, and grain. Bertin further characterized variables according to how rapidly the variable could be perceptually processed: a variable was considered selective, for example, if its meaning could be perceived instantly, rather than sequentially processed in concert with other marks.

Tufte (2001) developed general principles for effective visualizations. For example, graphical excellence could be achieved with a number of principles, such as presenting a large amount of data in a small space, or providing multiple layers of detail. Data ink maximization was a technique for presenting the largest amount of data with the smallest amount of ink, since excessive use of graphics could distract the observer. Data density was a metric used to measure the number of data elements by the entire graphical area.

Ware's (2004) research was grounded in cognitive psychology research. Ware identified additional marks that were processed pre-attentively, such as blur and flicker. Gestalt laws described the features of pattern recognition, such as relative size and symmetry. These principles arose from theories in Gestalt psychology, such as the proposition that humans simplify visualizations by clustering and connecting elements of a scene (Wagemans, et al., 2012). For example, viewers may cluster together elements that are moving in the same direction. Further, humans have a tendency to add closure to shapes that are not completely closed, and to divide regions according to whether they fall inside or outside a closure (Ziemkiewicz & Kosara, 2010).

Ware (2008) expanded on the perceptual aspects of visualization. According to Ware, the brain uses a nested loop for information gathering and problem solving. The outer loop deals with generalized problems, which it breaks down into individual tasks. For example, finding a route on a bus line can be decomposed into tasks such as locating the starting point, the terminal point, and identifying candidate routes between these points. A middle loop executes a series of eye movements, or “visual queries” to gather information from the environment. Finally, when the eye comes to rest for a brief period, or fixates, on an object, an inner loop initiates a series of visual tests to identify patterns. A fixation typically lasts less than two-tenths of a second, and the brain can process approximately 20 patterns per second; accordingly, the brain can process up to four patterns per fixation. A pattern is detected through a process known as binding, where neurons that trace the contour of a particular pattern are stimulated, and emit electrical signals. The brain then distills patterns into objects. In visualization design, features such as color, orientation, and texture can be tuned to assist the brain in pattern recognition, which enhances the cognitive process.

In humans and other animals, light falling on a central part of the retina known as the fovea receives preferential treatment over peripheral areas (Eckstein, 2011). The fovea possesses a higher number of photoreceptors, and has more neurons dedicated to foveal processing than peripheral regions of the retina. During visual processing, the fovea does not move directly toward the target; instead, the eye moves in rapid, jerky steps known as saccades. Saccades act as a form of sampling; the brain uses the information perceived at each saccade location in order to inform a decision-making process (such as object detection and classification). The saccade pattern is influenced by a number of factors, such as the frequency and characteristics of distracters, the presence of context cues; or the prevalence of targets in the visual field.

Evaluating the Effectiveness of Uncertainty Visualization

Zuk and Carpendale (2006) employed heuristic evaluation to assess the effectiveness of visualization techniques for conveying uncertainty. The authors selected the contributions of three researchers in perceptual design theory: Bertin, Tufte, and Ware, and focused on each contributor’s perceptual and cognitive principles.

For example, a strategy for representing uncertainty in archeological reconstruction is to use markings that are sketch-like, as opposed to photo-realistic, as well as to use transparency to denote levels of uncertainty. The authors evaluated this technique in light of Bertin, Tufte, and Ware’s principles to assess the effectiveness of the technique. For example, using Bertin’s principles, transparency was an effective technique because the absence of marking indicates absence of data. Further, when evaluated according to Tufte’s principles, portraying uncertainty with photo-realistic effects could increase the “lie factor” (Tufte, 2001) of the depiction, overstating the confidence in the representation. Finally, graphical aspects such as contour can contribute to the viewer’s cognitive model, satisfying Ware’s perceptual theories. The heuristic evaluation of an air traffic control system centered on the use of alerts that notified the observer of significant events. A color scheme was used to denote the level of uncertainty. The color technique satisfied Bertin's principles, because of the variable "length", that is, the spectrum of color changes between green (low uncertainty) and red (high uncertainty) allowed for a large number of uncertainty levels. However, according to Tufte's principles, the data density was very low, indicating more data could be displayed in the same space, and numerical representations of uncertainty could be replaced with colors to enhance the user's scanning strategy. According to Ware's principles, however, the reliance on red and green colors could exclude color-blind people from completing tasks, and the high degree of color saturation could increase the observer's stress level. Finally, since alert systems rely heavily on visual monitoring, the system could make better use of scanning strategies by employing motion and flicker (Zuk & Carpendale, 2006).

Antifakos et al. (2004) analyzed the effectiveness of displaying uncertainty using a four-factorial model that focused on task difficulty, cost (that is, the risk/reward ratio of achieving a task), knowledge, and level of uncertainty (by determining whether to display the uncertainty, and if so, the quality of the display). The effectiveness of displaying uncertainty was proportional to the quality of the tip, the level of task complexity, and the benefit of a correct response.

Van der Kleij et al. (2009) studied the effects of a network-aware system on a user's mental effort in a mobile environment. In a mobile environment, network connectivity is not always reliable. When connectivity was sporadic or unreliable, the study participants reported low levels of process and outcome satisfaction. Further, mental effort was found to be higher when network connectivity was sporadic and location uncertainty was displayed than when the network was not connected. However, study subjects commented that the visualization interface was not useful in decision-making, and was not perceived as making the participants more effective in their tasks. Given that the interface was not perceived as useful or effective, the experiment likely increased the participant's extraneous cognitive load without any improvement in intrinsic load.

Applications with high levels of certainty can positively affect user impressions by displaying the uncertainty (Lim & Dey, 2011). The threshold for a high level of certainty was identified at 80-90% for non-critical applications. However, for applications with low levels of certainty, the user's impression is dependent on whether the application takes the appropriate action (given the circumstances). If the application takes appropriate action, displaying uncertainty can compromise the user's impression of the system; however, if the application fails to take appropriate action, displaying the level of uncertainty can actually improve the user's impression of the system because the user becomes more aware of the complexity required to decide which action is appropriate.

Defining Uncertainty

Schunn et al. (2003) developed a taxonomy for uncertainty, first classifying the sources of uncertainty. There can be uncertainty in measurement; uncertainty in computation (for example, stemming from stale data collection, or the introduction of artifacts in algorithms that cloud the results); visualization uncertainty (for example, where a visualization makes a false or misleading representation of a state, or omits critical information entirely), cognitive uncertainty, where the memory and perceptual limitations of the human problem solver may introduce uncertainties in a process. The problem solver may rely on several techniques for identifying systemic uncertainties, such as impossible representations (an object passing through a wall) or mismatched representations (for example, when two sensors provide conflicting information about the speed at which an object is traveling). When faced with uncertainties, the problem solver engages in a succession of strategies, such as checking for errors, identifying reliable inputs, calibrating the outputs of different sensors, and bounding the uncertainty in order to provide a resolution.

Henricksen and Indulska (2004) described four sources of imperfect information in a context-aware system: sensed, static, profiled, and derived. These imperfections are introduced by the computing system that interacts with the problem solver, rather than uncertainties in the environment. Imperfections can be unknown (when there is no sensor data), ambiguous (when two sensors report conflicting readings), imprecise (when sensors cannot report to a degree of precision) or erroneous. Henricksen and Indulska modeled the uncertainties using Object Role Modeling (ORM) by associating facts with one or more quality indicators, and these indicators are classified with concrete metrics.

Thomson et al. (2005) suggested that the term uncertainty denotes more than the lack of knowledge about the presence of error; instead, error is only one characteristic of uncertainty, and that uncertainty can describe situations with insufficient clearness or distinctiveness, accuracy or reliability; in short, where the user has imperfect knowledge about information, a task, or the outcome. Uncertainty can be quantitative, such as positional and temporal errors; but can also include abstract factors, such as the reliability of information sources or the degree of coverage.

Uncertainty can be understood using a probabilistic representation. Analysts make an assumption about the state of a system, which can be observed by collecting inputs from a variety of sources. The inputs from these sources can include measurements and locations, which are quantitative, but can also include statements and propositions. In a probabilistic model, uncertainty is the probability distribution of each source as compared to the actual system state. Consequently, Bayesian networks can be generated using the probability characteristics of each uncertainty type; for example, completeness uncertainty is subject to sampling error, resulting in variance and bias; interrelatedness uncertainty results from source correlation. Based on this probabilistic understanding of uncertainty, researchers can combine and propagate uncertainties, as well as identify composites of multiple uncertainties, to model more complex real-world situations. In addition, researchers can correlate visualization techniques that are most effective at representing each category of uncertainty, allowing each uncertainty to be displayed in its own dimension in order to improve the consistency of a visual model (Thomson, Hetzler, MacEachren, Gahegan, & Pavel, 2005).

Fout and Ma (2011) proposed a framework for uncertainty propagation that encoded the source of uncertainty (e.g., whether uncertainty arose from source data or from an algorithm). Each stage of data processing contributes another layer of uncertainty, so the uncertainty layers are encoded in a range number. A range number is a hybrid structure that normalizes uncertainty factors and assigns the uncertainty a credibility rating. Another way to express a range number is value = approximate value ±

deviation. Range numbers can be presented as bounds around a central tendency. For example, a bar chart can be topped with a solid line to represent the reading with the highest probability, and multi-colored bands displayed above and below the solid line to represent probability bounds.

Visualization, Uncertainty, and Problem Solving

According to Tversky and Kahneman (1974) decision makers use a heuristic process to simplify problem solving. The heuristics can be categorized as representativeness, availability, adjustment, and anchoring. Each of these simplifications can lead to biases that reduce the effectiveness of the decision process. For example, a reader may place more weight in crime statistics if they happen to live in a high-crime area (availability); a person who is handed a one hundred dollar bill and asked to estimate the weight of a nearby object is likely to start with a guess of one hundred pounds. Tversky and Kahneman (1982) later expanded on this problem solving bias, called anchoring, or the suggestion effect. Suggestions may be warranted because they provide information, but the decision maker's reliance on, rather than questioning of, the validity of a suggestion represents a bias. In the context of bounded rationality (Tversky & Kahneman, 1981), decision makers may choose to accept a simplistic frame of reference for a decision, in order to conserve mental activity. Accordingly, in the context of uncertainty, problem solvers tend to interpret probability subjectively, because uncertainty is not sufficiently codified and formalized.

Kahneman (2011) described a dual process for decision-making. System 1 is the name given to a process that is automatic, responds quickly, and requires little effort; a System 1 process can react autonomously to an external stimulus. System 2, on the other hand, is slower to react than System 1, but acts more deliberately; brings more resources, such as memory, to bear when solving a problem; and can solve complex computations that System 1 cannot solve. When confronted with a challenge, System 1 will attempt to solve the problem; however, if the problem is too complex for System 1, System 2 intervenes in order to bring more resources to bear. Both systems are effort conserving; System 1 resorts to heuristics to simplify the decision-making process, frequently by substituting a simpler question for a more complex one. For example, when confronted with a question requiring statistical knowledge a person does not possess (such as predicting how popular a politician would be in six months) System 1 will instead substitute an easier question (for example, by responding with how popular the politician is at the current time). System 2 also seeks to conserve effort during problem solving. System 2 will intercede during problem solving when a person is confronted with a problem that is too complex for System 1 to solve; however, System 2 will not intervene when System when makes a sub-optimal decision due to its reliance on a faulty heuristics process.

Zuk and Carpendale (2007) analyzed the effects of uncertainty on cognition in light of knowledge constructs, reasoning heuristics, and reasoning time frames. Uncertainty affects higher order knowledge constructs, such as arguments, which is the means by which a problem solver formalizes the problem-space for inferences and judgments. Uncertainty introduces ambiguity, lack of relevance, and incomplete knowledge of operation, resulting in partial solutions, or representational refinement, which increase cognitive load. Further, uncertainty may affect reasoning heuristics, leading to overconfidence when evidence strength is high and predictiveness is low or under-confidence when evidence strength is low and predictiveness is high. Finally, time constraints can subject the problem solver to biases, and that biases can increase uncertainty.

Juhnke et al. (2007) analyzed the effects of the human-computer dyad on problem solving in complex environments. The researchers describe interaction models between the human and the system. First, there is situational awareness, in which the human must be aware of the equipment used in the environment (systems awareness), as well as awareness of one's goals and how to achieve them (task awareness), and awareness of one's location (spatial awareness). The second interaction model is the action loop, which is a compressed process based on Norman's (1988) stages of action, including perception, evaluation, and execution. Perception begins when the participant recognizes an event that requires the participant to respond. Evaluation occurs when the participant considers the event and identifies a response. Execution occurs when the participant responds.

MacEachren et al. (2005) observed that the representation of uncertainty in geographic data tends to focus more on representational techniques than whether the representations contribute to better decision-making. For example, does uncertainty visualization encourage analysts to make better decisions in light of the levels of uncertainty, or cause the analysts to discount the uncertainty, even when that is not the most effective strategy? Does revealing uncertainty cause analysts to miss important relationships and associations, or does it encourage them to find patterns and relationships that do not really exist?

Hancock, et al. (2005) contrasted the degree of uncertainty to the level of performance for processing information. The higher the level of uncertainty, the more energy the subject must devote to searching for innovative solutions. The subject can process less information, and as a result, the performance of information processing is compromised.

Atoyan et al. (2011) examined uncertainty visualization in dynamic environments, such as automated systems. Human decision makers approach problem solving with reasoning strategies that are fine-tuned to the situation. The heuristics decision makers follow can be classified as compensatory or non-compensatory. In a compensatory strategy, the decision maker allows a high-scoring attribute to compensate for a low-scoring attribute. For example, a driver choosing a longer route with faster driving speeds is following a compensatory strategy (where travel distance and travel speed are two attributes of a route). Using a non-compensatory strategy, on the other hand, a decision maker does not make trade-offs between different attributes, and instead chooses the option having the highest value. For example, a driver who only chooses a route with the shortest travel distance is following a non-compensatory strategy.

Compensatory strategies impose a higher degree of cognitive load than non-compensatory strategies. Decision makers are more likely to use a non-compensatory strategy when faced with complex problems; however, under time-constraints, the compensatory strategy produces poorer results than the non-compensatory strategy because the decision maker does not have sufficient time to process alternatives. Visualizations can influence a user’s compensatory strategy. For example, an application that does not display multiple attributes concurrently, or does not allow the user to re-order attributes discourages the user from following a compensatory strategy. When a decision maker integrates uncertainty information into problem solving, the decision maker is following a compensatory strategy (the decision maker uses knowledge of missing or incomplete data to compensate for inappropriate system behavior).

Cognitive Load Theory

The problem solver must recognize the current problem state, and identify the differences between the problem state and the goal state. The act of problem solving can impose a substantial cognitive load on the problem solver (Sweller, 1988). Cognitive load provides a model for understanding the mental resources a problem solver can draw on when completing tasks – attention, and working memory. These resources are limited, but can be distributed between competing tasks (Wickens, 2002; Baddely, 2003). Sweller et al. (2011) argued that tasks that are biologically primary, such as human movement, can be easily acquired without undue burden on cognitive load; working with mechanical systems, however, may impose a higher burden on working memory because humans have not evolved the capacity to handle non-biological tasks.

Wickens (2002) proposed a four-dimensional model for timesharing multiple resources. Each dimension was described with two opposite levels. Two tasks demanding resources from the same level would experience interference; however, two tasks demanding resources from opposing levels would be less likely to experience interference. The dimensions included staging, characterized by perception, and response; modalities, such as visual and auditory; visual channels, characterized by focal and ambient; and processing codes, characterized by spatial and verbal. For example, speech recognition is a different cognitive activity from speech production (perception versus response), and take place in different sections of the brain (frontal versus posterior). On the other hand, studies have shown that subjects can divide their attention between auditory and visual inputs better than they can divide their attention between two auditory inputs, or two visual inputs (cross-modal versus intra-model). As Wickens observed, this observation may not be attributable to auditory and visual processing occurring in different parts of the brain; rather, two inputs in the same modality (for example, visual) may require scanning (if too far apart) or suffer from masking, if too close together (Wickens, 2002). Further, there is evidence that working memory is dedicated to different modalities, and as a result, tasks that a subject executes concurrently will only interfere with each other when the tasks share the same storage modality (Parasuraman & Caggiano, 2005).

Kalyuga (2011) identified four general situations that can increase a user's extrinsic cognitive load:

1. Split-attention – occurs when graphical and textual elements are separated spatially or temporally; requires recall for the user to integrate separated elements.
2. Redundancy – occurs when different sources provide the same information; for example, when explanatory text describes the elements of a graph or diagram.
3. Transiency – occurs when elements are displayed to the user for an insufficient length of time to process the information; increases the load on the user's working memory.
4. Expert versus novice – presenting information with more detail than is required for an expert user, or insufficient detail for a novice user, burdens the user's working memory.

Each of these deficiencies is mitigated through proper user interface design. For example, a split-attention scenario can be relieved by physically integrating spatially separated elements, or displaying elements concurrently; a transient scenario can be relieved by increasing, or allowing the user to customize the amount of time information is displayed.

The goal of Cognitive Load Theory (CLT) advocates is to design effective interfaces that minimize the problem solver's cognitive load (Oviatt, 2006). Advocates of Cognitive Load Theory distinguish between two types of complexity: intrinsic complexity, which arises directly from the execution of a task, and extrinsic complexity, which is introduced by mismatches in the interface. Oviatt proposed an interface design that enhanced user performance to reduce cognitive load, for example, by following interface principles, including accommodating the user's extant workspace and work practices, and minimizing interruptions.

Hollender et al. (2010) identified areas of convergence between CLT and Human Computer Interface (HCI) design principles. For example, the HCI design principles of recognition rather than recall, displaying only relevant information, or minimizing the amount of information the user must retain between dialog flows, are methods to reduce the load on working memory, which is also a key CLT objective. Furthermore, core CLT principles, such as the split-attention principle, infuse many usability guidelines, such as not requiring the user to remember information when looking at different sections of the same dialog. Not all CLT principles are matched in HCI research, however. The worked-example effect, which describes how novice learners may benefit more from studying examples solved by experts rather than struggling through ineffective problem solving exercises, is closely tied to learning effects rather than general usability issues.

Garrabrants (2009) conducted an experiment using a software simulation of a battlefield. In order to achieve a high level of situational awareness, the decision maker had to process a large number of variables in a short period of time, a situation that frequently leads to cognitive overload. In the simulation, a hypervariate display was introduced. The design of this display observed visualization strategies to reduce the observer's cognitive load. In particular, the hypervariate display was designed to take advantage of the observer's pre-attentive processing using symbology so the observer could quickly gain situational awareness. Cognitive load was measured using three factors: workload, comprehension, and efficiency. Using the hypervariate display, the researcher found that participants showed improved cognitive processing using the hypervariate display, as opposed to a multivariate display.

Summary

There is a considerable body of research to explain the cognitive and perceptual factors that influence a person's effectiveness as a decision maker in a variety of diverse environments. While humans are adept at problem solving in challenging conditions, innovations in computer design have provided additional tools to aid people in making decisions.

Decision-making is especially challenging when incomplete or unreliable information is available. Visualization techniques have been used to generate a probabilistic model of an environment in order to encourage effective problem solving. The danger of these visualization techniques is that humans may ignore the probabilistic nature of the information they received, and make false or unsupported inferences. Accordingly, it is critical to provide cues to the user that information is projected or estimated, and explain the source or degree of the unreliability of the information. Uncertainty then becomes another factor in the decision-making process.

In stressful situations, uncertainty visualization can have diametric effects. While uncertain conditions can increase the cognitive load on a decision maker, so can the burden of added visualizations and decision points.

This research contributes to the body of knowledge by identifying techniques to increase the use of uncertainty visualization in stressful environments without increasing the user's cognitive load. As the use of computer systems proliferates into more environments, such as battlefields and emergency situations, research into improved decision-making will become increasingly critical.

Chapter 3

Methodology

Research Methods Employed

Research was conducted using both quantitative and qualitative measurements to determine whether adaptive uncertainty visualization has a significant impact on reducing a user's cognitive load. Software visualizations displaying uncertainty characteristics may also increase the user's cognitive load (Bunch & Lloyd, 2006); there is a trade-off between the reduction in cognitive load by displaying uncertainty factors, and the increase in cognitive load by displaying additional visual elements (Antifakos, Schwaninger, & Schiele, 2004). The study capitalized on research in the visualization community by adapting the display of the uncertainty aspects in software visualizations to the user's level of cognitive load (Zuk & Carpendale, 2007). A computer simulation was selected to test subjects under two conditions (using uncertainty visualization, and using a standard interface) in order to analyze dependent variables, such as memory and attention, as a measure of the user's cognitive load.

Experimental Design

The hypothesis (H) of the research is that knowledge workers exhibit better performance and improved decision-making using adaptive uncertainty visualization than when a standard interface without uncertainty visualization is employed. This hypothesis was selected in order to answer the following research questions:

1. Does adaptive uncertainty visualization improve the system operator's level of performance in completing assigned tasks?
2. Does adaptive uncertainty visualization reduce the system operator's level of attention?
3. Does adaptive uncertainty visualization reduce the system operator's workload?

Ultimately, can the adaptive uncertainty visualization be calibrated to reduce cognitive load that stems from uncertainty without increasing the system operator's overall cognitive load?

The null hypothesis (H_0) is that knowledge workers do not exhibit improved performance and decision-making using adaptive uncertainty visualization than when a standard interface was employed. For purposes of testing the statistical significance, alpha (α) is defined as .05.

An analogue experiment was conducted using a simulation. Analogue experiments closely emulate a real-world setting so that the results can be more readily generalized (Oulasvirta, 2009). One of the principal advantages to experimentation with a simulation, as opposed to a field or in situ experiment, is that the data generated by simulation is subject to less noise because the influence of external factors can be reduced (Goldspink, 2002). Further, simulations are effective when used to observe phenomena that cannot be directly detected (Goldspink, 2002), and are useful in capturing adaptive, problem-solving behaviors (Axelrod, 2003).

The experiment used a single independent variable, a number of controlled variables, and three dependent variables. Table 1 summarizes the independent, controlled, and dependent variables of this experiment. The independent variable governs whether adaptive uncertainty visualization is displayed in the simulation. For the sake of brevity, this condition will be referred to as “uncertainty on/off” to indicate whether the adaptive uncertainty visualization was displayed. The dependent variables are memory, attention, and workload. These variables were measured quantitatively (using the subject’s accuracy rate and response time to complete tasks) and qualitatively (with an after-test survey).

Table 1

List of Independent, Dependent, and Controlled Variables

Independent	Controlled	Dependent
Uncertainty on/off	Duration	Memory
	Number of objects	Attention
	Number of events	Workload
	Speed of objects	
	Uncertainty size and range	

The controlled variables include the duration of each round of the simulation; the number of objects the user interacted with, and the speed at which the objects moved about the simulation; the number of events the user responded to, and the duration and accuracy of uncertainty projections made by the simulation when adaptive uncertainty visualization was displayed. The controlled variables affected the degree of cognitive load imposed on the subject. For example, a longer-duration trial would impose a greater degree of fatigue on the operator; an increase in the number of objects, or the speed at which the objects travel, would increase the operator’s mental effort. The multidimensional variables are correlated with the operator's cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). A pilot phase (discussed further in the Instrumentation section) was used to determine the optimal setting for the controlled variables.

In order to ensure that only the independent and controlled variables influenced the outcome, all other variables were held constant. This included the hardware the user interacted with, as well as the training instruction and material each user received. In addition, the simulation advanced time, and executed the same events at the same frequency during every round. Time-stepped simulations are used for human-in-the-loop simulations in order to ensure the user perceives a consistent flow of time and events during the simulation (Smith, 2000).

The Population

The study involved knowledge workers in a corporate setting. Knowledge work is a cognitive activity requiring substantial concentration and attention (Davis, 1999). Analysts, managers, and researchers fall under the definition of knowledge workers. Knowledge workers are suitable subjects for research because visualization tools are used increasingly to augment knowledge workers in the knowledge discovery effort (Eick & Fyock, 1996; Kandogan, 2001) and to reduce information overload (Karr-Wisniewski & Lu, 2010). Further, as documented by Kirsch (2000) the workplace environment for knowledge workers is characterized by high levels of cognitive load that induces anxiety, stress, and poor health. A number of studies have been conducted recently that correlate the disruptive effects of the workplace environment on worker productivity (Mansi, 2011) and cognitive load (Speier & Vessey, 2003).

According to Huck (2008) a purposive sample starts with a large group of potential subjects; however, in order to be eligible to participate in the study, the subjects must satisfy certain criteria. Accordingly, the research restricted the sample to knowledge workers of adult age who have experience with computer software. A candidate was classified as having sufficient experience with computer software if they had used a computer for business or academic purposes for a period of two or more years, and if they used a computer for two or more hours a day. Participation was equally divided between male and female subjects. Demographic characteristics are summarized in the study's findings in Chapter 4.

Sample Size

Choosing an effective sample size is critical to the validity of the research. A sample size that is too large results in an inefficient use of time and resources; while a sample size that is too small compromises the validity of the results (Triola, 2009).

While conducting usability tests using the thinking-aloud technique, Nielsen (1990) observed that subjects were very adept at identifying usability issues. Virzi conducted three usability tests using small sample sizes and concluded that a sample size of five subjects was sufficient to identify 80% of known usability defects (Virzi, 1992). Virzi approximated the relationship between the mean probability of detecting a problem and the number of subjects with the formula$(1 - p)^n$, where *p* represented the mean probability of detecting a problem, and *n* represented the number of subjects. Accordingly, a researcher planning to isolate a problem experienced by 10% of the population at the 90% confidence level would choose a sample size of 22, according to the formula; at the 80% confidence level, the sample size would be 15 subjects.

A number of sources challenged this finding. Spool and Schroeder (2001) argued that e-commerce sites are significantly more complex than the systems tested in (Virzi, 1992), and recommended a progressive approach of increasing the sample size as the number of identified issues and possible paths through the system increase. Faulkner (2003) argued that an increased sample size improved the probability of identifying more critical usability issues, and allowed results to be generalized to a larger population.

Another method for calculating sample size is to refer to the sample size in previous studies (Ritter, Kim, Morgan, & Carlson, 2011). Table 2 lists a number of similar studies, as well as the sample size chosen for each study.

Table 2

Sample Size in Previous Studies

Study	Sample Size	Comments
1. Haapalainen et al. (2010)	20	Students
2. van der Kleij et al. (2009)	48	Students
3. Garrabrants (2009)	18	Experienced volunteers
4. Skeels et al. (2010)	18	
5. Girardin and Nova (2005)	60	Students
6. Antifakos et al. (2004)	24 10	Students
7. Healey et al. (1994)	12 15	
8. Speier and Vessey (2003)	136	Students
9. Spool and Schroeder (2001)	49	Single-task test
Mean	37.3	
Median	20	
Standard Deviation	36.9	

As Table 2 illustrates, the mean sample size in the previous studies is 37.3, while the median is 20. The mean is more sensitive to outliers than the median (Triola, 2009). The mean is higher than the median due predominantly to Speier and Vessey (2003) which used a sample size of 136 students. However, the authors of this research did not provide reasoning for such a high sample size, in terms of population variability, error rate, or confidence level.

In light of these findings, a sample size of 24 can be sufficiently generalized to describe the population of knowledge workers this study represents.

Specific research method(s) to be employed

Background

A simulation follows a discrete-event approach when the model changes from state to state in discrete time points (Schriber & Brunner, 2009). A simulation tracks the passage in time using an internal stored value called a simulation clock. While the simulation time is not necessarily synchronized with the wall clock time in all discrete-event simulations, the simulation and wall-clock time was synchronized for this research, following a time-stepped approach (Alt & Lieberman, 2010).

In the simulation model, an executive process is responsible for advancing the simulation clock, and for carrying out actions that are scheduled to occur at specific times. These actions are enumerated in a structure known as the calendar, or future events list. The actions may be unconditional; for example, "At 20 seconds, a person enters the crosswalk"; while other actions are conditional; for example, "When the light changes to green, the person in the crosswalk begins to cross the street." The executive is responsible for checking that conditions are satisfied before actuating a conditional action; if conditions are not satisfied, the action remains in the future events list.

Entities represent the actors in a simulation. An external entity represents an actor introduced by the modeler, such as a package or a vehicle; while an internal entity is created by the incidence of certain states in the simulation, such as a machine failure, or collision. Entities instigate events, and events change the state of the model. Events may trigger other events; for example, the event, "person enters crosswalk" may trigger the event, “vehicle stops before crosswalk”. An event is placed on the current event list, where it is actuated by the executive; in addition, events may also place other events on the future event list (Alt & Lieberman, 2010).

Global variables refer to the limits or constraints in a model (White & Ingalls, 2009). For example, a global variable can be used to limit the speed at which vehicles travel in a simulation. The combination of entities, events, rules, and global variables is referred to as an experiment. When a modeler requires unique statistical results, a random number generator is used to ensure distinct outcomes. This is known as a trial. A run refers to the initialization of the model, and the execution; the simulation then executes until a condition is met that terminates the execution (Schriber & Brunner, 2009).

The Simulation

The simulation was comprised of a security surveillance system. A monitor was used to display an area under surveillance. The area under surveillance mirrored a realistic setting, such as a power plant. In the pilot, four monitors were displayed concurrently in separate quadrants of the screen; however, the amount of activity proved to be too demanding for pilot subjects to interact with; consequently, the number of monitors was reduced to one.

The monitor displayed an overhead, photographic view of the area under surveillance. The view was comprised of fixed landmarks, such as buildings and streets, as well as restricted areas, denoted by dashed rectangle. An example of the monitor is illustrated in Figure 2.

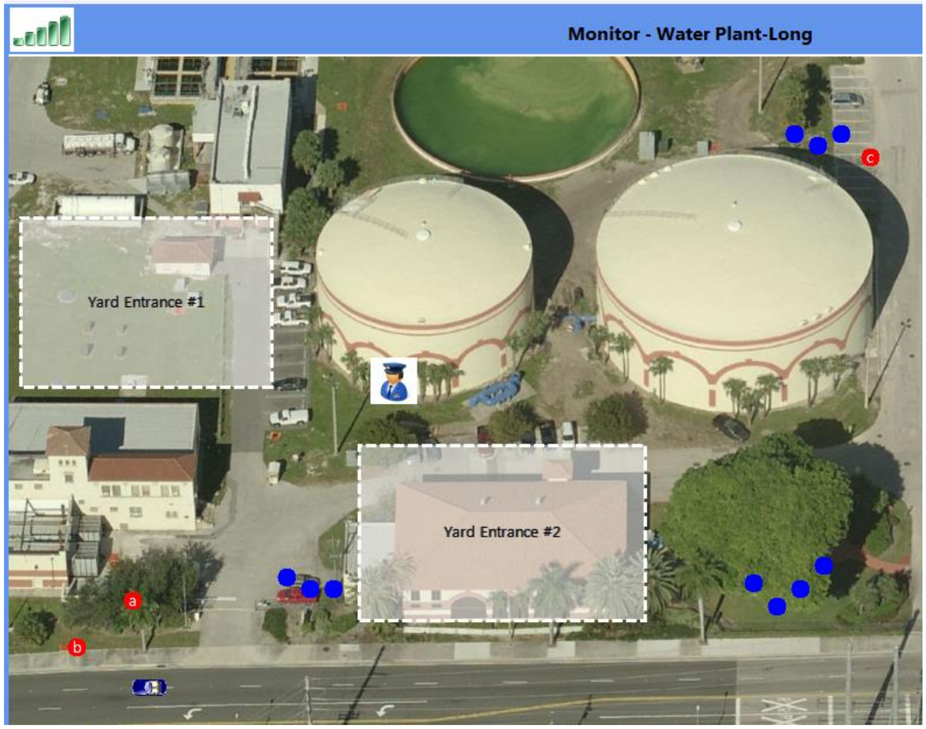


Figure 2. The monitor window displays an overhead, photographic view of a scene under surveillance (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission.

In addition, mobile "agents", such as people or vehicles, were represented using icons. Mobile agents represent entities in the simulation model. For example, the "view" of the power plant consisted of the overhead, photographic image of the power plant structure; an overlay of a fixed sensor that tracked movement; and an overlay of icons that represented agents, vehicles, and a security officer. The combination of overhead, photographic images of fixed landmarks, fixed sensors, and agents is called a scene.

Agents

Agents act in their own interests, according to a motive. A motive can be classified on a spectrum from "friendly" to "neutral" to "adversarial". Adversarial agents seek to cause harm or damage to property or other agents. A friendly agent seeks to monitor and maintain the safety and security of property. Neutral agents do not cause or prevent harm. Agents do not change their motives. A hostile agent's intrusion into a restricted area is an "event" according the simulation model. The presence of an intruder in the work plant setting constitutes a threat.

The Role of the Security Officer

The security officer is an agent in the simulation that interacts with the subject when threats arise in the simulation. The security officer interacts with the subject through the display of dialog boxes that prompt the subject for a response. The subject, who plays the role of a dispatcher, responds to the security officer's requests by selecting an option in the displayed prompt. Figure 3 illustrates the interaction between the security officer and the subject.

Figure 3. The computer simulation monitor window displays a prompt for the dispatcher to respond to a threat (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission.

When a hostile agent enters a restricted area for the first time, the appropriate response for the security guard is to issue a warning. The second time the hostile agent enters the restricted area; the appropriate response is once again to issue a warning. On the third intrusion, however, the appropriate response is to call for backup in order to intervene with the intruder. The appropriate response for the situation is described in a textual guide known as the Dispatcher Instructor Sheet (exhibited in Appendix A). The Dispatcher Instructor Sheet is a printed card that is positioned beside the computer terminal, within easy view of the subject. The security officer can detect which hostile agent has entered a restricted area, but is not able to determine how many times the agent has already entered the restricted area, and therefore cannot independently produce the appropriate response in accordance with the rules defined in the Dispatcher Instruction Sheet.

Sensors

In the scene, a sensor is used to track the location of all agents. Sensors transmit location information visually to the dispatcher. Sensor transmissions are subject to sporadic network disruptions. When a sensor transmission is disrupted, the scene no longer displays the real-time location of agents. During network disruptions, the subject will see either the uncertainty visualization interface, or a standard interface, according to which round of testing the subject is undergoing.

The Role of the Dispatcher (Subject)

The subject of the experiment takes the role of dispatcher. The dispatcher observes the situations on the four monitors and interacts with the on-scene security officer in each situation. When a threat occurs, the security officer will “ask” the dispatcher whether to proceed with a particular response. The dispatcher may refer to the Dispatcher Instruction Sheet at any time during the simulation. However, the dispatcher must respond to the security officer's request within ten seconds; otherwise, the security officer will make an independent (and possibly wrong) decision. The security officer can detect which hostile agent has entered a restricted area, but is not able to determine how many times the agent has already entered the restricted area, and therefore cannot independently produce the appropriate response in accordance with the rules defined in the Dispatcher Instruction Sheet.

The Secondary Task

The dispatcher's primary task is to interact with the security officer during threats in order to determine the most appropriate response, with a goal of restoring the situation to a normal state. In addition, the dispatcher has a secondary task. A table is displayed at the right of the simulation monitor. At preconfigured intervals, a record is inserted into the table, corresponding to a clock-in or clock-out activity for an "off-screen" security officer (not the security officer involved in the situation, however). The dispatcher must click on a section of the record to "approve" or "report" the activity. The determination to approve or report an activity is based on whether the officer is clocking in, or out, early, late or on time. According to the rules stated in the Dispatcher Instruction Sheet (exhibited in Appendix A) the subject should approve any activity that occurs on time; however, the subject should report any late clock-in, or early clock-out; and approve any early clock-in, or late clock-out. This exercise requires the subject to read and comprehend the text of the activity, and to recall which rule, as stated in the Dispatcher Instruction Sheet, is most appropriate. Once the activity is acknowledged, the record is removed from the table. The secondary task has no time limit for capturing a response; if the user fails to respond to more than ten activities, a vertical scroll bar appears in the table, and additional prompts are queued off-screen. The secondary task is illustrated in Figure 4.

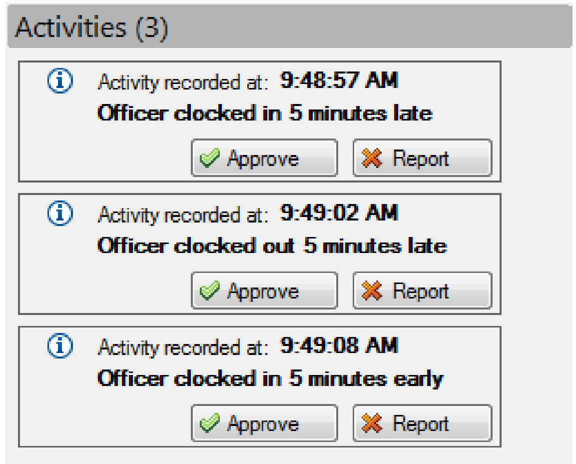


Figure 4. The Activities List displays activities the dispatcher must approve or report (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission.

The subject is advised that the monitor activities are high-priority, health, and safety issues that require “99%” of the subject’s attention. The clock-in activities are administrative, and the subject should only pay attention to these activities “as time permits”. Figure 5 displays both the primary task and secondary task panels side-by-side. The subject’s attention is split between the primary and secondary task panels.

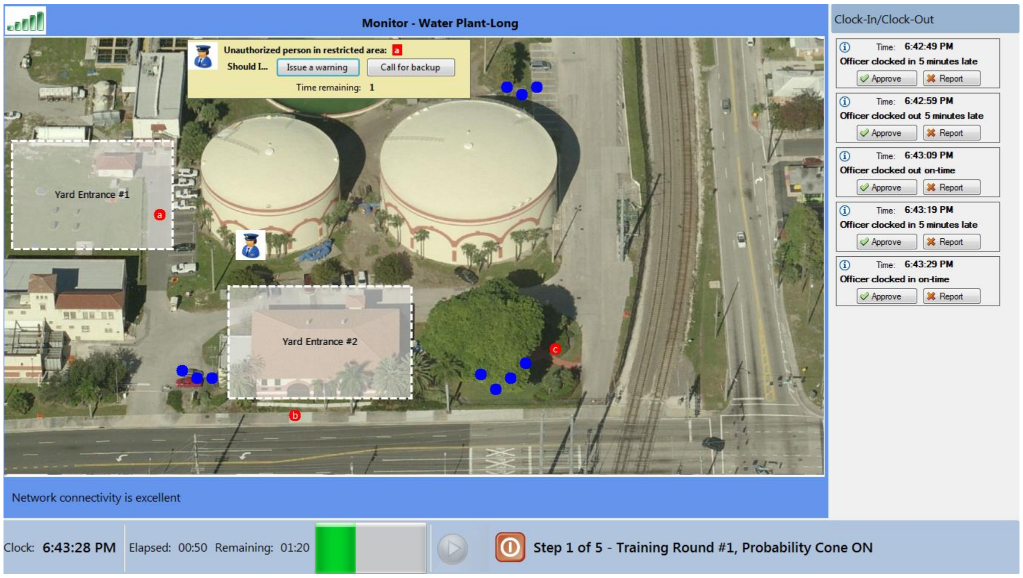


Figure 5. Primary and secondary tasks displayed in split panels (Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission.

The Simulation

A configuration script was used to configure the simulation at the start of each run. The configuration script determined the number and placement of agents. Time-variant features, such as agent movement was configured in the script as well. When the simulation is initialized, the actions defined in the configuration script are placed in the simulation's future events list. As the simulation progresses, the executive process moves actions from the future events list to the current events list, where the actions are executed.

Uncertainty Visualization

As mentioned before, disruptions in sensor transmissions are sources of uncertainty. Uncertainty was visualized as follows:

- A visual indicator denoted when network connectivity was disrupted.
- Visual indicators using color and a bounded polygon denoted the probabilistic location of mobile agents during network disruptions (Andre & Cutler, 1998).

Figure 6 illustrates the visual indicators used in the simulation during a disruption in network connectivity. During a network disruption, banners located at the top, bottom and edges of the monitor change from blue to red. Further, in the “uncertainty on” phase, a projection connects the last known location of each agent with the agent’s probable location.

Figure 6. Uncertainty visualization using bounded polygons Figure created by G. Block using Microsoft Windows snipping tool). Reprinted with permission.

The uncertainty interface provides a mechanism for modulating the effects of computer visualizations on the user’s cognitive load. When the uncertainty interface is employed, other less-essential visualizations are removed in order to offset the additional cognitive burden. That is, the salience of essential elements is increased by muting the display of non-essential elements. An element is considered essential if presentation contributes to achieving a goal. For example, vehicles in the simulation do not interact with agents and do not enter restricted areas; accordingly, vehicles are non-essential. Therefore, the display of vehicles is suppressed during network disruptions to minimize cognitive load. In addition, the satellite image switches from a color display to a black and white display to sharpen the contrast between items of interest and inanimate objects. By reducing the cognitive burden or other, less-essential visualizations, uncertainty visualizations can be added to the screen without unduly burdening the user's cognitive load. Accordingly, the decision maker can improve awareness of the environment with less danger of impairment from an increased cognitive burden.

There were two rounds of the simulation, each lasting ten minutes. In one round, there were no visual indicators of uncertainty. This round is referred to as the "uncertainty off" round, or the standard interface. In the "uncertainty on" round, the visual indicators of uncertainty were displayed. The "uncertainty on" and "uncertainty off" rounds were counterbalanced: half of the population started with the "uncertainty off" display, while the other half of the population started with the "uncertainty on" display.

Training

Due to the complexity of the user interface and the amount of interactions required for the primary and secondary tasks, each subject was asked to participate in two short training rounds lasting two minutes each round. The purpose of the training round was to familiarize the subject with the components on the screen (such as the scene monitor, and clock-in/clock-out table) and how to interact with the components. In addition, the training session introduced the uncertainty visualization elements. The subject was permitted to ask questions, and to repeat the training round if necessary. The training results were not included in the analysis.

Instrument development and validation

The simulation experiment was configured using a step-wise, iterative approach in order to increase the generalizability of the results. First, the researcher formulates a theory, and then builds a computational model that generates results to test the theory. The researcher then analyzes the results in order to refine further the theory, and repeats the process (Emond & West, 2004). Accordingly, the research used a pilot phase in order to validate the computational model. The pilot phase was used to define the configuration of the simulation and the setting of controlled variables. In addition, the pilot was used to develop benchmarks for subject responses, according to the configuration of controlled variables. When the pilot was completed, observations from this phase were used to specify design and interface changes in the simulation. For example, the number of monitors was reduced from four to one in order to better match the capabilities of pilot users.

Tests were conducted on the same hardware to ensure no variance in results was introduced by differing screen dimensions or resolutions; or differing keyboard or mouse layout. A Gateway NV54 laptop computer was used, running Windows 7 Home Premium[1]. No server, wireless, or internet connection was required.

The application recorded user inputs (both mouse and keyboard) in order to calculate response time and accuracy rates. Response time measures the duration, in seconds, between the initial display of a prompt, and when the user's input is detected in response to the prompt. If the user did not respond to a prompt within a specified period,

[1] Windows is a registered trademark of Microsoft Corporation in the United States and other countries.

or the testing round completes before the user was able to respond to a prompt, the application marked the prompt as not completed by the user.

Formats for presenting results

The subject's primary task was to respond to prompts posed by the security officer when a threat occurred. This is a relatively complex task because the subject was required to assess the situation, recall the number of times a particular agent had already trespassed on a restricted area, and then recall (or quickly refer to) the Dispatcher Instruction Sheet in order to provide the correct response. Because each subject was likely to choose different problem solving techniques to the primary task, variability in the execution of the primary task was expected.

The subject's secondary task was the acknowledgment of clock-in and clock-out activities. Because this task was less complex, and less variable, the secondary task was considered as a good candidate for measurement.

Immediately following the test, each subject was given a questionnaire to answer. The questionnaire was used to record subjective information from the subject. Similar to Garrabrants (2009) the questionnaire asked questions concerning the user's awareness of the environment, perception of workload, and the degree to which the subject was able to operate efficiently.

The subject's performance was scored in both the primary and secondary tasks. In the primary task, the subject was scored according to the amount of time the subject requires to respond to a security officer prompt when a threat occurred, as well as whether the subject chose an appropriate response according to the Dispatcher Instruction Sheet. In the secondary task, the subject was scored according to the amount of time the subject requires to acknowledge an activity record, as well as whether the subject provides the correct acknowledgment. The scores of both primary and secondary tasks were not displayed to the subject.

Further, a repeated measure of analysis of variance (RM ANOVA) was performed in order to assess the possibility that bias was introduced by repeating the experiment on each participant (that is, sequencing effects). The purpose of a RM ANOVA is to determine the degree to which the sample data may cast doubt on the null hypothesis of the analysis of variance, which focuses on whether the means differ between tests (Huck, 2008).

The dependent variables were measured as follows:

- Memory was quantitatively measured by the accuracy rate of responses in the primary task.
- Attention was quantitatively measured by the amount of time required to complete the primary task.
- Workload was quantitatively measured through accuracy rate of responses in the secondary task.

The “uncertainty on” and “uncertainty off” rounds were counterbalanced: half of the population started with the “uncertainty off” display, while the other half of the population started with the “uncertainty on” display. This step was to offset or minimize bias introduced by the ordering of each interface (sequencing effects). Because each subject was tested twice (in two rounds), the subject’s performance may improve in the second round because of additional experience with the simulation. A primary hypothesis (H_1) is that subject performance did improve with repeated testing; the null hypothesis

(H_0) is that subject performance did not improve with repeated testing. If the within-subject ANOVA results per subject are greater than the .05 confidence level, then the null hypothesis can be accepted.

Subjective results from the “uncertainty on” and “uncertainty off” rounds were measured using a questionnaire with Likert-type scales (see Appendix B). Almost half of the published articles in the field of human computer interfaces use Likert-type scales to measure the user’s qualitative experience (Kaptein, Nass, & Markopoulos, 2010). Questionnaire results were categorized according to the correlated dependent variables (memory, attention, and workload). A *t*-test is typically used to evaluate a hypothesis that deals with two means (Huck, 2008; Kaptein, Nass, & Markopoulos, 2010). Accordingly, a *t*-test was conducted to determine if responses relating to the “uncertainty on” scenario were statistically significant than responses relating to the “uncertainty off” scenario. Responses that were not statistically significant (that is, with less than a .05 confidence level) will be ignored.

Resource requirements

Software

The simulation software used was developed on the Microsoft Windows 7 platform using Microsoft Visual Studio 2010 as the Integrated Development Environment (IDE), the Microsoft .Net framework 4.0, and the C# programming language. The simulation was developed by the author of this report, who has 20 years of experience in developing Windows-based graphical user interfaces. Developmental testing of the simulation was conducted by the author.

The application recorded user inputs (both mouse and keyboard) in order to calculate response time and accuracy rates. Response time measures the duration, in seconds, between the initial display of a prompt, and when the user's input is detected in response to the prompt. If the user does not respond to a prompt within a specified time period, or the testing round completes, the application will mark the prompt as not completed by the user.

Hardware

The simulation was run on a Gateway NV54 with a Pentium Dual-Core processor running at 2.10 gigahertz, 4 gigabytes of random access memory (RAM) and Windows 7 Home Premium. The display is 15.6 inches and the screen resolution is 1366×768. The subjects were supplied with a mouse and built-in keyboard.

IRB Approval

Human subjects were used to conduct this experiment. The subjects did not require extensive experience with graphical user interfaces, surveillance, or monitoring software (Liu, Macchiarella, & Vincenzi, 2009). Institutional Review Board (IRB) approval was required for this research. The IRB Approval Memorandum is presented in Appendix C.

Summary

The purpose of this research is to use both quantitative and qualitative measurements to determine whether adaptive uncertainty visualization has a significant impact on reducing a user's cognitive load. Software visualizations displaying uncertainty characteristics may also increase the user's cognitive load (Bunch & Lloyd, 2006); there is a trade-off between the reduction in cognitive load by displaying uncertainty factors, and the increase in cognitive load by displaying additional visual elements (Antifakos, Schwaninger, & Schiele, 2004). This research builds on uncertainty research in the visualization community by adapting the display of the uncertainty aspects in software visualizations to the user's level of cognitive load (Zuk & Carpendale, 2007).

The rigorous methodology presented in this section is intended to ensure a high degree of reliability and validity. Reliability and validity was maintained through the calculation of the appropriate sample size, the selection of subjects from the population, and the use of a counterbalanced approach to ensure the order of experiments would not influence the outcome. Further, the method for collecting and analyzing the results of experiments was subjected to statistical tests to maintain a high degree of reliability and validity.

Chapter 4

Results

Findings

Three dependent variables were defined in Chapter 3: memory, attention, and workload. The dependent variables will be evaluated separately in order to assess the impact of adaptive uncertainty visualization on subjects with respect to each variable. Measurements were collected from seven primary tasks and three secondary tasks, including response time to complete each task, and the accuracy rate. The measurements contributed to the evaluation of each dependent variable; for example, measurement of the subject's response time was used to assess the subject's attention level. In addition, survey results were used to provide a qualitative assessment of each independent variable. Appendix D contains the table of quantitative measurements collected for each subject during the two successive runs of the study, while the table in Appendix E summarizes the qualitative survey results collected after the completion of the study for each subject, as well as all comments provided by the subjects.

As described in Chapter 3, the study was conducted on each subject twice: once with, and once without the adaptive uncertainty visualization. The order in which each visualization was scheduled was alternated in order to counterbalance any sequencing effects, such as learning, or fatigue. In order to determine whether repeated exposure to the study affected a subject's performance in terms of accuracy rates or response time, a repeated measures analysis of variance (RM ANOVA) was conducted using the quantitative measurements. RM ANOVA analysis was conducted using Minitab version 15, and the results of the analysis appear in the form of a Minitab session log in Appendix F.

The RM ANOVA analysis evaluates two factors – between-subjects and within-subjects. The within-subjects factor reflects the measurement of the dependent variable across all conditions for each subject, while the between-subjects factor reflects the measurement of the dependent variable across all subjects.

To evaluate the influence of sequencing effects, the hypothesis (H) is asserted that performance and accuracy (as measured by response time and accuracy rate) changed significantly, when the subject repeated the study, while the null hypothesis (H_0) is asserted that performance and accuracy did not change significantly. If the subject *p* values for within-subjects results is less than the confidence level alpha (α), which is defined as .05, the null hypothesis is rejected; if the subject *p* value exceeds the confidence level, then the null hypotheses is accepted. Within-subjects analysis was conducted on the measurements that contributed to the evaluation of each dependent variable, and the results appear in tables later in this chapter.

Finally, a *t*-test was conducted on the qualitative survey results to determine whether the responses regarding the uncertainty visualization display were significantly different from the responses regarding the standard display. The results of the analysis appear in the form of a Minitab session log in Appendix F. Minitab was also used to produce the charts and graphs included in this chapter.

About the Sample

Data was collected from a sample consisting of 24 subjects who used a computer for business or academic purposes for a period of two or more years, and for two or more hours a day. The sample was evenly divided among male and female participants (that is, 12 male and 12 female subjects). Subject age ranged from 18 to 62, with a mean age of 38. In addition, the subject's education ranged from subjects with high school only, to others with doctoral degrees, with a mean of 3.7 (just less than a 4-year degree). Education was scaled according to the number of years of college each subject had completed following high school; for example, three of the subjects had completed two years of college, while two of the subjects had completed eight years of college. The distribution of values for subject age and education are presented in Figure 7 and Figure 8.

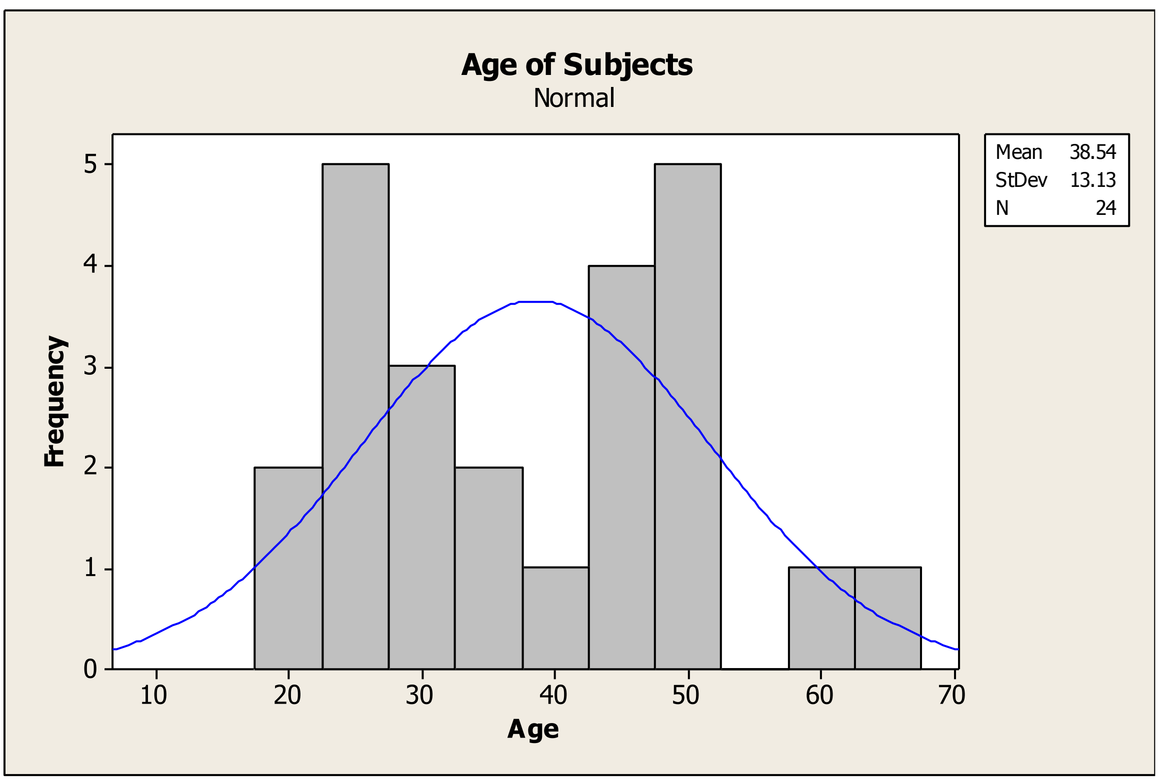


Figure 7. Histogram. Age of Participants

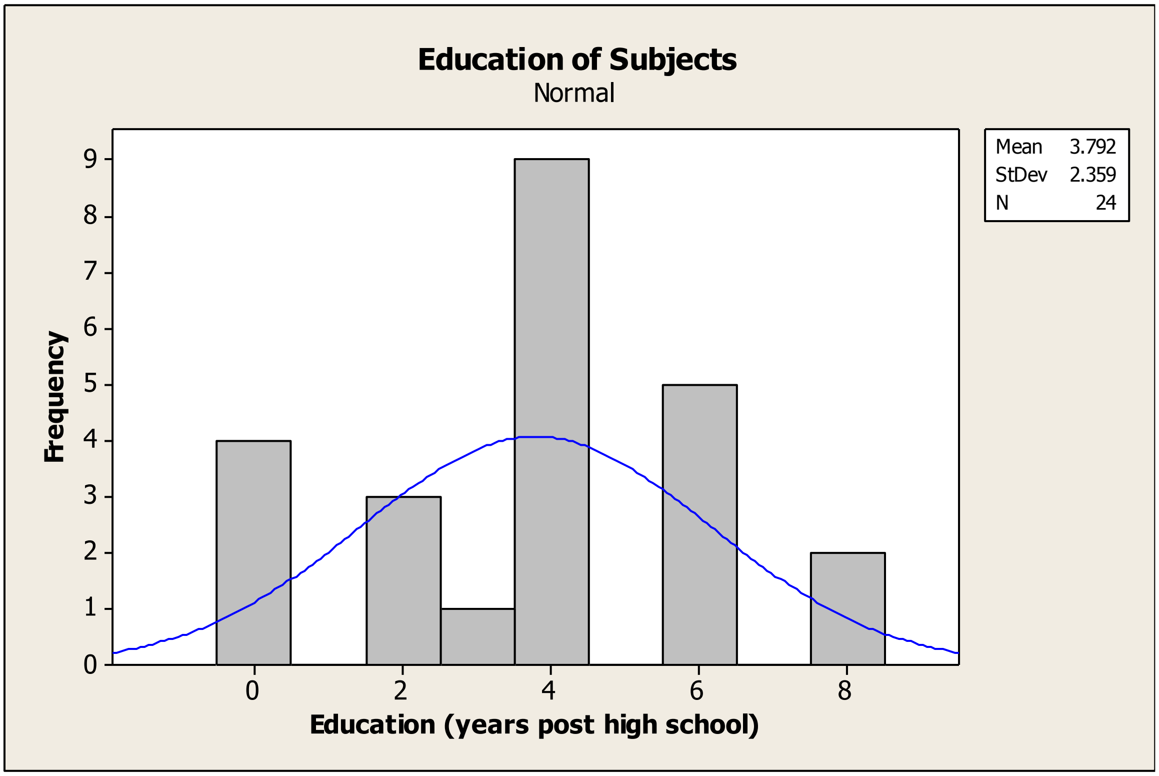


Figure 8. Histogram. Education of Subjects (years post high school)

Analysis of Memory Factors

The effects of the interface on memory was measured by the accuracy rate for completing seven primary tasks. The accuracy rate represented the number of correct responses the subject provided in completing each of the seven primary tasks. In order to provide an accurate response, the subject needed to remember how many times a particular agent in the simulation had received a warning for entering a restricted area. As outlined in the Dispatcher Instruction Sheet (included in Appendix A), the subject was to issue a warning the first two times an agent entering a restricted area, and request backup to remove the agent on the third attempt.

Table 3 lists the *p* values for both the between-subjects and within-subjects factors for the primary tasks examined in the memory analysis. As the table indicates, the within-subjects p values for all factors are greater than the *alpha* (α) of .05, supporting the null hypothesis that sequencing effects arising from repeated exposure to the study were not a significant factor. Accordingly, the factors can be evaluated to determine the effects of uncertainty visualization on the subject's memory.

Table 3 also lists the between-subject p values for tasks A-G. The between-subjects measurement assesses the effect of the interface on the subject's memory. As the table indicates, the between-subjects *p* values for tasks B, D, E, F, and G are less than the *alpha* (α) of .05, so the null hypothesis (H_0), that the interface did not significantly influence the subject's memory, is rejected. The hypothesis that the interface significantly influences the subject's memory is accepted. Appendix F lists the complete RM ANOVA results for the primary task in the form of the Minitab session log.

Table 3

Probability (*p*) Values for Memory Factors

Between-Subjects and Within-Subjects Accuracy Rate

Factor	Between Subjects (*p*)	Within Subjects (*p*)
Task A	0.328	0.630
Task B	0.029*	0.158
Task C	0.539	0.318
Task D	0.001*	0.216
Task E	0.004*	0.144
Task F	0.015*	0.824
Task G	0.002*	0.332

* indicates a *p* value < .05

Figure 9 and Figure 10 display the distribution of values for subject accuracy rate in the primary task for both the standard interface, and the uncertainty visualization interface. The mean accuracy rate for the uncertainty visualization display was significantly higher than the mean accuracy rate for the standard display, while the variance was significantly reduced.

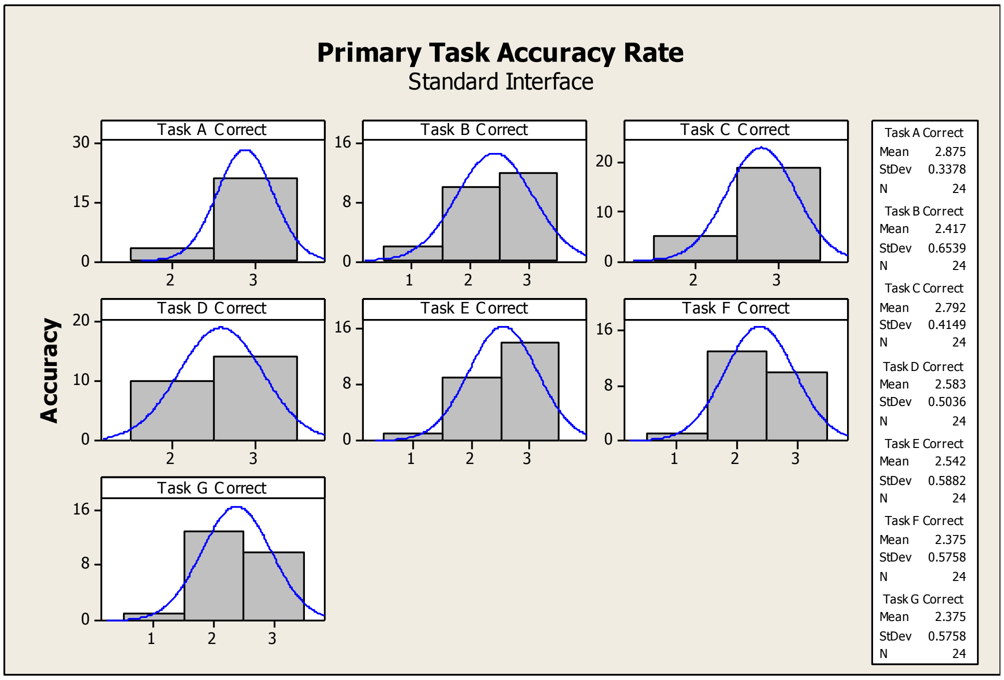


Figure 9. Histogram. Primary Task Accuracy. Standard Interface

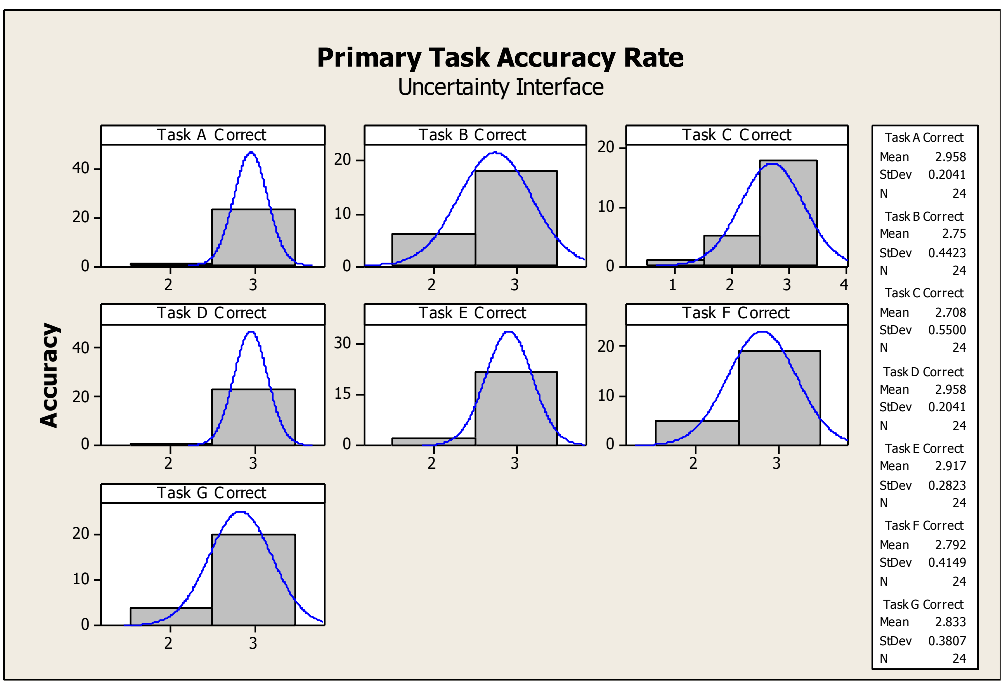


Figure 10. Histogram. Primary Task Accuracy. Uncertainty Interface

Figure 11 compares the mean accuracy rate for several representative primary tasks measured with both the standard display, as well as the uncertainty visualization display. Inspection of the graph illustrates that the uncertainty visualization interface improved the accuracy rate for completing the primary task. As Figure 11 illustrates, the mean accuracy for task C was slightly higher under the standard interface than the uncertainty interface, whereas the uncertainty interface improved the accuracy rate for the remaining tasks. Participant accuracy rates for task C may have differed from other tasks due to vagaries in the scheduling of tasks in the simulation; while other tasks were executed concurrently, dividing the operator's attention among multiple tasks, task C executed when no other tasks were running, reducing the burden on the operator's cognitive load.

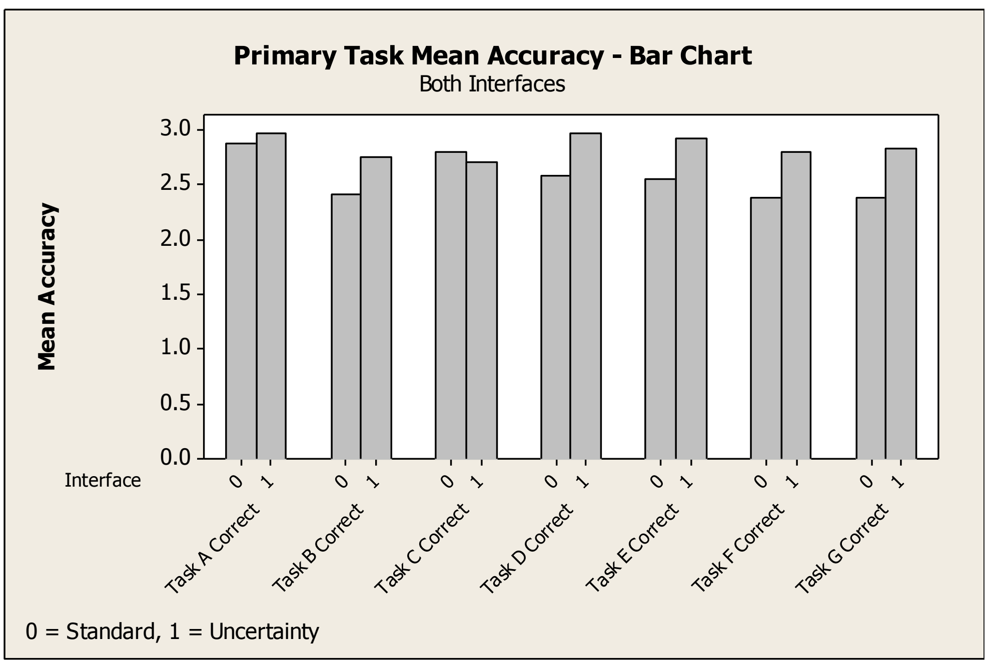


Figure 11. Bar Chart. Primary Task Mean Accuracy. Both Interfaces

Appendix B lists the survey questions asked of each subject immediately following the two test runs. Responses ranged from a value of one ("Strongly Disagree") to five ("Strongly agree"). Questions 2, 5, 7, 8, 13, 17, and 18 addressed the subject's perception of how the interface influenced their ability to remember the status of the agents in the simulation, and thus their accuracy rate in responded to prompts. A *t*-test was conducted on the survey responses, and the results are listed in Appendix F. Results greater than the *alpha* (α) of .05 indicated that there was no statistical significance between subject responses concerning the standard interface versus the uncertainty visualization interface. Results for question 13 were greater than the *alpha* (α) of .05, so responses for these questions were not considered. Responses for the remaining questions are listed in Table 4.

Table 4

Mean Responses for Survey Questions for Memory

Memory

Question	Standard Interface	Uncertainty Interface
2. Easily distinguish between critical and non-critical tasks	3.3	4.6
5. Knew the status of all individuals	2.8	3.9
7. Confident about decisions	3.1	4.0
8. Able to make good decisions	3.7	4.4
17. Made fewer errors	2.0	3.6
18. Easy to track individuals and remember who they were	2.9	4.1

The response scale is 1-5, where 1=strongly disagree, 5=strongly agree

Both quantitative and qualitative results conclusively demonstrated that the subject's memory was significantly improved in the uncertainty visualization display as opposed to the subject's ability to recall when using the standard display. The demand on the subject's memory was higher in standard display, resulting in a lower accuracy rate.

Analysis of Attention Factors

The effects of the interface on attention was measured by the response times for completing seven primary tasks. The response time represented the duration, in seconds, between the time a response was presented to the subject, and the time the subject responded in completing each of the seven primary tasks.

Table 5 lists the *p* values for both the between-subjects and within-subjects factors for the primary tasks examined in the attention analysis. As the table indicates, the within-subjects p values for A, B, C, D, E, and F factors are greater than the *alpha* (α) of .05, supporting the null hypothesis that sequencing effects arising from repeated exposure to the study were not a significant factor. Accordingly, the factors can be evaluated to determine the effects of uncertainty visualization on the subject's attention.

Table 5 also lists the between-subject p values for tasks A-G. The between-subjects measurement assesses the effect of the interface on the subject's level of attention. As the table indicates, the between-subjects *p* values for tasks A, B, D, E, and F are less than the *alpha* (α) of .05, so the null hypothesis (H_0), that the choice of interface did not significantly influence the subject's level of attention, is rejected. The hypothesis that the choice of interface significantly influences the subject's attention is accepted. Appendix F lists the complete RM ANOVA results for the primary task in the form of the Minitab session log.

Table 5

Probability (*p*) Values for Attention Factors

Between-Subjects and Within-Subjects Response Time

Factor	Between Subjects (*p*)	Within Subjects (*p*)
Task A	0.000*	0.462
Task B	0.003*	0.185
Task C	0.681	0.637
Task D	0.012*	0.764
Task E	0.017*	0.646
Task F	0.002*	0.881
Task G	0.630	0.011*

* indicates a p value < .05

Figure 12 and Figure 13 display the distribution of values for subject response time in the primary task for both the standard interface, and the uncertainty visualization interface. The mean response time for the uncertainty visualization interface was significantly lower than the mean response time for the standard interface, while the variance was significantly reduced.

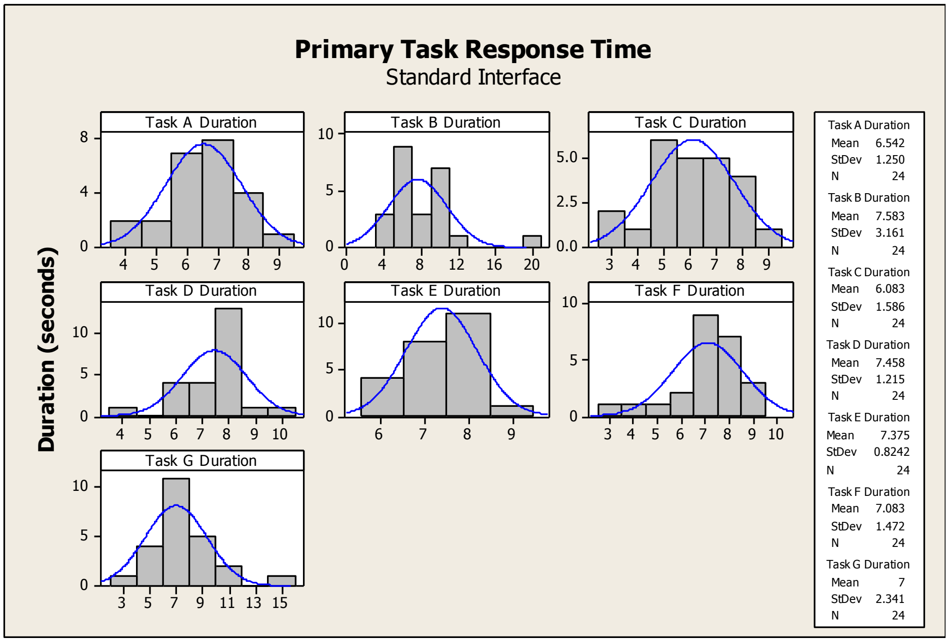


Figure 12. Histogram. Primary Task Response Time. Standard Interface

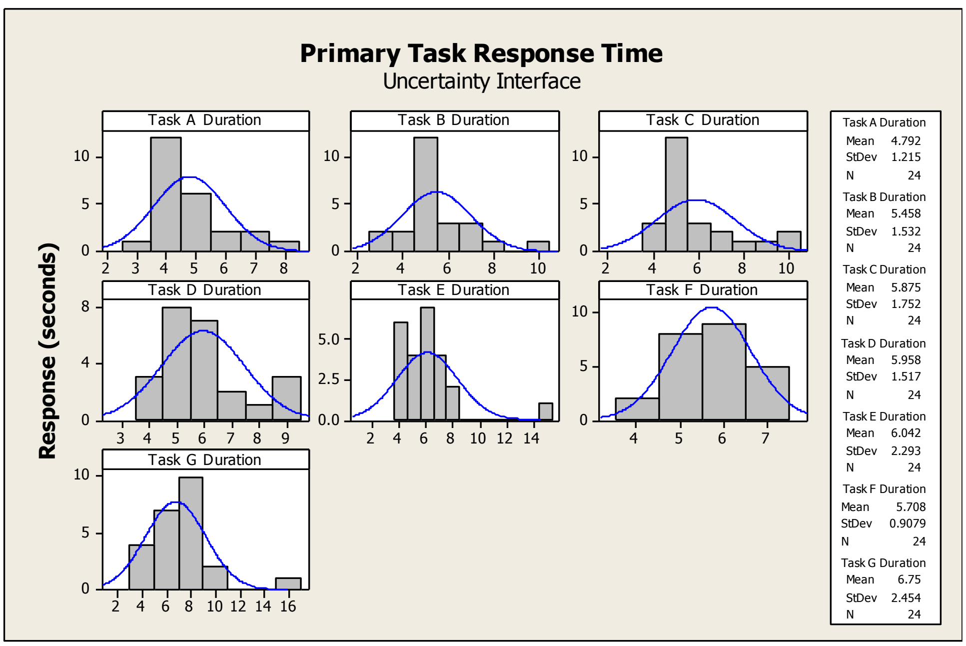


Figure 13. Histogram. Primary Task Response Time. Uncertainty Interface

Figure 14 compares the mean response times for several representative primary tasks measured with both the standard display, as well as the uncertainty visualization display. Inspection of the graph illustrates that the uncertainty visualization interface reduced the response time for completing the primary task.

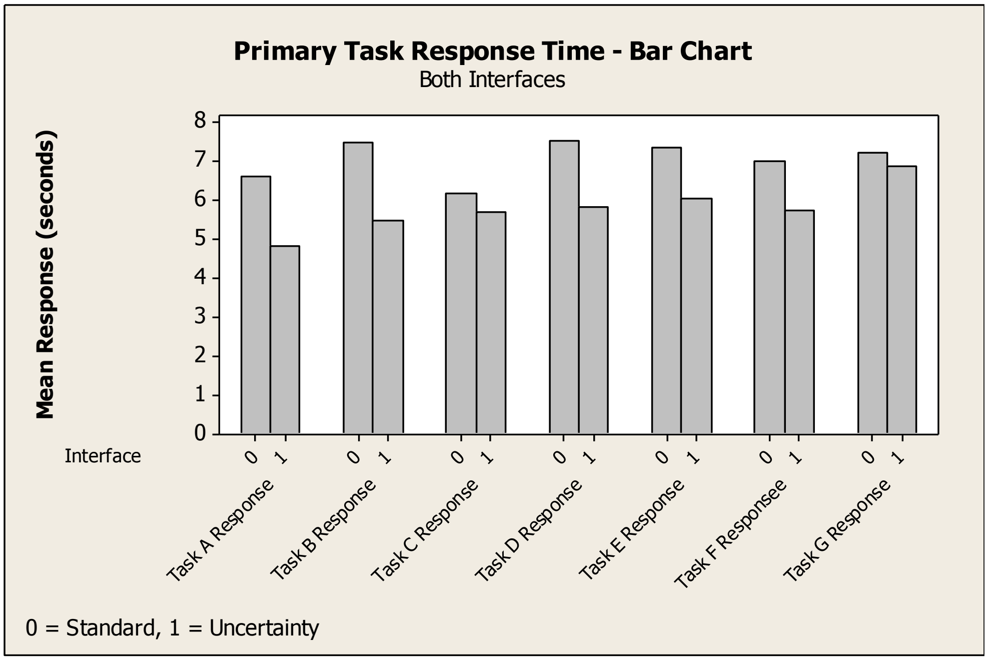


Figure 14. Bar Chart. Primary Task Mean Response Time. Both Interfaces

Appendix B lists the survey questions asked of each subject immediately following the two test runs. Responses ranged from a value of one ("Strongly Disagree") to five ("Strongly agree"). Questions 1, 3, 6, 11, 14, 19, and 20 addressed the subject's perception of how the interface influenced their ability to stay attentive to assigned tasks, and thus their response time in responded to prompts. A *t*-test was conducted on the survey responses, and the results are listed in Appendix F. Results greater than the *alpha* (α) of .05 indicated that there was no statistical significance between subject responses concerning the standard interface versus the uncertainty visualization interface. Results for questions 11, 14, and 20 were greater than the *alpha* (α) of .05, so responses for these questions were not considered. The responses for the remaining questions are listed in Table 6.

Table 6

Mean Responses for Survey Questions for Attention

Response Time

Question	Standard Interface	Uncertainty Visualization Interface
1. Knew when a situation required attention	3.6	4.9
3. Could concentrate on critical decisions	3.7	4.4
6. Easy to switch between tasks	3.2	4.1
19. Movement improved target detection	3.0	4.8

The response scale is 1-5, where 1=strongly disagree, 5=strongly agree

Both quantitative and qualitative results conclusively demonstrated that the subject's level of attention was significantly improved in the uncertainty visualization display as opposed to the subject's ability to recall when using the standard display. The demand on the subject's attention was higher in standard display, resulting in a lower response time.

Analysis of Workload Factors

The effects of the interface on workload was measured by the accuracy rate for completing three secondary tasks. The accuracy rate represented the number of correct responses the subject provided in completing each of the three secondary tasks. In order to provide an accurate response, the subject needed to read and comprehend a notification in the simulation indicating a security officer had arrived for, or departed the security post. As outlined in the Dispatcher Instruction Sheet (included in Appendix A), the subject was to report any late arrival or early departure; approve any early arrival or late departure; and approve any on-time arrival or departure.

Table 7 lists the *p* values for both the between-subjects and within-subjects factors for the secondary tasks examined in the workload analysis. As the table indicates, the within-subjects p values for all secondary factors are less than the *alpha* (α) of .05, supporting the hypothesis that sequencing effects arising from repeated exposure to the study were a significant factor. Accordingly, the factors cannot be evaluated to determine the effects of uncertainty visualization on the subject's workload.

Table 7 also lists the between-subject p values for all secondary tasks. The between-subjects measurement assesses the effect of the interface on the subject's workload. As the table indicates, the between-subjects *p* values for all secondary tasks exceed the *alpha* (α) of .05, so the hypothesis that the choice of interface significantly influences the subject's attention is rejected. Appendix F lists the complete RM ANOVA results for the secondary task in the form of the Minitab session log.

Table 7

Probability (*p*) Values for Workload Factors

Between-Subjects and Within-Subjects Accuracy Rate

Factor	Between Subjects (*p*)	Within Subjects (*p*)
Task 1	0.317	0.000*
Task 2	0.580	0.000*
Task 3	0.150	0.000*

* indicates a *p* value < .05

Figure 15 and Figure 16 display the distribution of values for subject accuracy rate in the secondary task for both the standard interface, and the uncertainty visualization interface.

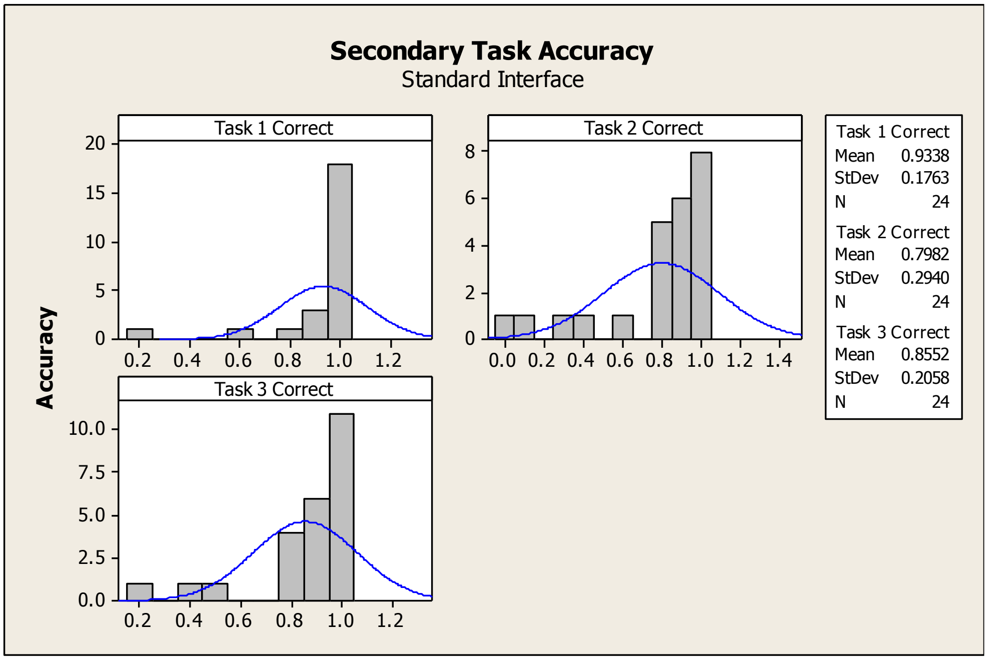


Figure 15. Histogram. Secondary Task Accuracy. Standard Interface

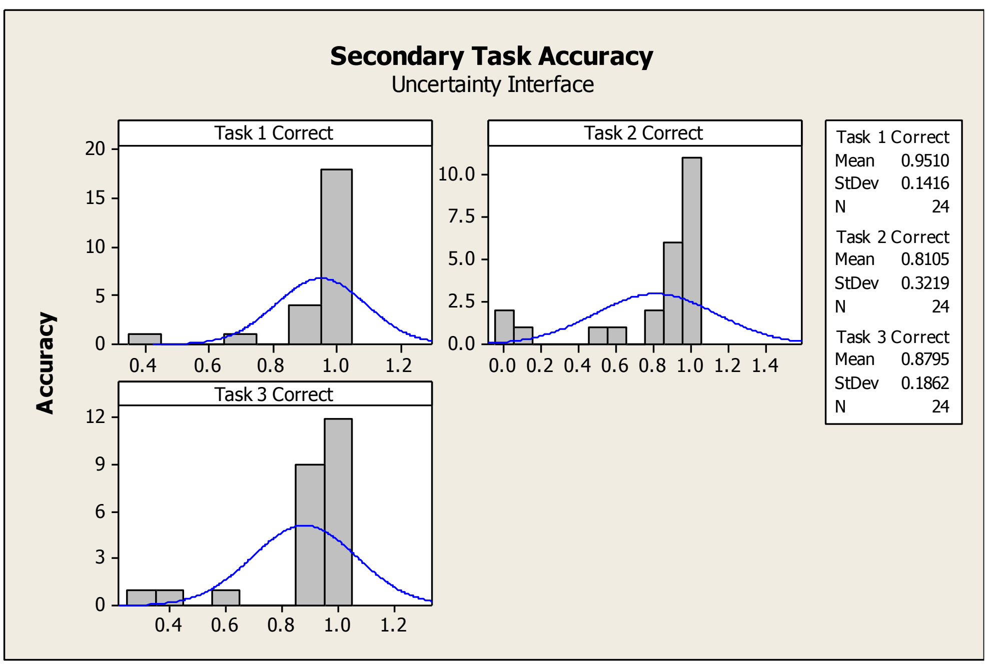


Figure 16. Histogram. Secondary Task Accuracy. Uncertainty Interface

During the simulation study, the researcher observed a number of subjects execute "strategies" using the secondary tasks. For example, one subject ignored secondary tasks that described a security officer as arriving or departing late or early. These tasks were two-factored tasks, because the subject first had to comprehend whether the officer was arriving or departing; then the subject had to comprehend whether the officer was early or late. This task was computationally more extensive than the on-time task, which only required that the subject determine whether the officer arrived on time. By concentrating on the on time tasks, the subject was able to increase the number of correct responses for that task category, but at the expense of the other categories. Another subject chose to ignore all of the secondary tasks when two or more agents were present on the screen (as part of the primary task). As a result, the accuracy rate and response time variance was high for all secondary tasks. This was manifested in very low *p* values for the within-subjects factors.

Appendix B lists the survey questions asked of each subject immediately following the two test runs. Responses ranged from a value of one ("Strongly Disagree") to five ("Strongly agree"). Questions 4, 9, 10, 12, 15, and 16 addressed the subject's perception of their situational awareness, and thus their accuracy rate in reading and comprehending prompts in the secondary task. A *t*-test was conducted on the survey responses, and the results are listed in Appendix F. Results greater than the *alpha* (α) of .05 indicated that there was no statistical significance between subject responses concerning the standard interface versus the uncertainty visualization interface. Results for questions 12, and 16 were greater than the *alpha* (α) of .05, so responses for these questions were not considered. The responses for the remaining questions are listed in Table 8.

Table 8

Mean Responses for Survey Questions for Workload

Question	Standard Interface	Uncertainty Visualization Interface
4. Prioritize critical and non-critical tasks	3.2	4.5
9. Able to make good decisions during active times	3.4	4.4
10. Able to make good decisions during critical events	3.2	4.2
15. Certainty about actions required	3.1	4.2

The response scale is 1-5, where 1=strongly disagree, 5=strongly agree

While quantitative results did not support the hypothesis, qualitative results conclusively demonstrated that the subject's workload was significantly improved in the uncertainty visualization display as opposed to the subject's ability to recall when using the standard display. The demand on the subject's workload was higher in standard display, resulting in a lower response time.

Summary of Results

A study was conducted using 24 subjects who were identified as knowledge workers who use computers each day for academic or professional purposes. The sample was evenly divided between male and female subjects, and ages ranged from 18 to 62 years. A computer simulation was chosen for the study to maintain a constant set of independent variables across all subjects. The computer simulation was designed to increase the subjects' cognitive load in order to determine whether adaptive uncertainty visualizations significantly affected dependent variables, such as memory, attention, and workload that represented the cognitive load each subject experienced. The dependent variables were assessed quantitatively by measuring the response time and accuracy rates of completing simulation tasks, and qualitatively by examining responses to a survey conducted at the end of each experiment. Because subjects were tested twice (once with a standard interface and once with the uncertainty visualization interface) results were validated using repeated measures of analysis of variance (RM ANOVA) to determine whether subjects were influenced by sequencing effects.

The results indicated that the cognitive burden on the subjects' memory, attention, and workload was significantly diminished when using the uncertainty visualization interface, in contrast to when the subjects used the standard interface. Subjects were able to complete tasks in less time, and with a higher accuracy rate. Further, the subjects perceived they were more effective when using the uncertainty visualization interface, in contrast to when the subjects used the standard interface

Chapter 5

Conclusions, Implications, Recommendations, and Summary

Conclusions

The study simulated an environment with an elevated level of activity in order to stress the subject's cognitive load. A primary task exercised the subject's ability to recall the status of up to four individual agents at a time, while requiring the subject to scan visually a region in order to detect when agents entered certain restricted areas. A secondary task exercised the subject's ability to read and comprehend alerts. Because both tasks were presented concurrently, the subjects were required to exercise both divided and sustained attention.

An analysis of the research using the computer simulation demonstrated that subjects were able to perform their primary tasks with a higher accuracy rate using the adaptive uncertainty display than when using a standard display. In addition, subjects were able to complete the primary tasks in less time using the adaptive uncertainty display. This finding was reinforced by qualitative results from a survey conducted immediately after each simulation test; subjects accorded the adaptive uncertainty display a higher degree of user satisfaction than the standard display.

A repeated measure of analysis of variance (RM ANOVA) indicated that sequencing effects significantly influenced the subject's accuracy rate and response time in completing the secondary tasks. Nonetheless, subjects were able to complete the secondary tasks in substantially less time with a similar level of accuracy. For example, the mean response time for completing Task A using the standard interface was 17.05 seconds, and with the uncertainty interface, the mean response time was 12.94 seconds.

Analysis of Research Questions

Chapter 1 presented the hypothesis that adaptive uncertainty visualization significantly reduce a user's cognitive load in an environment where both stress and uncertainty abound. The hypothesis (H) asserted that knowledge workers exhibit better performance and improved decision-making using adaptive uncertainty visualization than when a standard interface is employed. The null hypothesis (H_0) was that knowledge workers exhibit no better performance or improved decision-making using adaptive uncertainty visualization than when a standard interface was employed.

Chapter 1 also presented the following research questions:

1. Does adaptive uncertainty visualization improve the system operator's level of performance in completing assigned tasks?
2. Does adaptive uncertainty visualization improve the system operator's level of attention in handling multiple activities?
3. Does adaptive uncertainty visualization reduce the burden on the system operator's workload?

The purpose of the first research question was to determine whether the uncertainty interface improved performance factors, such as memory. The computer simulation tested the subject's memory by requiring the subject to remember the status of up to four agents at a time, and to provide a response in accordance with simple rules. Both the quantitative data, as measured by the subject's accuracy rate, and the qualitative responses from survey questions, indicate the subjects were able to recall the status of all agents at a higher rate or accuracy using the uncertainty display than with the standard display.

The second question was related to the user's level of attention. Because the subject was challenged with two concurrent tasks, the subject's attention was divided by the demands of each task. Consequently, when the simulation displayed an alert that required the subject's attention, the subject had to divert attention from their current point of focus to the section of the screen where the alert was displayed. The amount of time required for the subject to regain focus on a target area was reflected in the response time for providing a response to a particular task. Both the quantitative data, as measured by the subject's response time, and the qualitative responses from survey questions, indicate the subjects were able to regain focus on a target area with a significantly lower response time using the uncertainty display than with the standard display.

The final question was concerned with the user's workload. The computer simulation tested the subject's workload by requiring the subject to read and comprehend the text of an alert as part of the secondary task, and to provide a response in accordance with simple rules. A repeated measure of analysis of variance (ANOVA) indicated that subject performance on the secondary task was unduly influenced by sequencing events. However, the qualitative responses from survey questions indicate the subjects had greater comprehension of the text of alert notifications using the uncertainty display than with the standard display.

The study clearly supports the hypothesis that adaptive uncertainty visualization significantly reduces a user's cognitive load in an environment where both stress and uncertainty abound. Subjects exhibit better performance and improved decision-making using adaptive uncertainty visualization than a standard interface as measured by memory, attention, and workload. Accordingly, the null hypothesis (H_0) that knowledge workers exhibited no better performance or improved decision-making using adaptive uncertainty visualization than when using a standard interface can be rejected.

Implications

The proliferation of large-area networks, inexpensive sensors, and mobile devices has increased demand for context-aware applications that integrate the user with the environment in which the user is located; yet data from the environment can be faulty and unpredictable (Mühlhäuser & Hartmann, 2009). The combination of increased data flow, and greater and more diverse types of uncertainty, poses a challenge for user experience designers (Santos, Cardoso, Diniz, & Ferreira, 2010).

The unique characteristics of real-world settings can impose a cognitive burden on decision-makers that reduces the effectiveness of their decisions. These characteristics include time constraints, high stakes, and ill-structured problems. Decision-makers adapt strategies in order to cope with the effects of these uncertainties, using such techniques as reduction (to reduce the level of uncertainty), forestalling (preparing a contingency plan), and suppression (that is, increase risk-taking by ignoring the effects of uncertainty). Further, novices and experts respond with different strategies; for example, experts are more likely to build stories to account for phenomena, while novices are more likely to use checklists (Atoyan, Robert, & Duquet, 2011).

Emerging applications for pervasive, context-aware systems include emergency first-responder services, military, security, and disaster relief scenarios (Arabo, Shi, &

Merabti, 2011). Each of these scenarios require humans to make high stake decisions in an environment with a high degree of uncertainty. For example, military personnel are increasingly deployed to combat terrorism, counterinsurgencies, or for extended peacekeeping operations. These engagements are often protracted, with diffuse source of threats, and ill-defined rules of engagement. Soldiers are often fatigued, subject to numerous types of stress, and suffer limited cognitive functioning. The combination of uncertain situational factors and psychological factors can have adverse consequences for soldiers. and the successful fulfillment of their mission (Sharma & Sharma, 2012).

First-responders operate in a high stake environment with a compressed timeframe, where lives may be lost in a few minutes. Emergency medical technicians (EMTs) form a mental model prior to arriving at the scene of an emergency that is based on information provided by a dispatcher, and the EMT's own past experiences. Once the EMT arrives on the scene, additional information from onlookers, the victim, or the EMT's own observations refine and shape the mental model. The EMT then develops a response, called a situated action that is more reflective of unforeseeable contingencies rather foreseeable outcomes. Finally, the EMT makes decisions about care and treatment of the victim that balance the exigency of providing immediate care with the urgency of transporting the victim to a hospital or emergency unit. While formation of the mental model is critical to timely and accurate decision making, the EMT may be hampered by a number of factors inherent to emergencies. For example, situational stress can degrade the EMT's perception of the mental model; the EMT may have difficulty transforming the mental model of past situations to the present; unforeseen contingencies may cause deviations from mental models (Rahman, 2012).

The current study demonstrated the hypothesis that adaptive uncertainty visualization significantly reduces a user's cognitive load in an environment where both stress and uncertainty abound. This hypothesis has clear implications for situations where participants are subject to high cognitive load, such as soldiers on the battlefield or on peacekeeping missions, or first-responders in emergencies. There are many scenarios where a human decision maker could benefit from having visibility to uncertainty factors; for example, a soldier on a peacekeeping mission may determine that a sensor providing stale or out-of-date information has been tampered with or disabled, indicating the presence of insurgents. A fire official at the scene of a fire may revert to line-of-sight verification when an application showing the interior of a building indicates a low degree of confidence in presenting the structural integrity of interior walls. Sacrificing the presentation of uncertainty factors in order to spare a system user from additional burdens on cognitive load may also deprive the user of essential inputs that could improve decision-making, or enable intuitive heuristics or compensatory actions that lead to problem solving.

Adaptive uncertainty visualization provides a mechanism for modulating the effects of computer visualizations on the user's cognitive load. When an uncertainty visualization is added to a screen, other less-essential visualizations are removed in order to offset the additional cognitive burden. That is, the salience of essential elements is increased by muting the display of non-essential elements. An element is considered essential if presentation contributes to achieving a goal. For example, inanimate objects in the background of an image can be depicted with a wireframe rather than a textured image; or background colors can be muted to sharpen the contrast between items of interest and inanimate objects. By reducing the cognitive burden or other, less-essential visualizations, uncertainty visualizations can be added to the screen without unduly burdening the user's cognitive load. Accordingly, the decision maker can improve awareness of the environment with less danger of impairment from an increased cognitive burden.

Recommendations

While this study conclusively demonstrates that adaptive uncertainty visualization can significantly reduce cognitive load, the nexus of cognitive load and uncertainty visualization is fertile with areas for further investigation and refinement. The study was confined to a sample of 24 knowledge workers; this population is subject to information overload (Karr-Wisniewski & Lu, 2010) and high levels of stress and anxiety that can lead to high degrees of cognitive load (Kirsch, 2000). The knowledge worker classification encompasses a large population with a great degree of variety in terms of background, capabilities, and motivation. This can lead to a large variability in experiment performance, which can reduce the reliability of research findings. For example, two of the research subjects remarked that the simulation was "too slow", while two others remarked that the simulation was "too fast". While knowledge workers frequently interact with visualization tools for decision-making (Reinhardt, Schmidt, Sloep, & Drachsler, 2011) there are cognitive load factors unique to real-world settings, such as time compression and an ability to form a mental model (Atoyan, Robert, & Duquet, 2011). As a result, the findings of this study may not be generalized to other populations or situations, such as a combat setting, or with first-responders, since the cognitive load characteristics of these settings are unique (Arabo, Shi, & Merabti, 2011);

Rahman, 2012). Accordingly, an important area of future study would be to test adaptive uncertainty visualization in simulated combat or first-responder settings.

Another important area of future study is adaptable versus adaptive modulation of the cognitive effects of visualizations. Atoyan et al. (2011) define an adaptive system as one in which the system makes the decision to automate activities on behalf of the operator, while an adaptable decision allows the operator to grant or revoke the automation. Certain aspects of adaptive systems can increase uncertainty (such as when adaptive effects are unpredictable) or lower user acceptance (when the user resents or is frustrated by automation). Adaptable systems can increase user acceptance and foster trust in the system. In the context of uncertainty visualization, an adaptable system could allow the operator to decide which visual elements are cognitively burdensome but add little to decision-making, or which techniques to use to minimize the cognitive burden of non-essential visualizations. For example, the operator could choose to shade a moving object or reduce the refresh interval so the object's movement is not continuous. Adaptable uncertainty visualization methods could also be applied to the emerging field of augmented cognition, where a system modulates the type and volume of information provided to an operator based on an assessment of the operator's state (Juhnke, Mills, & Hoppenrath, 2007).

Confidence levels exert a strong influence on the operator's acceptance of a system (Cohen & Warren, 1990). There is a threshold of system confidence that influences the operator's acceptance of the system; below the threshold, it may be harmful for the system to disclose factors affecting uncertainty. A confidence threshold for non-critical applications is within an 80-90% confidence range, but this range may vary according to the situation (Lim & Dey, 2011). In the current study, the confidence level was fixed at 85% as a control variable. Consequently, the study did not examine the influence of low or varying confidence levels on cognitive load. Low or varying confidence levels could be cognitively burdensome to the operator, an outcome that could negate the benefits of displaying uncertainty factors.

Several subjects in the study observed that the simulation did not sufficiently suppress non-essential elements. While the display of a number of distractors was suppressed during simulated outages (such as passing vehicles), other non-critical elements were not suppressed (such as neutral targets that were not moving). According to the rules of the simulation, both non-moving objects and passing vehicles contributed little to decision-making (since they did not enter restricted areas) but because the weight of the cognitive assigned to movement was higher than stationary objects, the simulation suppressed vehicles rather than neutral objects. Further study on classification of the cognitive burden of visual elements would support a more sophisticated algorithm for determining which visual elements to modulate and in which priority, in order to reduce or maintain the same level of cognitive load. Using the same approach to classifying the cognitive burden of uncertainty visualizations support a more fine-grained approach to offsetting the additional cognitive burdens of uncertainty by reducing the cognitive burden of non-essential elements. Further research could identify whether a multi-faceted approach to the classification of cognitive burden of visual elements is appropriate (for example, by classifying a composite cognitive burden of different facets of a visual element, such as its movement, size, color, or brightness).

Humans are successful in solving complex problems in multimodal environments every day. For example, the task of driving involves a combination of visual, auditory, and haptic stimuli. Multimodal stimuli can have a detrimental effect in high workload environments, but for some cognitively challenging tasks, such as vehicle navigation, audiovisual stimuli can improve performance (Sigrist, Rauter, Riener, & Wolf, 2012). This finding is consistent with the multiple resource model proposed by Wickens (2002). Examples of auditory signals include auditory alarms, which alert an operator to an urgent condition, or sonification of system variables, where the volume or tone of a sound is modulated according to system parameters. The current study is unimodal, focusing solely on visual stimuli. Future research could explore the other modalities in which uncertainty can be represented (for example, haptically), or how cognitively burdensome tasks could be offloaded to underused modalities (such as by switching to auditory stimuli when the visual senses are overloaded).

Summary

Uncertainty is inherent in many real-world settings; for example, an emergency medical technician arriving on the scene of an emergency may have incomplete or inaccurate information regarding the source or severity of injuries, preventing the technician from making a valid assessment concerning treatment. In an environment plagued with uncertainty, decision-support systems, such as sensor-based networks, may make faulty assumptions about field conditions, especially when information is incomplete, or sensor operations are disrupted. Because humans are adept at problem solving under uncertain conditions (Tversky & Kahneman, 1974) the presentation of uncertainty can lead to more effective decision-making.

In stressful settings, such as a battlefield or an air traffic control tower, the user's cognitive load is already strained. Visualizing uncertainty elements add to the visual clutter that competes for the operator's limited attention. The operator must invest increased effort to process probabilistic assessments. Consequently, cognitive load increases, degrading the operator's performance and problem solving effectiveness.

A considerable body of research exists to explain the cognitive and perceptual factors that influence a person’s effectiveness as a decision maker in diverse environments. While humans are adept at problem solving in challenging conditions, innovations in computer design have provided additional tools to aid people in making decisions. For example, visualization techniques have been used to generate a probabilistic model of an environment in order to encourage effective problem solving. The danger of these visualization techniques is that humans may ignore the probabilistic nature of the information they receive, and make false or unsupported inferences. Accordingly, it is critical to provide cues to the user that information is projected or estimated, and explain the source or degree of the unreliability of the information. Uncertainty then becomes another factor in the decision-making process. However, in stressful situations, uncertainty visualization can have diametric effects. While uncertain conditions can increase the cognitive load on a decision maker, so can the burden of added visualizations and decision points.

Given the importance of uncertainty in effective decision-making, there is a critical need for research that demonstrates how uncertainty visualization can be used without straining the operator's cognitive load (Antifakos, Schwaninger, & Schiele, 2004). This research demonstrated that an adaptive visual system could provide relevant visualization of uncertainty to improve decision-making without further straining the operator's cognitive load. This research contributes to the body of knowledge by identifying techniques to increase the use of uncertainty visualization in stressful environments without increasing the user's cognitive load.

This research used both quantitative and qualitative techniques to measure whether adaptive uncertainty visualization has a significant impact on reducing a user's cognitive load. Since software visualizations that displaying uncertainty characteristics may also increase the user's cognitive load (Bunch & Lloyd, 2006); there is a trade-off between the reduction in cognitive load by displaying uncertainty factors, and the increase in cognitive load by displaying additional visual elements (Antifakos, Schwaninger, & Schiele, 2004). The research builds on uncertainty research in the visualization community by adapting the display of the uncertainty aspects in software visualizations to the user's level of cognitive load (Zuk & Carpendale, 2007).

In this study, 24 subjects (identified as knowledge workers who use computers each day for academic or professional purposes) were subjected to a computer simulation. The sample was evenly divided between male and female subjects, and ages ranged from 18 to 62 years. The computer simulation was chosen as the model for testing the hypothesis of the study to maintain a constant set of independent variables across all subjects. The computer simulation was designed to increase the subjects' cognitive load in order to determine whether adaptive uncertainty visualizations significantly affected dependent variables, such as memory, attention, and workload that represented the cognitive load each subject experienced. The dependent variables were assessed quantitatively by measuring the response time and accuracy rates of completing simulation tasks, and qualitatively by examining responses to a survey conducted at the end of each experiment. Because subjects were tested twice (once with a standard interface and once with the uncertainty visualization interface) results were validated using repeated measures of analysis of variance (RM ANOVA) to determine whether subjects were influenced by sequencing effects.

The results indicated that the cognitive burden on the subjects' memory, attention, and workload was significantly diminished when using the uncertainty visualization interface, in contrast to when the subjects used the standard interface. Subjects were able to complete tasks in less time, and with a higher accuracy rate. Further, the subjects perceived that they operated more effectively when using the uncertainty visualization interface, in contrast to when the subjects used the standard interface.

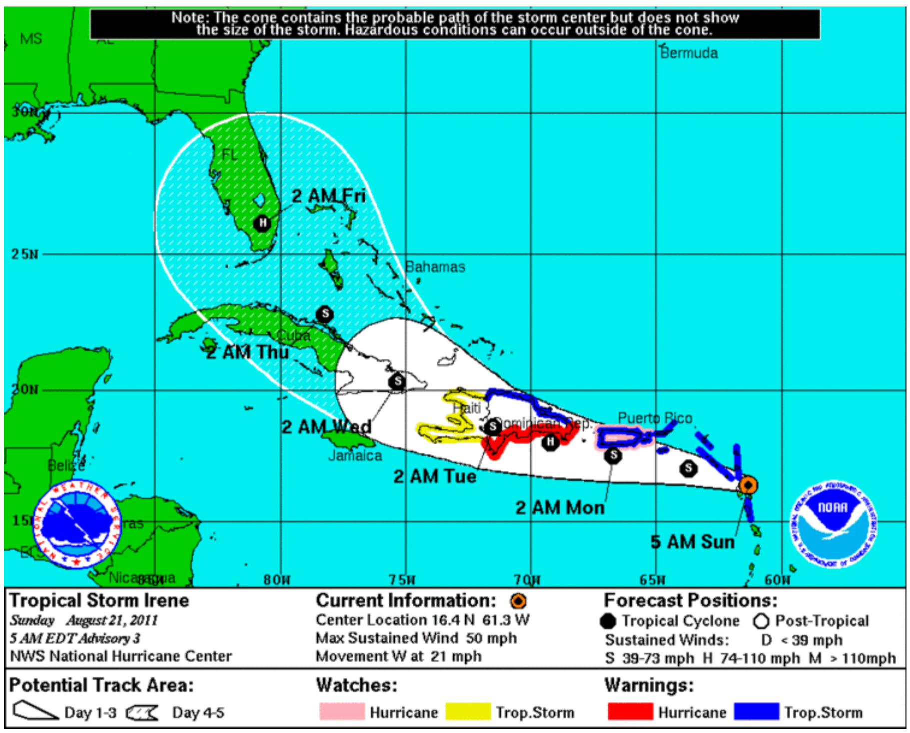
The research supported the hypothesis that adaptive uncertainty visualization significantly reduces a user's cognitive load in an environment where both stress and uncertainty abound. This hypothesis has clear implications for situations where participants are subject to high cognitive load, such as soldiers on the battlefield or on peacekeeping missions, or first-responders in emergencies. There are many scenarios where a human decision maker could benefit from having visibility to uncertainty factors; for example, a soldier on a peacekeeping mission may determine that an erratically behaving sensor may have been tampered with or disabled, indicating the presence of insurgents.

While this study demonstrated conclusively that adaptive uncertainty visualization significantly reduces cognitive load, the implications of uncertainty visualization on cognitive load merit further research and investigation. For example, humans are highly successful in completing complex tasks, such as driving, in a multimodal setting that includes auditory, visual, and haptic stimuli (Sigrist, Rauter, Riener, & Wolf, 2012). Presenting uncertainty factors across multiple modalities may provide different effects to a user's cognitive load than a single modality. In addition, researching uncertainty visualization in the context of adaptable systems may offer insights on how a system operator can offset the cognitive burden of coping with uncertainty by minimizing stimuli from other, less essential sources. These insights can inform and refine the design of adaptive systems, such as the emerging field of augmented cognition.

Appendix A. Participant Instructions

Background

The “probability cone” is a familiar sight to Florida residents. Because weather forecasters cannot predict the path of a hurricane with a high degree of certainty, a “cone” is used to portray the hurricane’s probable path.



Projecting the path of a hurricane involves a great deal of uncertainty, since wind, current and temperature patterns can alter a hurricane’s course in unpredictable ways. The probability cone conveys to the reader the hurricane’s likely path, and allows the reader to make an informed decision about whether they should prepare, or plan to evacuate.

To a person unfamiliar with the meaning of the cone, however, it’s easy to misinterpret the size of the cone as indicating the expected size of the hurricane! This illustrates how a graphic that was intended to convey one piece of information – the hurricane’s probably path – may inadvertently misinform the reader by suggesting it portrays the hurricane’s size.

Purpose

The purpose of this research is to determine the effectiveness of a graphic, such as the “probability cone” in a high-stress situation. For example, imagine you’re the captain of a fire team working to extinguish a blazing building. You have to make many decisions in a short period of time with imperfect information, and a wrong decision can endanger lives. If you knew how far the fire had spread inside the building, you could make a better decision on whether your crew could safely enter the building. Using your tablet to explore a 3-D image of the building, you map out the likely path of the blaze based on where the fire originated, the layout of the building’s ventilation system, etc. Would a visual “cone of probability” help you make better decisions about where the fire was headed, or simply add to your stress and anxiety, possibly resulting in poorer decisions?

What to Expect

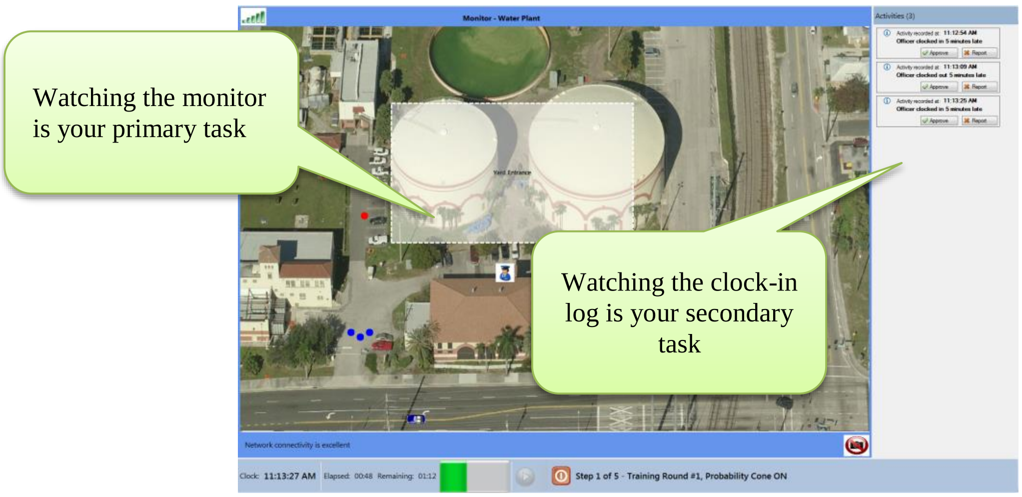
First, I’d like to thank you for agreeing to participate in this research! The sample will include twenty-three other “subjects”. The process will be identical for all twenty-four subjects, with only minor variations. I will ask you to read and sign some paperwork (there’s *always* paperwork!); then we’ll review the “simulation”.

The simulation involves a fictional security company that patrols different restricted sites. You will play the role of a “dispatcher” in two timed “rounds of ten minutes each. The dispatcher is not on-site, but works in the call center in the company’s headquarters. The dispatcher seems an overhead “view”, or “camera” of the site, and can watch the activities on the site in order to provide instructions to the security officer who is on the premises. There are sporadic network outages that prevent the camera from transmitting images to the dispatcher. In one round, you will see the equivalent of the “probability cone” which predicts where people have moved during the network outage. In another round, you’ll see a static “freeze frame” because the simulation makes no effort to predict where people have moved while the system is offline.

During these rounds, your responses will be timed, and the accuracy of your responses will be measured. These measurements will be used to analyze the effectiveness of the “probability cone” in improving your decision-making. Finally, you’ll be asked to complete a quick survey (the last piece of paperwork – I promise!).

Dispatcher Instruction Sheet

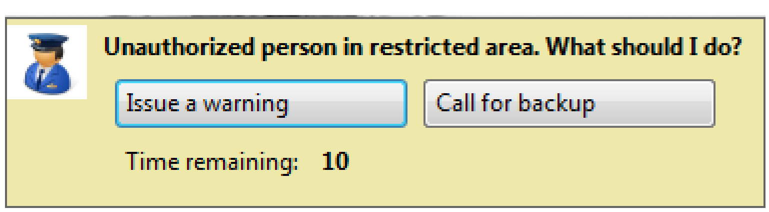
As a dispatcher, you will have two tasks. Your primary task is to watch the monitor on the left side of the screen.



Your secondary task is to monitor the clock-in log. However, this is only as time permits. For example, as you scan the monitor, you may see that nothing is happening that requires your attention; then, you may then scan the clock-in log for activities to approve or report.

<u>**The Monitor**</u>

A security officer is posted on-site to guard a restricted area. If an intruder enters the restricted area, the security officer will prompt you, the dispatcher, for instructions on how to best respond. You have ten seconds to respond.

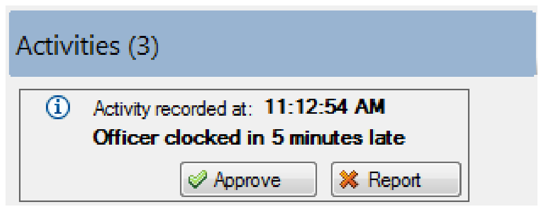


The rule for responding to the security officer's prompt is as follows:

A person entering a restricted area must be given two warnings; after two warnings, you must call for backup

The Clock-In Log

The clock-in log records when security officers clock in, or out, of their post.



When the officer clocks in on time, you should approve the activity; however, when the officer clocks out early, you must report the activity. The rule for approving or reporting activity is as follows:

- Clocking **in** early or on-time
- Clocking **out** on-time or later

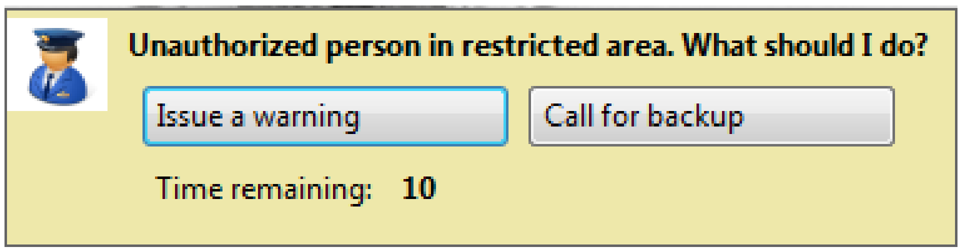
- Clocking **in** late
- Clocking **out** early

Dispatcher Instructions

All dispatchers must follow the instructions below.

Security Monitor

*A person entering a restricted area must be given **__two__** warnings; after two warnings, you must **__call for backup__***



Clock-in/Clock-out log

1. 

Clocking in and clocking out **on time**

__Clocking in__

2. 

Clocking in **early**

Clocking in **late**

__Clocking out__

3. 

Clocking out **late**

Clocking out **early**

Appendix B. User Survey

ID: _______________

In the table below, please circle the number (1-5) that most accurately describes your experience while using the system. A response of “1” means you strongly disagree; a response of “5” means you strongly agree.

Questions	Probability Cone On ●	Probability Cone Off ○
	Disagree Agree	Disagree Agree
1. I knew when a situation required my attention	1 2 3 4 5	1 2 3 4 5
2. I could easily distinguish between critical and non-critical situations	1 2 3 4 5	1 2 3 4 5
3. I could concentrate on a critical decision when it required my attention	1 2 3 4 5	1 2 3 4 5
4. I could prioritize critical and non-critical tasks	1 2 3 4 5	1 2 3 4 5
5. I knew the status of all individuals at all times	1 2 3 4 5	1 2 3 4 5
6. It was easy to switch back and forth between tasks	1 2 3 4 5	1 2 3 4 5
7. I was confident my decisions would not result in errors	1 2 3 4 5	1 2 3 4 5
8. I was able to make good decisions even when network errors prevented me from seeing timely information	1 2 3 4 5	1 2 3 4 5
9. I was able to make good decisions even when a lot of things were happening	1 2 3 4 5	1 2 3 4 5
10. I was able to make good decisions even when critical things were happening	1 2 3 4 5	1 2 3 4 5
11. Only essential information was displayed to me	1 2 3 4 5	1 2 3 4 5
12. I was frustrated when network disruptions prevented me from understanding what was going on	1 2 3 4 5	1 2 3 4 5
13. I was overwhelmed with the information displayed	1 2 3 4 5	1 2 3 4 5
14. I was distracted by the information displayed	1 2 3 4 5	1 2 3 4 5
15. I was uncertain what actions to take	1 2 3 4 5	1 2 3 4 5
16. The amount of information displayed prevented me from making good decisions	1 2 3 4 5	1 2 3 4 5
17. I made more errors when network errors prevented me from seeing timely information	1 2 3 4 5	1 2 3 4 5
18. It was easy to track individuals and remember who they were	1 2 3 4 5	1 2 3 4 5
19. The level of movement made it easy for me to distinguish individuals from the background	1 2 3 4 5	1 2 3 4 5
20. The level of movement made it easy for me to be distracted by less critical tasks	1 2 3 4 5	1 2 3 4 5

Appendix C. IRB Approval Memorandum

NOVA SOUTHEASTERN UNIVERSITY
Office of Grants and Contracts
Institutional Review Board

NSU

MEMORANDUM

To: Gregory Block

From: Ling Wang, Ph.D.
Institutional Review Board

Date: July 16, 2012

Re: *Reducing Cognitive Load Using Adaptive Uncertainty Visualization*

IRB Approval Number: wang07151201

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms these must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE REACTIONS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and 954-262-2020 respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Protocol File

Appendix D. Simulation Results

Simulation Results - Primary Task

		Task A		Task B		Task C		Task D		Task E		Task F		Task G	
Subject	Interface	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration
1	Standard	3	8	2	6	2	7	2	8	2	7	3	8	2	5
1	Uncertainty	3	4	3	7	3	9	3	6	3	4	2	4	3	7
2	Standard	3	6	3	9	3	8	2	8	2	7	2	7	2	6
2	Uncertainty	3	7	3	7	3	5	3	6	3	5	3	5	3	8
3	Standard	3	8	3	8	3	5	3	8	2	7	3	8	2	8
3	Uncertainty	3	6	3	5	3	4	3	4	3	8	3	5	3	7
4	Standard	3	4	2	6	2	6	2	8	2	6	3	6	2	6
4	Uncertainty	3	4	3	5	3	5	3	6	3	6	3	6	3	4
5	Standard	3	6	2	6	2	3	3	8	3	8	3	7	2	7
5	Uncertainty	3	4	3	5	2	5	3	5	3	6	2	4	3	6
6	Standard	3	5	3	5	3	5	3	4	3	8	3	7	3	6
6	Uncertainty	3	5	2	5	3	5	3	5	3	6	2	5	3	6
7	Standard	3	7	1	5	3	8	3	8	3	8	2	4	3	10
7	Uncertainty	3	3	2	8	2	4	3	4	3	4	2	7	2	7
8	Standard	3	7	2	6	3	5	3	7	3	7	3	6	3	8
8	Uncertainty	3	4	2	6	2	6	3	5	3	4	3	6	3	6
9	Standard	3	6	3	4	3	3	2	6	2	8	2	3	2	8
9	Uncertainty	3	5	3	4	2	4	3	7	2	4	3	5	3	5
10	Standard	2	7	1	9	3	7	3	8	2	8	2	7	3	5
10	Uncertainty	3	4	3	6	3	7	3	6	3	7	3	6	3	9
11	Standard	2	6	2	8	2	7	2	8	2	8	2	9	2	7
11	Uncertainty	3	7	2	5	2	6	3	5	3	5	3	6	2	8
12	Standard	3	7	3	9	3	6	3	7	3	7	2	7	3	5
12	Uncertainty	3	4	3	4	3	6	3	4	3	6	2	5	3	8
13	Standard	3	7	2	9	3	6	2	8	3	9	2	8	2	8
13	Uncertainty	3	4	2	5	3	5	3	7	3	4	3	6	3	4
14	Standard	3	6	3	4	3	8	3	8	1	6	3	8	3	7
14	Uncertainty	3	4	3	3	3	5	3	6	3	6	3	6	3	7
15	Standard	3	7	3	6	3	5	3	8	3	8	2	7	2	7
15	Uncertainty	3	4	3	5	3	5	3	5	3	5	3	7	2	5
16	Standard	3	9	3	19	3	8	3	7	2	8	2	9	3	8
16	Uncertainty	3	5	3	5	1	5	3	9	2	6	3	5	3	6
17	Standard	3	7	2	9	3	9	3	8	3	8	2	7	2	6
17	Uncertainty	3	5	3	7	3	5	3	5	3	7	3	7	3	7
18	Standard	3	8	2	6	2	7	2	6	3	8	2	7	2	7
18	Uncertainty	3	4	3	6	3	5	3	6	3	8	3	6	2	10
19	Standard	3	6	2	9	3	6	3	8	3	7	3	8	3	6
19	Uncertainty	2	4	3	5	3	5	3	6	3	7	3	5	3	8
20	Standard	3	8	3	11	3	7	2	10	3	7	2	8	2	14
20	Uncertainty	3	8	3	10	3	7	3	8	3	15	3	6	3	15
21	Standard	3	7	2	4	3	5	2	7	3	7	2	8	2	5
21	Uncertainty	3	5	3	3	3	5	2	5	3	5	3	6	3	3
22	Standard	3	4	3	6	3	5	3	9	3	6	3	5	1	11
22	Uncertainty	3	6	3	5	3	8	3	5	3	4	3	7	3	5

23	Standard	2	6	3	8	3	6	3	6	3	6	3	7	3	6
23	Uncertainty	3	5	2	5	3	10	3	9	3	7	3	7	3	7
24	Standard	3	5	3	10	3	4	2	6	2	8	1	9	3	2
24	Uncertainty	3	4	3	5	3	10	3	9	3	6	3	5	3	4

Simulation Results - Secondary Task

		Task 1		Task 2		Task 3	
Subject	Interface	Accuracy	Duration	Accuracy	Duration	Accuracy	Duration
1	Standard	1.00	19	0.83	24	0.89	22
1	Uncertainty	1.00	14	0.83	14	0.89	14
2	Standard	1.00	3	0.83	3	1.00	3
2	Uncertainty	1.00	2	0.92	5	1.00	4
3	Standard	1.00	11	0.92	8	0.94	11
3	Uncertainty	1.00	16	0.75	12	0.94	13
4	Standard	1.00	2	0.83	4	1.00	2
4	Uncertainty	1.00	2	0.92	3	0.89	4
5	Standard	1.00	3	0.92	4	1.00	4
5	Uncertainty	1.00	3	0.92	6	1.00	5
6	Standard	0.24	116	0.12	89	0.19	96
6	Uncertainty	0.35	50	0.12	33	0.27	47
7	Standard	1.00	5	1.00	5	0.85	7
7	Uncertainty	1.00	3	0.94	6	0.88	6
8	Standard	1.00	34	0.29	39	0.92	37
8	Uncertainty	0.94	34	0.00	30	0.96	34
9	Standard	1.00	4	1.00	4	0.85	4
9	Uncertainty	1.00	13	0.94	8	0.96	10
10	Standard	1.00	8	1.00	9	0.96	8
10	Uncertainty	1.00	5	1.00	5	0.96	6
11	Standard	1.00	6	0.94	6	0.96	7
11	Uncertainty	1.00	6	0.94	7	0.96	6
12	Standard	0.94	8	0.94	8	0.92	9
12	Uncertainty	1.00	13	1.00	22	0.92	20
13	Standard	1.00	9	1.00	11	0.96	15
13	Uncertainty	0.94	11	1.00	12	0.92	11
14	Standard	1.00	3	1.00	4	0.96	4
14	Uncertainty	1.00	2	1.00	3	0.96	3
15	Standard	1.00	3	1.00	5	0.96	5
15	Uncertainty	1.00	3	1.00	4	0.96	4
16	Standard	0.59	42	0.59	49	0.50	44
16	Uncertainty	0.94	24	0.53	55	0.38	35
17	Standard	1.00	4	0.00	3	0.96	4
17	Uncertainty	1.00	3	0.00	2	0.96	3
18	Standard	1.00	4	1.00	6	0.96	7
18	Uncertainty	1.00	5	1.00	6	0.96	9
19	Standard	1.00	37	0.94	17	0.92	43
19	Uncertainty	1.00	29	1.00	26	0.92	24
20	Standard	0.94	37	0.94	34	0.88	33
20	Uncertainty	1.00	28	1.00	31	0.96	26
21	Standard	1.00	14	0.82	12	0.81	13
21	Uncertainty	1.00	15	1.00	9	0.92	13

22	Standard	0.94	11	0.82	13	0.77	13
22	Uncertainty	1.00	10	1.00	12	0.96	9
23	Standard	1.00	9	1.00	8	0.96	10
23	Uncertainty	0.94	6	1.00	6	0.92	7
24	Standard	0.76	88	0.41	74	0.38	70
24	Uncertainty	0.71	18	0.65	66	0.62	61

Appendix E. Survey Results and Participant Comments

Survey Results – Standard Interface

Subject	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Average
1	2	2	2	1	1	2	2	3	3	2	5	2	3	2	1	1	1	1	1	5	2.1
2	5	4	5	5	3	4	3	4	4	5	5	4	5	5	5	4	4	3	2	1	4.0
3	3	3	4	3	2	3	3	3	2	2	5	2	5	2	2	2	2	3	2	4	2.9
4	4	3	3	4	2	3	3	3	3	3	5	3	2	2	2	5	2	4	3	2	3.1
5	2	4	4	4	2	3	3	4	2	2	3	4	3	4	5	4	1	3	3	2	3.1
6	3	2	3	3	4	3	3	5	4	2	4	3	5	2	5	5	2	3	3	3	3.4
7	4	5	4	4	5	5	5	5	5	4	5	1	5	3	5	5	4	5	4	5	4.4
8	5	5	5	4	2	1	3	2	2	3	1	1	1	1	3	1	1	1	2	3	2.4
9	2	3	5	5	5	5	4	5	5	5	5	5	5	5	2	1	1	3	4	5	4.0
10	5	4	3	4	2	1	3	3	2	3	4	3	5	1	1	3	1	5	3	2	2.9
11	3	2	5	2	2	5	2	2	3	3	2	4	5	1	1	4	1	2	3	4	2.8
12	2	3	5	1	4	2	4	4	4	4	5	1	2	4	1	1	4	1	1	2	2.8
13	3	2	2	3	2	4	3	4	2	1	4	2	2	5	2	2	1	2	4	3	2.7
14	4	4	4	5	4	4	5	5	5	4	5	5	5	3	5	5	2	5	4	4	4.4
15	5	4	4	3	4	5	4	4	4	4	1	3	3	2	5	3	5	4	5	3	3.8
16	3	3	2	4	2	2	1	2	1	1	1	1	2	2	3	1	1	1	3	3	2.0
17	4	2	5	2	2	1	4	5	5	4	1	2	2	1	3	2	2	1	4	2	2.7
18	3	2	2	2	3	2	2	4	4	3	4	5	2	2	5	5	1	2	1	1	2.8
19	4	3	4	3	4	4	5	5	5	5	5	5	5	5	5	5	5	5	4	5	4.6
20	5	3	3	3	1	2	2	2	2	2	5	1	1	2	1	1	1	2	1	4	2.2
21	4	4	3	4	3	4	1	4	3	3	2	4	5	3	3	4	2	2	5	4	3.4
22	5	5	5	5	3	5	4	5	5	5	1	5	4	5	5	5	2	5	5	1	4.3
23	4	4	3	2	3	4	4	4	4	4	3	2	3	2	2	5	2	4	4	4	3.4
24	2	2	3	1	1	3	1	1	3	2	2	3	4	2	2	1	1	3	2	3	2.1
Average	**3.6**	**3.3**	**3.7**	**3.2**	**2.8**	**3.2**	**3.1**	**3.7**	**3.4**	**3.2**	**3.5**	**3.0**	**3.5**	**2.8**	**3.1**	**3.1**	**2.0**	**2.9**	**3.0**	**3.1**	**3.2**

Survey Results – Uncertainty Visualization Interface

Subject	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Average
1	5	5	5	5	5	4	5	5	5	5	5	5	4	4	5	5	5	4	5	1	**4.6**
2	5	4	5	5	5	1	3	4	5	5	5	5	5	5	5	4	4	5	4	1	**4.3**
3	4	5	5	5	4	5	4	4	4	4	5	3	5	5	5	5	4	3	4	2	**4.3**
4	5	5	5	5	4	5	4	4	4	3	5	4	3	3	4	5	3	4	5	4	**4.2**
5	5	4	4	4	4	5	5	5	4	4	3	5	5	4	5	5	5	5	4	5	**4.5**
6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	**5.0**
7	5	5	5	5	5	5	5	5	5	5	5	1	5	5	5	5	4	5	5	1	**4.6**
8	5	5	5	5	2	5	3	2	5	5	1	1	2	1	3	1	5	5	5	4	**3.5**
9	5	5	5	5	5	5	4	5	5	5	5	5	5	5	2	1	2	5	5	5	**4.5**
10	5	5	3	5	2	1	5	5	2	3	4	1	5	1	3	3	5	5	5	5	**3.7**
11	5	4	5	3	2	5	3	2	4	4	3	4	5	2	2	4	2	2	5	1	**3.4**
12	5	5	5	5	5	5	4	5	3	5	5	4	2	1	5	5	2	5	5	4	**4.3**
13	5	5	2	5	5	4	5	5	5	5	4	4	3	5	5	4	3	3	4	3	**4.2**
14	5	5	5	5	5	5	5	5	5	4	5	5	5	3	5	5	4	5	5	2	**4.7**
15	5	5	4	3	4	5	4	4	4	4	1	3	3	2	5	3	5	4	5	3	**3.8**
16	4	4	3	4	2	3	1	3	2	1	1	1	2	2	3	1	2	1	4	5	**2.5**
17	5	2	5	2	4	1	4	5	5	1	2	3	3	1	4	3	3	1	5	2	**3.1**
18	5	5	5	5	3	3	2	4	5	5	4	5	3	3	5	5	3	5	5	5	**4.3**
19	5	3	5	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	3	**4.7**
20	5	4	4	3	2	4	4	4	5	4	5	2	2	5	2	1	2	4	5	4	**3.6**
21	4	5	3	5	3	4	3	4	4	4	2	4	5	3	4	4	3	3	5	5	**3.9**
22	5	5	5	5	4	5	5	5	5	5	1	5	4	5	5	5	3	5	5	1	**4.4**
23	5	5	3	4	5	5	5	5	5	5	3	3	3	5	4	5	4	4	5	2	**4.3**
24	5	5	5	5	4	3	4	5	4	4	5	3	4	4	4	5	5	5	5	3	**4.4**
Average	**4.9**	**4.6**	**4.4**	**4.5**	**3.9**	**4.1**	**4.0**	**4.4**	**4.4**	**4.2**	**3.7**	**3.6**	**3.9**	**3.5**	**4.2**	**3.9**	**3.6**	**4.1**	**4.8**	**3.2**	**4.1**

Respondent	Interface	Remarks
1	Standard	During the network outage it was easy to be distracted because there was no movement
2	Both	If you have a video game background, the system was pretty easy to operate
5	Standard	Required me to pay more attention because I could not tell what was going on; couldn't look away to take care of other things
5	Uncertainty	I was able to relax rather than stare at the screen
6	Both	Too many things were happening; it was hard to keep up with everything
7	Both	It's like a video game only slower
8	Uncertainty	I felt safer when I saw the cones because I had a better idea what was going on

Respondent	Interface	Remarks
9	Standard	I had to work harder because I had no idea where people were moving and it was harder to find them
11	Both	It was hard to take the simulation seriously because the task was so different from what I usually do; photographs or video that was on the ground would be more realistic
12	Standard	I could only do one thing at a time when there was a network disruption
12	Uncertainty	I knew how much time I had to attend to other tasks, so it was easier to look away from the monitor during a network disruption
13	Standard	I got a false sense of security when nothing moved on the screen
13	Uncertainty	I definitely had more time to respond
14	Both	Graphic component of the simulation wasn’t necessary to making a decision, but it did help in deciding how much time I had
16	Both	Things moved too fast and there were too many things happening at the same time
17	Both	I was required to handle tasks, like counting/memorization or evaluating simple rules, that the computer could easily do; the system should have recognized my stress level and disabled any clock-in prompts when the network was down
19	Uncertainty	It would have been stressful if the probability cone were really large, or the predictions about direction were incorrect; I didn’t really need the graphics to make decisions
22	Both	The simulation wasn’t realistic – they were just “dots”; that made it hard to treat tasks as if they were critical
23	Both	The clock-out tasks were annoying and made it really hard to concentrate
24	Standard	I have a hard time focusing and paying attention; not seeing any movement forced me to focus, which isn’t a bad thing; however, I couldn’t pay any attention to the other tasks then

Appendix F. Minitab Output

12/20/2012 1:26:40 PM

```
Welcome to Minitab, press F1 for help.
Retrieving project from file:
'C:\USERS\GREG\DOCUMENTS\COLLEGES\NSU\DISSERTATION\WIP\4.
RESEARCH\MINITAB\DISSERTATION.MPJ'
```

Results for: Task 1 Results.MTW

ANOVA: Task A Correct, Task B Correct, ... versus Interface, Participant

```
Factor       Type    Levels
Interface    fixed        2
Participant  random      24

Factor       Values
Interface    0, 1
Participant   1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19, 20, 21, 22, 23, 24


Analysis of Variance for Task A Correct

Source       DF       SS       MS     F      P
Interface     1  0.08333  0.08333  1.00  0.328
Participant  23  1.66667  0.07246  0.87  0.630
Error        23  1.91667  0.08333
Total        47  3.66667


S = 0.288675   R-Sq = 47.73%   R-Sq(adj) = 0.00%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant   -0.00543      3  (3) + 2 (2)
3  Error          0.08333         (3)


Analysis of Variance for Task B Correct

Source       DF       SS      MS     F      P
Interface     1   1.3333  1.3333  5.41  0.029
Participant  23   8.6667  0.3768  1.53  0.158
Error        23   5.6667  0.2464
Total        47  15.6667
```

S = 0.496364   R-Sq = 63.83%   R-Sq(adj) = 26.09%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.06522	3	(3) + 2 (2)
3	Error	0.24638		(3)

Analysis of Variance for Task C Correct

Source	DF	SS	MS	F	P
Interface	1	0.0833	0.0833	0.39	0.539
Participant	23	6.0000	0.2609	1.22	0.318
Error	23	4.9167	0.2138		
Total	47	11.0000			

S = 0.462351   R-Sq = 55.30%   R-Sq(adj) = 8.66%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.02355	3	(3) + 2 (2)
3	Error	0.21377		(3)

Analysis of Variance for Task D Correct

Source	DF	SS	MS	F	P
Interface	1	1.6875	1.6875	13.80	0.001
Participant	23	3.9792	0.1730	1.41	0.206
Error	23	2.8125	0.1223		
Total	47	8.4792			

S = 0.349689   R-Sq = 66.83%   R-Sq(adj) = 32.22%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.02536	3	(3) + 2 (2)
3	Error	0.12228		(3)

Analysis of Variance for Task E Correct

Source	DF	SS	MS	F	P
Interface	1	1.6875	1.6875	10.18	0.004
Participant	23	5.9792	0.2600	1.57	0.144
Error	23	3.8125	0.1658		

```
Total        47  11.4792


S = 0.407137   R-Sq = 66.79%   R-Sq(adj) = 32.13%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant    0.04710      3  (3) + 2 (2)
3  Error          0.16576         (3)


Analysis of Variance for Task F Correct

Source       DF       SS      MS     F      P
Interface     1   2.0833  2.0833  6.93  0.015
Participant  23   4.6667  0.2029  0.67  0.824
Error        23   6.9167  0.3007
Total        47  13.6667


S = 0.548384   R-Sq = 49.39%   R-Sq(adj) = 0.00%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant   -0.04891      3  (3) + 2 (2)
3  Error          0.30072         (3)


Analysis of Variance for Task G Correct

Source       DF       SS      MS      F      P
Interface     1   2.5208  2.5208  11.64  0.002
Participant  23   5.9792  0.2600   1.20  0.332
Error        23   4.9792  0.2165
Total        47  13.4792


S = 0.465280   R-Sq = 63.06%   R-Sq(adj) = 24.51%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant    0.02174      3  (3) + 2 (2)
3  Error          0.21649         (3)
```

ANOVA: Task A Durat, Task B Durat, ... versus Interface, Participant

```
Factor       Type    Levels
```

```
Interface    fixed        2
Participant  random      24

Factor       Values
Interface    0, 1
Participant   1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19, 20, 21, 22, 23, 24


Analysis of Variance for Task A Duration

Source       DF       SS      MS      F      P
Interface     1   36.750  36.750  24.68  0.000
Participant  23   35.667   1.551   1.04  0.462
Error        23   34.250   1.489
Total        47  106.667


S = 1.22030   R-Sq = 67.89%   R-Sq(adj) = 34.39%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant    0.03080      3  (3) + 2 (2)
3  Error          1.48913         (3)


Analysis of Variance for Task B Duration

Source       DF       SS      MS      F      P
Interface     1   54.187  54.187  10.81  0.003
Participant  23  168.479   7.325   1.46  0.185
Error        23  115.312   5.014
Total        47  337.979


S = 2.23910   R-Sq = 65.88%   R-Sq(adj) = 30.28%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant      1.156      3  (3) + 2 (2)
3  Error            5.014         (3)


Analysis of Variance for Task C Duration

Source       DF       SS     MS     F      P
Interface     1    0.521  0.521  0.17  0.681
Participant  23   59.479  2.586  0.86  0.637
Error        23   68.979  2.999
Total        47  128.979


S = 1.73179   R-Sq = 46.52%   R-Sq(adj) = 0.00%
```

Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1 Interface		3	(3) + 24 Q[1]
2 Participant	-0.2065	3	(3) + 2 (2)
3 Error	2.9991		(3)

Analysis of Variance for Task D Duration

Source	DF	SS	MS	F	P
Interface	1	27.000	27.000	12.42	0.002
Participant	23	36.917	1.605	0.74	0.764
Error	23	50.000	2.174		
Total	47	113.917			

S = 1.47442   R-Sq = 56.11%   R-Sq(adj) = 10.31%

Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1 Interface		3	(3) + 24 Q[1]
2 Participant	-0.2844	3	(3) + 2 (2)
3 Error	2.1739		(3)

Analysis of Variance for Task E Duration

Source	DF	SS	MS	F	P
Interface	1	21.333	21.333	6.66	0.017
Participant	23	62.917	2.736	0.85	0.646
Error	23	73.667	3.203		
Total	47	157.917			

S = 1.78966   R-Sq = 53.35%   R-Sq(adj) = 4.67%

Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1 Interface		3	(3) + 24 Q[1]
2 Participant	-0.2337	3	(3) + 2 (2)
3 Error	3.2029		(3)

Analysis of Variance for Task F Duration

Source	DF	SS	MS	F	P
Interface	1	22.687	22.687	12.19	0.002
Participant	23	25.979	1.130	0.61	0.881
Error	23	42.813	1.861		
Total	47	91.479			

```
S = 1.36434   R-Sq = 53.20%   R-Sq(adj) = 4.36%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                  Variance  Error  restricted
   Source        component   term  model)
1  Interface                    3  (3) + 24 Q[1]
2  Participant    -0.3659       3  (3) + 2 (2)
3  Error           1.8614          (3)


Analysis of Variance for Task G Duration

Source       DF       SS     MS     F      P
Interface     1    0.750  0.750  0.24  0.630
Participant  23  192.250  8.359  2.66  0.011
Error        23   72.250  3.141
Total        47  265.250


S = 1.77237   R-Sq = 72.76%   R-Sq(adj) = 44.34%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                  Variance  Error  restricted
   Source        component   term  model)
1  Interface                    3  (3) + 24 Q[1]
2  Participant      2.609       3  (3) + 2 (2)
3  Error            3.141          (3)
```

Results for: Task 2 Results.MTW

ANOVA: Task 1 Correct, Task 2 Correct, ... versus Interface, Participant

```
Factor       Type    Levels
Interface    fixed        2
Participant  random      24

Factor       Values
Interface    0, 1
Participant   1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19, 20, 21, 22, 23, 24


Analysis of Variance for Task 1 Correct

Source       DF        SS        MS      F      P
Interface     1  0.003532  0.003532   1.04  0.317
Participant  23  1.098545  0.047763  14.12  0.000
Error        23  0.077783  0.003382
Total        47  1.179860


S = 0.0581537   R-Sq = 93.41%   R-Sq(adj) = 86.53%
```

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.02219	3	(3) + 2 (2)
3	Error	0.00338		(3)

Analysis of Variance for Task 2 Correct

Source	DF	SS	MS	F	P
Interface	1	0.00180	0.00180	0.32	0.580
Participant	23	4.24074	0.18438	32.25	0.000
Error	23	0.13149	0.00572		
Total	47	4.37403			

S = 0.0756099   R-Sq = 96.99%   R-Sq(adj) = 93.86%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.08933	3	(3) + 2 (2)
3	Error	0.00572		(3)

Analysis of Variance for Task 3 Correct

Source	DF	SS	MS	F	P
Interface	1	0.007037	0.007037	2.21	0.150
Participant	23	1.698260	0.073837	23.23	0.000
Error	23	0.073100	0.003178		
Total	47	1.778398			

S = 0.0563762   R-Sq = 95.89%   R-Sq(adj) = 91.60%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		3	(3) + 24 Q[1]
2	Participant	0.03533	3	(3) + 2 (2)
3	Error	0.00318		(3)

ANOVA: Task 1 Durat, Task 2 Durat, ... versus Interface, Participant

Factor	Type	Levels
Interface	fixed	2
Participant	random	24

Factor	Values
Interface	0, 1

```
Participant   1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19, 20, 21, 22, 23, 24


Analysis of Variance for Task 1 Duration

Source       DF       SS     MS     F      P
Interface     1    562.1  562.1  2.95  0.100
Participant  23  17622.1  766.2  4.02  0.001
Error        23   4388.6  190.8
Total        47  22572.8


S = 13.8133   R-Sq = 80.56%   R-Sq(adj) = 60.27%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant      287.7      3  (3) + 2 (2)
3  Error            190.8         (3)


Analysis of Variance for Task 2 Duration

Source       DF        SS      MS     F      P
Interface     1     65.00   65.00  0.84  0.370
Participant  23  16768.69  729.07  9.38  0.000
Error        23   1788.45   77.76
Total        47  18622.15


S = 8.81809   R-Sq = 90.40%   R-Sq(adj) = 80.37%


                                  Expected Mean
                                  Square for Each
                                  Term (using
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant     325.66      3  (3) + 2 (2)
3  Error            77.76         (3)


Analysis of Variance for Task 3 Duration

Source       DF        SS      MS      F      P
Interface     1    194.52  194.52   3.15  0.089
Participant  23  16635.81  723.30  11.73  0.000
Error        23   1418.42   61.67
Total        47  18248.75


S = 7.85306   R-Sq = 92.23%   R-Sq(adj) = 84.12%


                                  Expected Mean
                                  Square for Each
                                  Term (using
```

```
                 Variance  Error  restricted
   Source       component   term  model)
1  Interface                   3  (3) + 24 Q[1]
2  Participant     330.81      3  (3) + 2 (2)
3  Error            61.67         (3)
```

Results for: Survey Results.MTW

ANOVA: Q1, Q2, ... versus Interface

```
Factor     Type   Levels  Values
Interface  fixed       2  0, 1


Analysis of Variance for Q1

Source     DF      SS      MS      F      P
Interface   1  20.021  20.021  30.24  0.000
Error      46  30.458   0.662
Total      47  50.479


S = 0.813718   R-Sq = 39.66%   R-Sq(adj) = 38.35%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error         0.6621         (2)


Analysis of Variance for Q2

Source     DF      SS      MS      F      P
Interface   1  21.333  21.333  25.60  0.000
Error      46  38.333   0.833
Total      47  59.667


S = 0.912871   R-Sq = 35.75%   R-Sq(adj) = 34.36%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error         0.8333         (2)


Analysis of Variance for Q3

Source     DF      SS     MS     F      P
Interface   1   6.750  6.750  6.58  0.014
Error      46  47.167  1.025
Total      47  53.917
```

S = 1.01260   R-Sq = 12.52%   R-Sq(adj) = 10.62%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		2	(2) + 24 Q[1]
2	Error	1.025		(2)

Analysis of Variance for Q4

Source	DF	SS	MS	F	P
Interface	1	18.750	18.750	15.42	0.000
Error	46	55.917	1.216		
Total	47	74.667			

S = 1.10253   R-Sq = 25.11%   R-Sq(adj) = 23.48%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		2	(2) + 24 Q[1]
2	Error	1.216		(2)

Analysis of Variance for Q5

Source	DF	SS	MS	F	P
Interface	1	15.188	15.188	11.07	0.002
Error	46	63.125	1.372		
Total	47	78.313			

S = 1.17144   R-Sq = 19.39%   R-Sq(adj) = 17.64%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		2	(2) + 24 Q[1]
2	Error	1.372		(2)

Analysis of Variance for Q6

Source	DF	SS	MS	F	P
Interface	1	9.187	9.187	4.93	0.031
Error	46	85.792	1.865		
Total	47	94.979			

S = 1.36566   R-Sq = 9.67%   R-Sq(adj) = 7.71%

```
                                Expected Mean
                                Square for Each
                                Term (using
                Variance  Error  restricted
   Source      component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error           1.865         (2)
```

Analysis of Variance for Q7

```
Source     DF      SS      MS     F      P
Interface   1  11.021  11.021  8.34  0.006
Error      46  60.792   1.322
Total      47  71.813
```

S = 1.14959   R-Sq = 15.35%   R-Sq(adj) = 13.51%

```
                                Expected Mean
                                Square for Each
                                Term (using
                Variance  Error  restricted
   Source      component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error           1.322         (2)
```

Analysis of Variance for Q8

```
Source     DF      SS     MS     F      P
Interface   1   6.021  6.021  5.23  0.027
Error      46  52.958  1.151
Total      47  58.979
```

S = 1.07297   R-Sq = 10.21%   R-Sq(adj) = 8.26%

```
                                Expected Mean
                                Square for Each
                                Term (using
                Variance  Error  restricted
   Source      component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error           1.151         (2)
```

Analysis of Variance for Q9

```
Source     DF      SS      MS     F      P
Interface   1  11.021  11.021  9.14  0.004
Error      46  55.458   1.206
Total      47  66.479
```

S = 1.09801   R-Sq = 16.58%   R-Sq(adj) = 14.76%

```
                                Expected Mean
                                Square for Each
```

```
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          1.206         (2)
```

Analysis of Variance for Q10

Source	DF	SS	MS	F	P
Interface	1	12.000	12.000	8.28	0.006
Error	46	66.667	1.449		
Total	47	78.667			

S = 1.20386   R-Sq = 15.25%   R-Sq(adj) = 13.41%

```
                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          1.449         (2)
```

Analysis of Variance for Q11

Source	DF	SS	MS	F	P
Interface	1	0.750	0.750	0.29	0.593
Error	46	118.917	2.585		
Total	47	119.667			

S = 1.60784   R-Sq = 0.63%   R-Sq(adj) = 0.00%

```
                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          2.585         (2)
```

Analysis of Variance for Q12

Source	DF	SS	MS	F	P
Interface	1	4.688	4.688	2.18	0.146
Error	46	98.792	2.148		
Total	47	103.479			

S = 1.46548   R-Sq = 4.53%   R-Sq(adj) = 2.45%

```
                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
```

```
1  Interface                  2  (2) + 24 Q[1]
2  Error          2.148          (2)


Analysis of Variance for Q13

Source     DF      SS     MS     F      P
Interface   1   1.688  1.688  0.94  0.337
Error      46  82.625  1.796
Total      47  84.313


S = 1.34022   R-Sq = 2.00%   R-Sq(adj) = 0.00%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error          1.796          (2)


Analysis of Variance for Q14

Source     DF       SS     MS     F      P
Interface   1    6.750  6.750  3.03  0.088
Error      46  102.500  2.228
Total      47  109.250


S = 1.49274   R-Sq = 6.18%   R-Sq(adj) = 4.14%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error          2.228          (2)


Analysis of Variance for Q15

Source     DF       SS      MS     F      P
Interface   1   14.083  14.083  7.27  0.010
Error      46   89.167   1.938
Total      47  103.250


S = 1.39227   R-Sq = 13.64%   R-Sq(adj) = 11.76%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                  2  (2) + 24 Q[1]
2  Error          1.938          (2)
```

```
Analysis of Variance for Q16

Source     DF       SS     MS     F      P
Interface   1    7.521  7.521  2.92  0.094
Error      46  118.458  2.575
Total      47  125.979


S = 1.60474   R-Sq = 5.97%   R-Sq(adj) = 3.93%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          2.575         (2)


Analysis of Variance for Q17

Source     DF       SS      MS      F      P
Interface   1   30.083  30.083  19.61  0.000
Error      46   70.583   1.534
Total      47  100.667


S = 1.23872   R-Sq = 29.88%   R-Sq(adj) = 28.36%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          1.534         (2)


Analysis of Variance for Q18

Source     DF       SS      MS     F      P
Interface   1   16.333  16.333  8.77  0.005
Error      46   85.667   1.862
Total      47  102.000


S = 1.36467   R-Sq = 16.01%   R-Sq(adj) = 14.19%


                                Expected Mean
                                Square for Each
                                Term (using
               Variance  Error  restricted
   Source     component   term  model)
1  Interface                 2  (2) + 24 Q[1]
2  Error          1.862         (2)


Analysis of Variance for Q19
```

Source	DF	SS	MS	F	P
Interface	1	36.750	36.750	39.39	0.000
Error	46	42.917	0.933		
Total	47	79.667			

S = 0.965904   R-Sq = 46.13%   R-Sq(adj) = 44.96%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		2	(2) + 24 Q[1]
2	Error	0.9330		(2)

Analysis of Variance for Q20

Source	DF	SS	MS	F	P
Interface	1	0.021	0.021	0.01	0.920
Error	46	93.958	2.043		
Total	47	93.979			

S = 1.42919   R-Sq = 0.02%   R-Sq(adj) = 0.00%

	Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1	Interface		2	(2) + 24 Q[1]
2	Error	2.043		(2)

References

Alt, J. K., & Lieberman, S. (2010). Agent Frameworks for Discrete Event Social Simulations. *Proceedings of the 19th Conference on Behavior Representation in Modeling and Simulation,* (pp. 134-139). Charleston, South Carolina.

Andre, A. D., & Cutler, H. A. (1998). Displaying Uncertainty in Advanced Navigation Systems. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 31-35). Santa Monica, US: Human Factors and Ergonomics Society.

Antifakos, S., Schwaninger, A., & Schiele, B. (2004). Evaluating the Effects of Displaying Uncertainty in Context-Aware Applications. *Proceedings of the 6th International Conference on Ubiquitous Computing* (pp. 54-69). Nottingham, UK: Springer-Verlag.

Arabo, A., Shi, Q., & Merabti, M. (2011). ContextRan: Begetting Order to Usage of Context information and Identity Management in Pervasive Ad-hoc Environments. In J. Symonds, *Emerging Pervasive and Ubiquitous Aspects of Information Systems: Cross-Disciplinary Advancements* (pp. 275-297). Aukland, New Zealand: Auckland University of Technology.

Atoyan, H., Robert, J.-M., & Duquet, J.-R. (2011). Uncertainties in complex dynamic environments. *Journal d'Interaction Personne-Système, 2*(1), 1-24.

Axelrod, R. (2003). Advancing the Art of Simulation in the Social Sciences. *Japanese Journal for Management Information Systems, 12*(3).

Back, J., & Oppenheim, C. (2001). A Model of Cognitive Load for Information Retrieval: Implications for User Relevance Feedback Interaction. *Information Research, 6*(2).

Baddeley, A. (2003). Working Memory: Working Backward, Working Forward. *Neuroscience, 4,* 829-839.

Baldauf, M. (2007). A Survey on Context-Aware Systems. *International Journal of Ad Hoc and Ubiquitous Computing, 2*(4), 263-277.

Barrett, L. (2011). *Beyond the Brain: How Body and Environment Shape Animal and Human Minds.* Princeton: Princeton University Press.

Barthelmé, S., & Mamassian, P. (2010). Flexible Mechanisms Underlie the Evaluation of Visual Confidence. *Proceedings of the National Academy of Sciences, 107*(48), 20834-20839.

Beck, D., & Kastner, S. (2009). Top-down and Bottom-up Mechanisms in Biasing Competition in the Human Brain. *Vision Research, 49*(10), 1154-1165.

Bertin, J. (1983). *The Semiology of Graphics: Diagrams, Networks, Maps.* (W. J. Berg, Trans.) University of Wisconsin Press.

Broad, K., Leiserowitz, A., Weinkle, J., & Steketee, M. (2007). Misinterpretations of the "Cone of Uncertainty" in Florida During the 2004 Hurricane Season. *Bulletin of the American Meteorological Society, 88*, 651–667.

Bunch, R. L., & Lloyd, R. E. (2006). The Cognitive Load of Geographic Information. *The Professional Geographer, 58*(2), 209-220.

Carrasco, M. (2011). Visual Attention: The Past 25 Years. *Vision Research, 51*, 1484-1525.

Cavanagh, P. (2011). Visual Cognition. *Vision Research, 51*(13), 1538-1551.

Chen, C. (2005). Top 10 Unsolved Information Visualization Problems. *IEEE Computer Graphics and Applications, 25*(4), 12-16.

Chong, C.-Y., & Kumar, S. P. (2003). Sensor Networks: Evolution, Opportunities, and Challenges. *Proceedings of the IEEE, 91*(8), 1247-1256.

Cohen, M., & Warren, P. H. (1990). Factors that Influence User Confidence in Expert Systems. *Poster Session, Annual Conference on Ergonomics,* (pp. 486-490). Leeds, England.

Csinger, A. (1992). *The Psychology of Visualization.* Vancouver, BC, Canada, Canada: University of British Columbia.

Davis, G. B. (1999). A Research Perspective for Information Systems and Example of Emerging Area of Research. *Information Systems Frontiers, 1*(3), 195-203.

de Jong, T. (2009). Cognitive Load Theory, Educational Research, and Instructional Design: Some Food for Thought. *Instructional Science, 38*(2), 105-134.

Definition of the NHC Track Forecast Cone. (2011, August 3). Retrieved April 21, 2012, from National Weather Service: http://www.nhc.noaa.gov/aboutcone.shtml

Eckstein, M. P. (2011). Visual Search: A Retrospective. *Journal of Vision, 11*(5), 1-36.

Eick, S., & Fyock, D. (1996). Visualizing Corporate Data. *AT&T Technical Journal, January/February*, 74-99.

Emond, B., & West, R. L. (2004). Using Cognitive Modelling Simulations for User Interface Design Decisions. *Lecture Notes in Computer Science, 3029/2004*, 305-314.

Estrin, D., Govindan, R., Heidemann, J., & Kumar, S. (1999). Next Century Challenges: Scalable Coordination in Sensor Networks. *Proceedings of the International Conference on Mobile Computing and Networking (MOBICOM)*, (pp. 263-270).

Faulkner, L. (2003). Beyond the Five-User Assumption: Benefits of Increased Sample Sizes in Usability Testing. *Behavior Research Methods, Instruments, & Computers, 35*(3), 379-383.

Finger, R., & Bisantz, A. M. (2000). Using Graphical Formats to Convey Uncertainty in a Decision Making Task. *Proceedings of the IEA 2000/HFES 2000 Congress* (pp. 13-16). Santa Monica, United States: Human Factors and Ergonomics Society.

Fout, N., & Ma, K.-l. (2011). *Reliable Visualization: Verification of Visualization based on Uncertainty Analysis.* Davis, California: UC Davis Computer Science Department.

Garrabrants, W. M. (2009). Reducing Cognitive Load Using Hypervariate Display. *Doctoral Dissertation*. Nova Southeastern University.

Girardin, F., & Nova, N. (2005). Getting Real with Ubiquitous Computing: the Impact of Discrepancies on Collaboration. *International Journal on Human-Computer Interaction*, *1*(1), 60-64.

Goldspink, C. (2002). Methodological implications of Complex Systems Approaches to Sociality: Simulation as a Foundation for Knowledge. *Journal of Artificial Societies and Social Simulation, 1*.

Haapalainen, E., Kim, S., Forlizzi, J. F., & Dey, A. K. (2010). Psycho-Physiological Measures for Assessing Cognitive Load. *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (pp. 301-310). Copenhagen, Denmark: ACM.

Hancock, P., Szalma, J., & Oron-Gilad, T. (2005). Time, Emotion, and the Limits to Human Information Processing. In D. McBride, & D. Schmorrow (Eds.), *Quantifying Human Information Processing* (pp. 157-175). Lanham, Maryland: Lexington Books.

Healey, C. G., Booth, K. S., & Enns, J. T. (1994). High-Speed Visual Estimation Using Preattentive Processing. *ACM Transactions on Computer-Human Interaction, 3*, 107-135.

Henricksen, K., & Indulska, J. (2004). Modelling and Using Imperfect Context Information. *Workshop Proceedings of the 2nd IEEE Conference on Pervasive Computing and Communications*, (pp. 33-37). Orlando, FL.

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating Cognitive Load Theory and Concepts of Human–Computer Interaction. *Computers in Human Behavior, 26*(6), 1278-1288.

Huck, S. W. (2008). *Reading Statistics and Research* (5th ed.). Pearson.

Juhnke, J., Mills, T., & Hoppenrath, J. (2007). Designing for Augmented Cognition - Problem Solving for Complex Environments. *Foundations of Augmented Cognition.* Springer.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York, New York, USA: Farrar, Straus, and Giroux.

Kahneman, D., & Tversky, A. (1982). On the Study of Statistical Institutions. *Cognition, 11*, 123-141.

Kalyuga, S. (2011). Cognitive Load Theory: Implications for Affective Computing. *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference* (pp. 1-6). Palm Beach, Florida: Association for the Advancement of Artificial Intelligence.

Kandogan, E. (2001). Visualizing Multi-Dimensional Clusters, Trends, and Outliers Using Star Coordinates. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 107-116). New York, NY, USA: ACM.

Kaptein, M., Nass, C., & Markopoulos, P. (2010). Powerful and Consistent Analysis of Likert-Type Rating Scales. *Proceedings of the 28th International Conference on Human Factors in Computing Systems* (pp. 2391-2394). Atlanta, Georgia: ACM.

Karr-Wisniewski, P., & Lu, Y. (2010). When more is too much: Operationalizing technology overload and exploring its impact on knowledge worker productivity. *Computers in Human Behavior, 26*(5), 1061-1072.

Kirsch, D. (2000). A Few Thoughts on Cognitive Overload. *Intellectia, 1*(30), 19-51.

Lapinski, A.-L. S. (2009). *A Strategy for Uncertainty Visualization Design.* Defence Research and Development Atlantic Dartmouth (Canada).

Lim, B. Y., & Dey, A. K. (2009). Assessing Demand for Intelligibility in Context-Aware Applications. *Proceedings of the 11th International Conference on Ubiquitous Computing* (pp. 195-204). New York: ACM.

Lim, B. Y., & Dey, A. K. (2011). Investigating Intelligibility for Uncertain Context-Aware Applications. *Proceedings of the 13th International Conference on Ubiquitous Computing* (pp. 415-424). Beijing, China: ACM.

Liu, D., Macchiarella, N., & Vincenzi, D. (2009). Simulation Fidelity. In D. Vincenzi, J. Wise, M. Mouloua, & P. Hancock, *Human Factors in Simulation and Training* (pp. 61-73). Boca Raton, Florida: Taylor & Francis.

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., & Hetzler, E. (2005). Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science, 32*(3), 139-160.

Mansi, G. (2011). An Assessment of Instant Messaging Interruptions on Knowledge Workers' Task Performance in E-learning-based Training. *Doctoral Dissertation*. Nova Southeastern University.

Marr, D. (1982). *Vision.* New York, New York: W.H. Freeman and Company.

Mühlhäuser, M., & Hartmann, M. (2009). Interacting with Context. *Proceedings of the 1st International Conference on Quality of Context* (pp. 1-14). Springer-Verlag Berlin, Heidelberg: ACM.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology.* W.H. Freeman and Company: New York, New York.

Nielsen, J. (1990). Evaluating the Thinking-Aloud Technique for Use by Computer Scientists. *Advances in Computer-Human Interaction, 3,* 197-216.

Norman, D. A. (1988). *The Design of Everyday Things.* New York: Doubleday.

Oulasvirta, A. (2009). Field Experiments in HCI: Promises and challenges. In P. Saariluoma, & H. Isomäki, *Future Interaction Design II* (pp. 87-103). Springer.

Oviatt, S. (2006). Human-Centered Design Meets Cognitive Load Theory: Designing Interfaces that Help People Think. *Proceedings of the 14th annual ACM international conference on Multimedia* (pp. 871-880). Santa Barbara, CA, USA: ACM.

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist, 38*(1), 63-71.

Parasuraman, R., & Caggiano, D. (2005). Neural and Genetic Assays of Human Mental Workload. In D. McBride, & D. Schmorrow (Eds.), *Quantifying Human Information Processing* (pp. 123-155). Lanham, Maryland: Lexington Books.

Petre, M., Blackwell, A., & Green, T. (1998). Cognitive Questions in Software Visualisation. In J. T. Stasko, J. Domingue, B. A. Price, & M. H. Brown, *Software Visualization: Programming as a Multimedia Experience* (pp. 453-480). MIT Press.

Piekarski, W., & Thomas, B. H. (2009). Through-Walls Collaboration. *IEEE Pervasive Computing*, *8*, 42-49.

Rahman, M. (2012). Emergency Medical Responders and Physicians: Diagnostics, Decision Making and Medical Management in High Stakes Situations. *Symposium on Human Factors and Ergonomics in Health Care*, (pp. 14-19). Baltimore, Maryland.

Reinhardt, W., Schmidt, B., Sloep, P., & Drachsler, H. (2011). Knowledge Worker Roles and Actions—Results of Two Empirical Studies. *Knowledge and Process Management, 18*(3), 150-174.

Ritter, F. E., Kim, J. W., Morgan, J. H., & Carlson, R. A. (2011). Practical Aspects of Running Experiments with Human Participants. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction* (pp. 119-128). Springer Berlin / Heidelberg.

Rukzio, E., Hamard, J., Noda, C., & De Luca, A. (2006). Visualization of Uncertainty in Context Aware Mobile Applications. *Mobile HCI.* Helsinki, Finland.

Sager, L., Grier, R., Jackson, C., Levchuk, Y., & Stelzer, E. (2007). Visualization Techniques for Revealing Uncertainty to Decision Makers. *Proceedings of the 2007 Human Systems Integration Symposium.* Annapolis, MD.

Santos, A. C., Cardoso, J. M., Diniz, P. C., & Ferreira, D. R. (2010). Challenges in the Development of Context-Inference Systems for Mobile Applications. *Programming Methods for Mobile and Pervasive Systems (PMMPS'10), workshop part of The Eighth International Conference on Pervasive Computing (Pervasive'2010).* Helsinki, Finland.

Sarter, M., & Lustig, C. (2009). Attention and Memory. In L. R. Squire, *New Encyclopedia of Neuroscience.* Oxford: Elsevier.

Sauer, J., Seibel, K., & Rüttinger, B. (2010). The Influence of User Expertise and Prototype Fidelity in Usability Tests. *Applied Ergonomics, 5*, 130-140.

Scerbo, M., Bliss, J., Freeman, F., Mikulka, P., & Robinson, S. (2005). Measuring Task-Related and Task-Unrelated Thoughts. In D. McBride, & D. Schmorrow (Eds.), *Quantifying Human Information Processing* (pp. 195-231). Lanham, Maryland: Lexington, Maryland.

Schriber, T. J., & Brunner, D. T. (2009). Inside Discrete-Event Simulation Software: How It Works and Why It Matters. *Proceedings of the 2009 Winter Simulation Conference* (pp. 216-229). Austin, Texas: Winter Simulation Conference.

Schunn, C., Kirschenbaum, S., & Trafton, J. (2003). *The Ecology of Uncertainty: Sources, Indicators, and Strategies for Informational Uncertainty.* Retrieved July 17, 2011, from US Airforce Air War College: http://www.au.af.mil/au/awc/awcgate/navy/nrl_uncertainty_taxonomy.pdf

Shagrir, O. (2010). Marr on Computational-Level Theories. *Philosophy of Science, 77*(4), 477-500.

Sharma, S., & Sharma, M. (2012). Low Intensity Conflict Stress in Soldiers: Building Coping Skills and Resources. *Psychological Studies, 57*(3), 260-268.

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2010). *Designing the User Interface* (5th ed.). Reading, MA: Addison-Wesley.

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2012). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin & Review*, 1-33.

Skeels, M., Lee, B., Smith, G., & Robertson, G. (2010). Revealing Uncertainty for Information Visualization. *Information Visualization, 9*(1), 70-81.

Smith, R. D. (2000). Simulation: The Engine behind the Virtual World. *Simulation 2000 Series*, 1-24.

Speier, C., & Vessey, I. (2003). The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance. *Decision Sciences, 34*(4), 771-797.

Spool, J., & Schroeder, W. (2001). Testing Web Sites: Five Users Is Nowhere Near Enough. *CHI '01 Extended Abstracts on Human factors in Computing Systems* (pp. 285-286). Seattle, Washington, USA: ACM.

Sweller, J. (1988). Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257-285.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory.* Springer.

Thomson, J., Hetzler, B., MacEachren, A., Gahegan, M., & Pavel, M. (2005). A Typology for Visualizing Uncertainty. *Conference on Visualization and Data Analysis 2005 (part of the IS&T/SPIE Symposium on Electronic Imaging, 2005)*, (pp. 16-20). San Jose, California, United States.

Triola, M. F. (2009). *Elementary Statistics* (11th ed.). Addison Wesley.

Tufte, E. (2001). *The Visual Display of Quantitative Information.* Cheshire, Connecticut: Graphics Press.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, 211*(4481), 453-458.

van der Kleij, R., de Jong, A., te Brake, G., & de Greef, T. (2009). Network-aware Support for Mobile Distributed Teams. *Computers in Human Behavior, 25*(4), 940-948.

Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Human Factors, 34*(4), 457-468.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure–Ground Organization. *Psychological Bulletin, 138*(6), 1172-1217.

Ware, C. (2004). *Information Visualization: Perception for Design.* Morgan Kaufmann.

Ware, C. (2008). *Visual Thinking for Design.* Morgan Kaufmann.

White, A. L., & Carrasco, M. (2011). Feature-based Attention Involuntarily and Simultaneously Improves Visual Performance across Locations. *Journal of Vision*, *11*(6), 1-10.

White, K. P., & Ingalls, R. G. (2009). Introduction to Simulation. *Proceedings of the 2009 Winter Simulation Conference* (pp. 12-23). Piscataway, New Jersey: IEEE, Inc.

Wickens, C. D. (2002). Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Zhao, W., Fang, T., & Jiang, Y. (2007). Data Fusion Using Improved Dempster-Shafer Evidence Theory for Vehicle Detection. *Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery* (pp. 487-491). IEEE Computer Society.

Zhou, F., Duh, H. B.-L., & Billinghurst, M. (2008). Trends in Augmented Reality Tracking, Interaction and Display. *Proceedings of the IEEE ISMAR* (pp. 193-202). Cambridge, UK: IEEE.

Ziemkiewicz, C., & Kosara, R. (2010). Laws of Attraction: From Perceived Forces to Conceptual Similarity. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, 16*(6), 1009-1016.

Zuk, T., & Carpendale, S. (2006). Theoretical Analysis of Uncertainty Visualizations. *Proceedings of the SPIE & IS&T Conference in Electronic Imaging: Visualization and Data Analysis*, *6060*, pp. 66-79.

Zuk, T., & Carpendale, S. (2007). Visualization of Uncertainty and Reasoning. *Smart Graphics* (pp. 164-177). Berlin Heidelberg: Springer-Verlag.