Use of Integrated Process Control Displays in Work System Design

Somchart Thepvongs

Thesis Submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial and Systems Engineering

Dr. Brian M. Kleiner, Chair
Dr. Robert H. Sturges
Dr. Robert C. Williges

June 4, 1998

Blacksburg, Virginia

Keywords: Macroergonomics, three-dimensional displays, control charts, mental workload, signal detection theory

Use of Integrated Process Control Displays in Work System Design

Somchart Thepvongs

Abstract

Given the continuing deployment of total quality control and total quality management initiatives by organizations, employees have seen changes in their work environment. Furthermore, the impact of downsizing has resulted in operators becoming responsible for the quality of their own processes. This study tested the impact of various display alternatives of control chart data on decision performance and mental workload. The control charts were shown as multiple two dimensional displays, a composite two dimensional display, and a composite three dimensional perspective display.

Multiple two dimensional displays were found to have significantly higher decision accuracy and decision confidence ratings than either composite displays. No significant difference in decision accuracy and decision confidence ratings was found among the composite displays. The type of display did not have a significant effect on decision time. Mental workload was also found to be significantly affected by the type of display used. Multiple two dimensional displays imposed significantly lower levels of mental workload than either composite display. No significant difference in mental workload was found among the composite displays. These results indicated that multiple two dimensional displays should be used when control chart data from multiple processes must be displayed.

Acknowledgements

I would like to dedicate this research to all of those who supported me along the way.

I would like to thank Dr. Brian Kleiner, Dr. Robert Williges, and Dr. Robert Sturges for their time and effort in advising and supporting me throughout this entire process. I cannot thank any of you enough for all the help you provided me in conducting this research.

I would like to dedicate this to Andy Thompson, a dear old friend who will never be forgotten.

I would like to thank my family. Mom and Dad, you provided me with so much support and encouragement along the way, I could have never done this without your help. I thank both of you for the guidance, patience, and your genuine belief in me. Your love has helped me come so far. I would also like to thank my sister and brother for their friendship and love during the years. I only hope I can be the same for both of you in the years to come.

I would like to thank Amanda for all the love and friendship throughout these many years. Your sincere belief in me has provided much of the strength and focus I needed along the way. I thank you for all the patience along the way.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Scope of Research

Organizations globally have been witnessing dramatic changes in their environment in the past decades. Ever since the introduction of computer technology into mainstream business, corporations have attempted to employ these technologies in the hopes of becoming more efficient in carrying out their business strategies. With the introduction of Internet technology and the growth of distributed computer systems, even smaller corporations are able to reach the global marketplace with relative ease. At the same time, organizations continue to change internally to make use of this technology to respond to the dynamic organizational environment.

The Changing State of the Work System

Many organizations have attempted to keep pace with ongoing environmental changes which have resulted in a variety of change efforts going as far as changing the organizational design. In recent years, TQM, and reengineering have received a great deal of notoriety in practitioner as well as academic publications. These change efforts have been used to guide organizational changes. With organizations adopting these change efforts, they have modified many internal operations to meet customer and stakeholder needs more efficiently. The typical vertical organizational design has slowly been replaced by horizontal designs to better meet environmental needs. The result of the changing organizational design has been the decentralization of decision making where decision making has been moved from top level management to the lower levels of the organization. The change in design has required a great deal of internal work for organizations to accommodate the new decision making authority.

Besides changes in organizational design, organizations have also attempted to develop performance measurement systems in order to better understand what is going on in their work system so they can react to the environmental changes. As a result, they are realizing that effective and efficient performance measurement systems are essential for long term success, whether through continuous improvement or through breakthrough innovations. As process improvement efforts, TQM and reengineering both played a role in the popularity of performance measurement. From TQM, experts such as Juran, Deming, and others introduced the importance of continuous improvement to attain incremental improvement. Before TQM, Shewhart and Crosby also made strides in the Quality Assurance area contributing to the body of knowledge of process improvement. These improvements would be attained using statistical procedures to measure process performance to support management by data. The ideas and theories behind these statistical procedures originated with some of the Total Quality Control gurus such as Feigenbaum and Shewhart. Similarly in reengineering, Hammer (1996) stated that, “if the process is to perform well, ‘well’ must be precisely defined in a way that is measurable, unambiguous, understood by everyone involved, and relatable to people’s own work.” (see Figure 1.1)

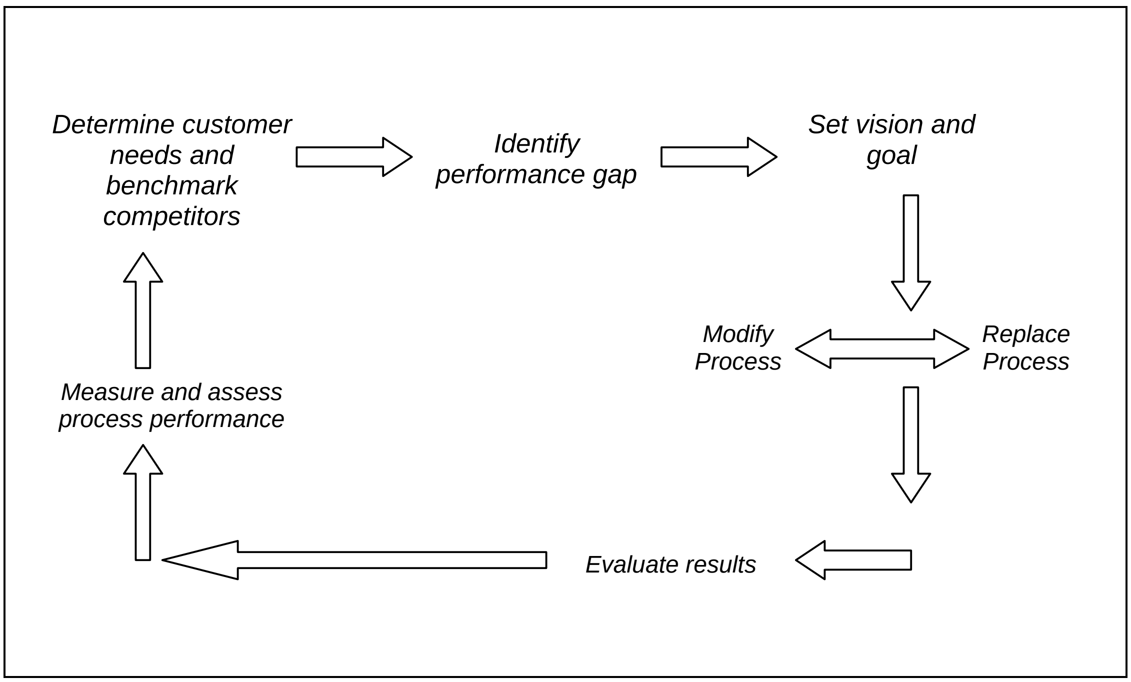


Figure 1.1: Role of Performance Measurement (Hammer, 1996)

Performance measurement systems act as tools for workers to make informed decisions. Employees still require the decision making authority to make use of these measurement systems. The move to decentralization of decision making allowed employees to make use of these measurement systems since they were empowered to act on the information from these systems.

However, a major organizational development that has also impacted work system decision making has been the downsizing of organizations. Downsizing has been used with the intent to improve the efficiency of the organization as well as to look better for organizational stakeholders. However, organizations which have gone through downsizing did not necessarily improve their organizational efficiency due to the lack of consideration of employee job design. The implication of a reduced work force was that the remaining employees were now required to assume the responsibilities and the necessary skills to make up for the downsized workforce. As a result, the job requirements were broadened demanding more from employees. An example of this can be illustrated in a manufacturing work environment. Previously, employees were asked to perform a narrow functional task along a manufacturing process. However with all of these organizational development, workers may now be responsible for a series of machines or processes as a result of the downsized workforce. With these new responsibilities the push to decentralize decision making added to the job requirement of employees since they were now given the autonomy to make the decisions regarding the multiple work processes for which they were responsible. In order to make these decisions, employees still needed the tools to make decisions based on information rather than intuition.

Decision making and the Work System

Simon (1968) defines management as a decision making process which encompasses a variety of activities:

1. *Intelligence activities*: activities dealing with environmental or system scanning
2. *Design activities*: activities dealing with the development of possible courses of action
3. *Choice activities*: activities relating to the selection of one or many alternatives from the design activities

The amount of time spent on these activities depends both on the characteristics of the decision as well as the decision maker. However, the decision making system not only consists of the decision maker, but also the tools to support information used in the decision making process. Mintzberg (1975) cites several reasons why decision makers may not use the tools which are provided to them. These reasons can be categorized into three areas:

Ineffective formal information systems: Decision tools do not portray rich data that managers require in a timely and effective manner.

Organizational constraints: Dysfunctional goals, political climate, and poor job design result in ineffective and inappropriate use of decision tools.

Human limitations: Poor human factors design do not take advantage of human capabilities.

Due to the commitment of organizations to information technology, many information systems provide instantaneous data to the decision maker. Software packages have been developed to provide real time data from the manufacturing floor. These packages go as far as to use tools, to convert raw data into a typical two dimensional portrayal. As information is portrayed, decision makers must perceive that information and arrive at a decision or an evaluation of the work system, from which action is taken. However, without careful consideration of the interface, and the decision maker's preferences, information can easily be misconstrued which would result in a poor decision or an inaccurate evaluation of a work system.

Macroergonomics: Theory for Work System Analysis and Design

Based on sociotechnical systems theory, macroergonomics is concerned with the optimization of organizational and work system design through consideration of relevant personnel, technological, and environmental variables and their interactions (Kleiner, 1997a). Macroergonomics has become especially important in investigating and optimizing the relationship between the human and the work system using ergonomic interventions to improve work system design. According to Hendrick (1991), these macroergonomic interventions can yield improvements of 60-90%. With sociotechnical systems design, computer supported work can benefit from this comprehensive approach.

STS theory is based on an open systems model which seeks to optimize the relationship between the personnel and the technological subsystems through joint design. The development of tools to support decision making lends itself to STS analysis and design and also addresses Mintzberg's concerns of decision tool usage. Control charts are a technological mechanism which converts data into information to support decision making.

As new portrayals are developed, the personnel subsystem will be impacted by these new decision tools through the interface between the personnel and technological subsystem. In developing and evaluating these tools, human processing capabilities must be addressed in order to maximize the effectiveness of the decision making tool.

With the changing work environment, organizations continue to look for ways to improve their work systems. Macroergonomics provides a systematic approach towards analysis and design of work systems to attain this improvement using change efforts. TQM, a practitioner oriented change effort, is aligned with the fundamental theories and concepts defined by sociotechnical systems theory (Kleiner, 1997b). At the same time, this theory also provides a methodology by which human factors can be introduced into work system design, such that the work system is compatible with requirements and limitations of people.

Given the changing nature of the work system environment, work systems can be improved by jointly considering both the personnel and technological subsystems to guide work system changes. As decision makers in the personnel subsystem, new job requirements and responsibilities require tools to facilitate their work and decision making. Technology can be used as a tool to meet these needs of the decision maker given the characteristics of the work system. Macroergonomics provides a framework that attempts to make the changing work environment, the people, and the technology all compatible to meet the goals and objectives of the work system.

Issues with Performance Measurement

A complex issue with performance measurement design, which has received less attention, has been the portrayal of information on these measurement systems. The portrayal of the data plays a vital role in any person’s decision making ability just based on the human’s information processing capabilities (Simon, 1990; O'Reilly, 1980). The ability of the decision maker to interpret the information in an efficient and accurate manner is an important characteristic of any decision making system. With the complexities involved with advancing technology, the importance of the interface has been shown to impact decision making performance (Morgan, Hershcler, Wiener, E., and Salas, 1993). As a result, a work system with a set of exceptional metrics can still fail any person who needs to make a decision. This failure can occur if the portrayal of the information does not accurately represent what is happening in the organizational system or if there is too much information for the user to process at once. Thus, the effectiveness of a performance measurement system depends on not only the work system metrics but also the portrayal system used. The portrayal system used represents the human-machine or human-system interface. A mismatch which occurs at this interface can limit the decision maker's ability to recognize and evaluate the results from the performance measurement system. An example of this interface can be found inside the cockpit of an airplane. Pilots have available to them a wide array of dials and gauges which represent the portrayal system used by the airplane manufacturer. This portrayal system helps the pilot to have a level of situation awareness of the airplane and make decisions based on the portrayal of information. However, humans do not possess an unlimited capability of processing information. Given the fact of these human limitations, a cockpit designer must decide how to design the portrayal system to ensure accurate, reliable and quick decision making. Design of this human-system interface can play a large role in the decision making capabilities of a human.

Development of Statistical Process Control

One of the major reasons many companies chose to implement TQM was due to the success of Japanese companies in the seventies and eighties. This success was quite remarkable considering their industrial capabilities immediately after World War II. Dr. W. Edwards Deming provided Japanese industry with some of the fundamental tools to implement TQM, although he would never recognize his teachings as Total Quality Management. With the help of Dr. Deming, Japan was able to rebuild their industries, and even more importantly shift from a country who was known for its cheap products immediately after the war, to a country with the capabilities of creating high quality manufactured products. However, this success could not be solely the result of one person's ideas. Shewhart and Feigenbaum provided many of the fundamental ideas to TQM through their development and implementation of scientific methods to measure and assess process and system performance through their Total Quality Control initiatives. After Japanese consumer products began reaching the United States marketplace, consumers would begin to take note of the quality of these products.

TQM prescribes various interventions which deal with both social and technical issues within the organization. Although many new interventions have been tacked onto the TQM movement, one of the tools, which has seen widespread application, is the use of Statistical Process Control (SPC). SPC allows organizations to chart data and make decisions based on data, not just intuition. Since work systems are inherently subjected to natural variation, decision makers must either track the progress of the system subconsciously, or through the help of SPC tools such as control charts. Control charts allows decision makers to identify trends, problem areas, and perhaps assign causes of unnatural phenomena of a work system. Control charts are a type of performance measurement system that attempts to provide a portrayal that meets the needs of the decision maker. Specifically, these control charts are graphical portrayals which attempt to help separate common causes and special causes using statistical methods. There has been a steady amount of research which has been concentrating on the statistical tools, such as control charts, which aim to provide the information a decision maker requires in a timely manner. The importance of implementing these tools has been shown to be a significant characteristic among successful organizations (Adhire, 1996). As a result, research in improving SPC has continued to improve the accuracy and efficiency of these tools.

With the continued usage of TQM in mainstream industry today, new statistical display tools can be developed to advance the movement given the changing work environment. As corporate downsizing continues, employees are forced to take on new job responsibilities adding to the complexity of their job design. At the same time decision making is being pushed down within the organization giving autonomy to those who may need the added responsibility. These new job responsibilities may include being accountable for more machines or processes, so that the workers can make decisions that have a more immediate impact on the process or system. Hence, a worker may require information from a variety of machines or processes to facilitate decision making. In order to make an informed decision, a worker must search for the proper charts or graphs relevant to the processes and machines of interest. A way to avoid the confusion associated with this search, is to create new displays which integrate information from the various processes and machines of interest. By matching the display, to the information needs of the decision maker, it is hypothesized that improved decision making and reduced workload will result.

Visualization

With recent technological advances, computer technology has become prevalent throughout organizations. Visualization provides a means to convert large complex data sets into information that people can understand. Software must be written, and hardware must be developed in order to provide a variety of visual representations of these complex data sets. At the same time, usability, training, and human information processing research must accompany these technological developments to insure that these technologies maximize the human's potential to process vital information. Over the past few years, AT&T (Eick and Fyock, 1996) has led many research initiatives in the area of visualization, as well as demonstrated the importance of these technologies to achieve their corporate objectives. Applications of visualization technology include identification of fraudulent calling, identification customers for a particular niche, customer migration tracking, and computer source code tracking which amount to billions of dollars in resources annually.   As organizations collect a vast amount of information for their databases, taking advantage of the massive data collection to provide a competitive advantage has been difficult for many organizations to achieve. The potential for visualization tools for organizational and work system decision making is enormous. Visualization is an enabling tool which allows decision makers to explore data through data mining techniques and to perceive information gathered from data in different perspectives in order to efficiently arrive at an informed decision. Decision makers who have access to information first will have a jump on making decisions to assess and evaluate work system performance, and take corrective action if necessary. For any work system, providing this invaluable information is vital in order to maintain the competitive edge.

As a result of the downsizing, people are assuming more responsibility at any given point in time. From a technology standpoint, this has meant feeding more information to decision makers to an almost crippling point, where the decision maker cannot possibly filter all the noise present in all the information they are provided. Visualization tools help consolidate and integrate information into a new portrayal. However, the major issue with visualization has been, whether decision performance measures are affected by the new portrayal tools.

Implications on Industrial Tasks

Today, a great deal of research continues in industry and academia dealing with quality and production. The research and work in this area includes the entire spectrum of the industrial process from its suppliers to its customers. However, one of the areas of most importance is the ability of companies to have the proper controls in order to make adjustments to their processes when products fail to meet their specifications. Control charts act as a tool to help provide some of the controlling functions within these processes, since they help illustrate a system's ability or inability to meet the specifications dictated by the customer of the process. Furthermore, these control charts may also provide details about the nature of variation within the system, whether it is due to natural causes or assignable causes. The ability of companies to control their processes is extremely important in a variety of ways. Companies have the knowledge of the work system's capabilities in order to understand the limitations of what can be offered to the customers. These control charts, among other SPC tools, are by no means perfect. Problems with human error in identifying trends or patterns are not unusual and continue to be investigated (Chih and Rollier, 1995).

The Need to Investigate the Effects of Integrated Portrayals of Control Charts

With the help of Shewhart (1931), control charts have typically been used as tools to help provide the vital information that companies require to control their work processes. Control is important in production and inspection tasks since defect parts that are beyond random variation need to be immediately addressed by the employees of the work system. Otherwise, the system will continue to make defect parts at the expense of the company, as well as risk the delivery of those defected parts to the customer. Tools such as control charts, which help provide insight on the system's capabilities, are important in providing information to decision makers of the work system.

Traditionally, these control charts have been two dimensional tools which have functioned to provide time related data pertaining to the quality of the product perhaps from a particular process or machine. By providing this data, decision makers were able to gain valuable information about their process in an attempt to achieve statistical control of their processes. However, for each and every machine or metric, a separate control chart has typically been used to represent the various states of each machine or metric. Organizations which implemented SPC tools provided information for a specific machine or process for which a particular employee was responsible. However, as global competition increased, organizations required more productivity from their processes. As a result, many employees in the manufacturing sector have undergone training and education in additional skills in order to maximize the potential unproductive time of an organization's workforce, as well as compensate for the downsized workforce. Employees have become multi-skilled, and have become responsible for a number of machines or processes, increasing the complexity of the employee's job requirements. At the same time, organizations have employed a variety of teaming interventions such as high performance work teams, cross-functional teams etc. to attempt to push decision making to the source of the work. As mentioned earlier, SPC has been used by those in management and on the manufacturing floor to measure work system performance.. The SPC tools developed by Feigenbaum and Shewhart provided the fundamentals for identifying the nature of variation for a particular process or machine.

Although some processes and systems have become more complex, it has been competition that has driven industries to improve their ability to adjust and respond to the voice of the process. The voice of the process is important since it provides a wealth of information regarding the capabilities and the state of the process. Those who cannot respond to the voice of the process produce rejected products which must be reworked or disposed, cutting into the organization's profit margin. By pushing decision making down to the source of the work, organizations hope to minimize the lag between the signals produced by the processes and the time required to make decisions. Reducing the lag reduces the amount of additional rejected products produced improving profits for the company, as well its reputation. In order to understand what the voice of the process is saying, a tool such as

SPC can be used to assist the human decision maker in making the right decision. Otherwise, the decision maker would rely on raw data which is difficult to interpret. SPC is a tool used to convert raw data into information, helping the decision maker understand the state of the process at relative points in time.

Given the added responsibilities of multiple processes or machines and more decision making authority of employees, new ways of providing information to the decision maker must be explored. One possible method of providing the vital information from these processes or machines is to integrate several control charts into one composite portrayal. Rather than having different control charts on separate portrayals, displaying all the charts on one composite portrayal may help improve the decision making ability of employees. The concept of integrating numerous displays into one display has been defined as the proximity compatibility principle (Wickens, 1992). The integration of several control charts into one display may minimize the visual search time required of decision makers to look at separate control chart data. This integration could come in the typical two dimensional coplanar display which would combine data from the various systems into one graph. The integration could also be the result of changing the perspective of the data to be a 3-D perspective rather than a 2-D coplanar representation of data. Haskell and Wickens (1993), Wickens and Todd (1990) have found different advantages for using both of these integrated approaches in different task environments. An advantage found for the tasks used in these studies using the 3-D perspective include the reduction of scanning and search time as compared to 2-D planar representations of similar data. Due to the advances in computer technology, integrating multiple charts can be done with relative ease.

Problem Statement

In the past, control charts have provided a wealth of data for decision makers. However, the displays for these control charts are by no means perfect, and research in improving these displays should be investigated. Given the nature of traditional control charts, development of modified control chart displays to provide decision makers with information more effectively must be explored. Additionally, consideration of the impact of these tools on humans should also be investigated.

Research Questions

This research focuses around work system design, specifically how integrated displays can be used to improve work system decision making.

This thesis will primarily focus around two questions:

1. How do the dimensions of a control chart portrayal effect task related outcomes?
2. What is the impact of the dimensions of a control chart portrayal on the mental workload of individuals?

Task Related Outcomes

Hypothesis 1: Individuals using composite three dimensional perspective control charts will have better decision accuracy than individuals using two dimensional control charts

Hypothesis 2: Individuals using composite three dimensional perspective control charts will require more decision time than individuals using two dimensional control charts.

Hypothesis 3: Individuals using composite three dimensional perspective control charts will have less decision confidence in their evaluations than individuals using two dimensional charts

Personnel Measures

Hypothesis 4: Individuals using composite three dimensional perspective control charts will exhibit more mental workload than individuals using two dimensional charts.

Research Model

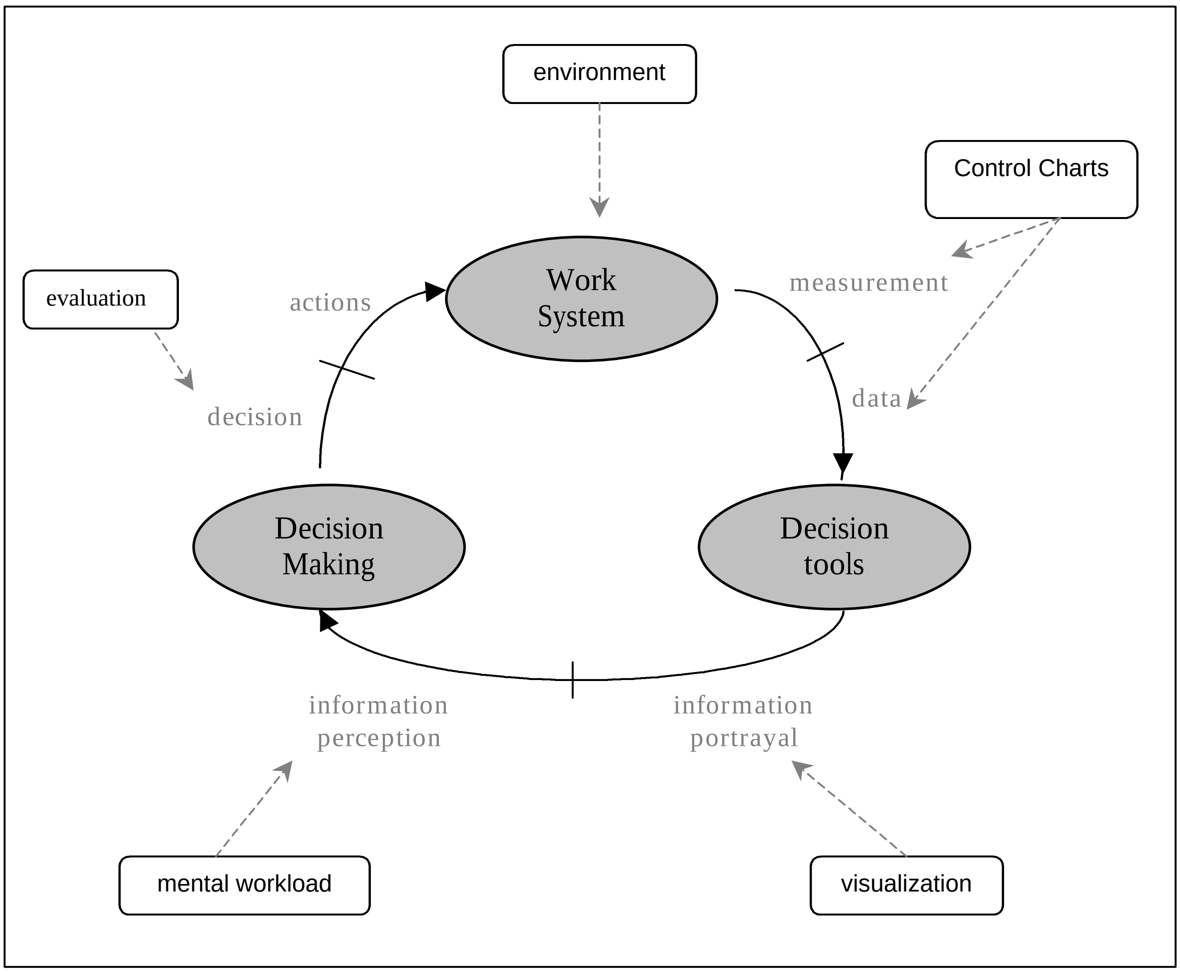
Figure 1.2 consists of the various components of the research model proposed in this thesis:

Independent Variables		Process Variables		Dependent Variables
• Multiple two dimensional portrayals • Composite two-dimensional portrayal • Composite three-dimensional perspective portrayal	→	• Mental Workload	→	• Decision Time • Decision Accuracy • Decision Confidence

Figure 1.2: Research Model

Conceptual Model

Figure 1.3 consists of the various relationships among the topics covered in this research:



Chapter 2: Literature Review

Introduction

This chapter provides a review of the literature and body of knowledge defined by the scope of this research. In order to provide the context by which the research will be conducted, a sociotechnical framework will be used to guide this research.

Sociotechnical Systems Theory and Macroergonomics

Sociotechnical systems emerged as a result of a set of British coal mining studies conducted by Emery and Trist in the early fifties. The importance of these studies brought forth the realization that organizational design affected its members in terms of performance and satisfaction. Beyond this realization, joint optimization considered both the demands of the social subsystem and the technical subsystem through joint design. Cherns (1987) defines several principles which dictate this joint design:

1. *Compatibility* between the final design and the intended design for which joint optimization is truly achieved.
2. *Minimal Critical Specification* states that tasks and/or roles should be specified only to the point to which it is necessary.
3. *Variance Control* indicates that variances which occur within a given system should be controlled closest to the source of that variance.
4. *Boundary Location* is essential to avoid the delay of information exchange, knowledge, and learning.
5. *Information Flow* must go to the correct people on time.
6. *Power and Authority* must go to those who require them in order to carry out their job effectively, or go to those who have expertise in a particular field.
7. *The Multifunction Principle* states that organizations must be able to adapt to their environment and perform additional functions which are dictated by their environment.
8. *Support Congruence* so that all aspects within an organization support synergy among functional units and teams within that organization.
9. *Transitional Organization* is part of the process of managing change which also requires careful planning and design in order to reach the goal of sociotechnical joint design.
10. *Incompletion or the Forth Bridge Principle* states that the job of design is never done, always require constant evaluation and revision.

These ten basic tenets provide the fundamentals of STS joint design and optimization. Macroergonomics utilizes these fundamental principles from STS as its basis to optimize the organizational and the work system. According to the Macroergonomics Technical Group of the Human Factors and Ergonomics Society, "Macroergonomics is concerned with the optimization of organizational and work system design through consideration of relevant personnel, technological, and environmental variables and their interactions."(Kleiner, 1997b) A conceptual model developed by Hendrick puts the various subsystems and their interrelationships into perspective:

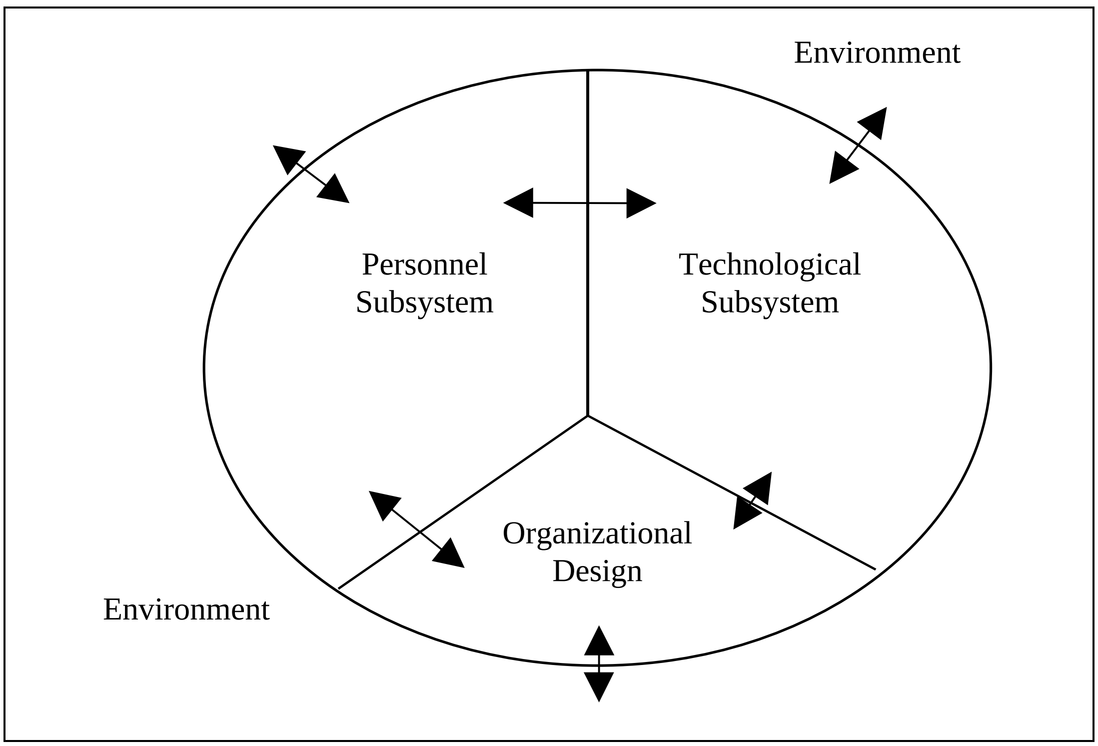


Figure 2.1: Components of Sociotechnical Theory (Hendrick, 1996)

This model (Figure 2.1) of a work system consists of the personnel subsystem (the people within the system), technological subsystem (the processes to get the work done), organizational design subsystem (the characteristics of the organization), and environmental subsystem (factors that are external to the system). The components of the system are free to interact and can have a profound effect on one another. This research focuses on the relationship between the personnel and technological subsystem in order to improve decision making within work systems. Due the technological interventions such as new portrayal techniques, considerations must be made on what impact these interventions have on people. Understanding exactly what people are able to handle in terms of information processing is paramount in designing new portrayal methods.

Personnel Subsystem

The personnel subsystem consists of the human aspects of an organization. This subsystem includes a broad range of attributes at the macro-level and the micro-level of the organization. These attributes include norms, procedures, values, organizational culture, motivation, rewards, etc. According to Taylor and Felten (1993), this subsystem functions to do the following:

> "1. Attain the system's primary goals
> 2. Adapt to the external environment for survival
> 3. Integrate internal environment for conflict management
> 4. Provide for the development and maintenance of the system's long-term needs." (p. 116)

Ideally, employees will be very motivated to perform these functions on behalf of the organization. In reality, the motivation does not always exist causing a work system to perform in a suboptimal manner. Organizations which design a job which is challenging and rewarding can help provide the necessary intrinsic motivation to perform these functions. Broad responsibility and decision making autonomy are ways in which organizations can make the job fulfilling. Similarly, rewards and recognition can be used to help supplement the intrinsic motivation incorporated in the work system. Together, these design options can be used to motivate employees to accomplish the functions of the subsystem which ultimately help the organization meet its goals. Without this motivation, unrealized performance improvement may result without the integration of work system design which considers both personnel and technological factors.

To characterize this subsystem, Hendrick (1991) stated that personnel subsystem consisted of two dimensions: degree of professionalism, and psycho-social characteristics. In dealing with work system evaluation and ultimately work system decision making, changes in the work system will ultimately effect the people. In this research, consideration of this subsystem will be made through the assessment of individual mental workload.

Technological Subsystem

This subsystem ultimately effects the social subsystem through the technical demands which create the roles people are engaged, and should be carefully considered when implementing technological changes (Pasmore, 1988). Technology can effect any combination of individuals, teams, departments, or the entire organization. The technological system consists of the tools, guides, procedures, machinery, processes, and technology used to transform inputs to outputs. Perrow (1967) classified this subsystem along two dimensions: task variability, and problem analyzability. Technological considerations for this research primarily rest upon the display tool, which can utilize two dimensional display space or three dimensional display space to portray data.

Environmental Subsystem

Joint optimization can only be achieved within the context of the system's environment. The entire sociotechnical system as well as its individual components must deal with the environmental conditions and constraints and react accordingly in order to maintain joint optimization. Work systems must deal with a variety of environmental factors. It is the work system's ability to deal with those factors, and to set up the proper interfaces, which will dictate the level of joint optimization and ultimately the level of work system success. Hendrick (1991), characterizes the environmental subsystem along two dimensions: degree of change and degree of complexity. Emery and Trist (1965) state that environmental uncertainty plays the largest role in work system effectiveness out of all the STS subsystems. Within the context of work system decision making, the environment can change the state of the work system at any given point in time and change the relationships between the environment and the sociotechnical components.

Organizational Design

An organization is a coordinate unit consisting of at least two people who function to achieve a common goal or set of goals (Gibson, Ivancevich, and Donnelly, 1994). The structure of organizations can be classified along three dimensions (Hendrick, 1991):

Complexity: the degree of differentiation and integration that exists in an organization, where complexity is subdivided into vertical differentiation, horizontal differentiation, spatial differentiation, and personnel differentiation.

Centralization: the degree to which formal decision making is concentrated in an individual, unit, or level.

Formalization: the degree to which jobs within the organizations are standardized (Hendrick, 1991)

Macroergonomics can provide a framework to justify this research using terminology mentioned in this section. As organizations have downsized, the implications on the manufacturing floor have been far reaching. With a downsized workforce, there are fewer people to perform the functions required by the organizations. On the manufacturing floor this has resulted in less complexity, since one worker is now responsible for multiple machines or processes. At the same time, organizations were attempting to push decision making down to the manufacturing floor in order to be in a position to respond to the needs of the environment. Furthermore, by pushing decision making authority down, employees were also better able to correct some of the problems which originated on the manufacturing floor. By pushing decision making down, organizations were also providing some of the intrinsic motivation necessary for employees to carry out the goals of the work system. Therefore, individuals became more empowered to make decisions on the manufacturing floor regarding multiple machines and processes. However, in order to maintain order within the organizational system, some centralization of decision making was required to maintain order within the manufacturing system. Decision making requires information from the processes for which employees are responsible. Due to this changed work system environment, technology can be used to accommodate the needs of the decision maker. One possible way to fulfill these needs is to provide integrated displays, which provide information from multiple processes or machines onto one display so that the employee can make an informed decision. Without the necessary decision making authority, the effectiveness of integrated displays becomes minimal.

Total Quality Management (TQM)

Although there is no universal definition of TQM, there are many components that are common among definitions. Kleiner and Hertweck (1996) define it as a people-focused approach whereby teams that use quality tools and technique continuously improve cross-functional processes with the aim of increasingly satisfying customers. Similarly Taveira and Smith (1996) define it as:

> "an approach for continuously improving the quality of goods and services delivered through the participation of all levels and functions of the organization. Three main principles in TQM are customer focus, continuous improvement, and teamwork"

However, one area of agreement among most TQM practitioners is the importance of SPC tools such as control charts to help achieve the goals of continuous improvement through the management by fact. In order to achieve continuous improvement, tools such as control charts must be used to gather necessary data to understand where a system has been, what the system is doing now, and what can be expected in the future. The notion of utilizing scientific methods to help provide insight into system or process performance was popularized by Feigenbaum and Shewhart through the Total Quality Control(TQC) movement, a perspective which shares many of the same ideals as the TQM movement. Juran (1988) also stated that quality control is an essential mechanism for sound quality improvement as well as for sound quality planning. One such application would be in the Plan-Do-Study-Act (PDSA) cycle used by many quality improvement initiatives. This cycle provides a basic framework for implementing system change using an incremental approach. The cycle begins in the plan phase where necessary plans are developed to improve or to innovate a system. Once the plan is completed, the plan is executed in the Do phase, and then monitored in the Study phase. In the study phase, control charts can function as powerful tools to monitor system performance, especially when measurements are taken before the new plan are implemented. Taking historical data before the plan is implemented provides a baseline to which the new changes can be compared. Once the data is studied, then necessary corrective action can be taken based on the data in the Act phase. The PDSA cycle provides a systematic methodology for designing and implementing process or system change.

Theory of Variation

According to Deming, variation is a part of every process or system. Variation may be visible in the number of defects in a batch of products, the temperature of a furnace, the suppliers' goods, the accuracy of quality control measurement systems, and ultimately the process or system's final product. Although variation is inherent to all systems and processes, variation can be both understandable and controllable. The nature of variation has been categorized into two areas by Deming (1986):

1. Special causes deal with variation beyond what is naturally inherent in a system. This is synonymous with the term "assignable causes" used by Shewhart.
2. Common causes are those factors which are inherent to the system and are present day to day.

The nature of variation is extremely important in identifying the sources of quality problems. Furthermore, identifying the type of variation within the system is essential for taking the necessary corrective action if action is required at all. Ultimately, Deming (1986) estimates that ninety-four percent trouble in a work system or opportunities of improvement can be linked to common causes, while the remaining six percent can be attributed to special causes. Management must deal with common causes since they are the ones who are empowered to make changes which affect an entire process or system. It is management's duty to listen to the voice of the process which will dictate much of the variation resulting from common causes. Special causes are to be dealt with by the worker since they are the ones who are directly involved with the process and have a better understanding of the variance which may occur in the specific process or machine.

In order to control variation, one must first begin to understand it. Typically, variation can be understood through the use of a control chart (see Figure 2.2) which is a graph that displays data within the context of limits based on process variation calculated from gathered data. Shewhart (1931) explains the concept of control as the condition from which past experience may help predict future behavior of a process or system within certain limits.

Based on these limits, various data patterns can be identified by a decision maker of the system to determine whether the system is in control or out of control. In order to determine this, the decision maker must identify data on a control chart as a run, trend, periodical pattern, or data that has a tendency to hug a control line. Based on a recognizable pattern, corrective action must be taken.

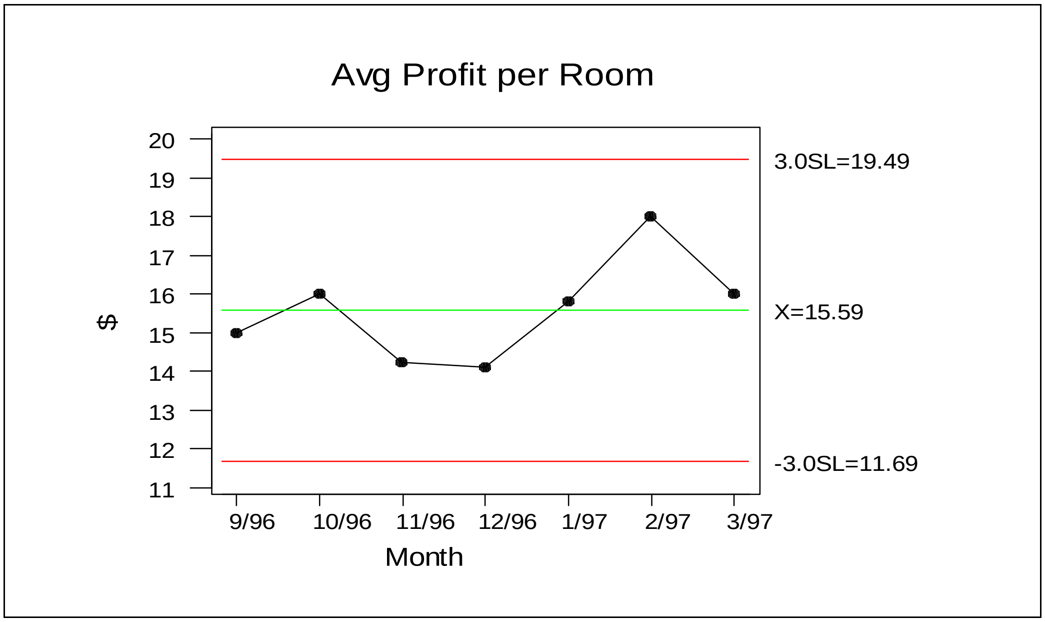


Figure 2.2: Example of a Control Chart

However, without control charts, managing a system becomes a greater challenge. Assuming that a manager decides to take corrective action due to variation from common and special causes, managers are capable of making two mistakes. According to Deming(1986), people can either:

> "1. Ascribe a variation or a mistake to a special cause when in fact the cause belongs to the system (common causes)
>
> 2. Ascribe a variation or a mistake to the system (common causes) when in fact the cause was special)" (p. 318)"

To help managers in their quest for organizational success, Shewhart devised three sigma control limits to shed light on the complexity of system behavior. If the process yields data that falls outside of the control limits, the variation is more likely due to special causes which can be controlled by the line workers. However, should no data fall outside these limits, the process is said to be in statistical control. Thus, variation which does occur is inherent in the system, and due to common causes. These common causes can be removed by management through the implementation of new technology, new processes, employee training, or other improvement initiatives.

Typically, processes which are in statistical control exhibit variation which is normally distributed around an arithmetic mean. This normal behavior is possible due to the Central Limit Theorem which is valid given certain conditions explained by Wapole and Myers (1993):

> "The normal approximation for X bar will generally be good if $n \geq 30$ regardless of the shape of the population. If $n < 30$, the approximation is good only if the population is not too different from a normal distribution and, as stated above, if the population is known to be normal, the sampling distribution of X bar will follow a normal distribution exactly, no matter how small the size of the samples" (p. 216)

Given the data, the control limits, and the mean of the process, a decision maker must decide if a non-random behavior exists (out of control) or if the process is behaving randomly (in control). In order to recognize the existence of non-random behavior, several rules have been developed to guide the decision maker in assessing the state of the process. Although other specific rules may exist, Hoyer and Ellis (1996) provide a general set of rules for identifying the existence of special causes of variation:

1. A point is above the upper control limit or below the lower control limit
2. Of three consecutive points, two points are more than two standard deviations above(or below) the centerline
3. Two consectuive points are more than two standard deviations above(or below below)the centerline
4. Of five consecutive points, four are more than one standard deviation above(or below) the centerline
5. Four consecutive points are more than one standard deviation above(or below) the centerline
6. Seven consecutive points above(or below) the centerline
7. Six conseuctive points are in a monotone increasing(or decreasing) pattern (count the first and last points in the pattern)

8. Of 10 consecutive values, there is a subset of eight that are in a monotone increasing(or decreasing) pattern
9. Given two consecutive points, the second is at least four standard deviations above(or below) the first

Variation and its effect on Quality

Although management may hold the key to affecting process or system change, the sources of the quality problems reside in a variety of categories according to Feigenbaum (1983):

1. Markets – The need for companies to be responsive to the organization's environment through organizational flexibility.
2. Money – Problems which arise from international competition as well as organizational investments have impacted profits and ultimately quality programs.
3. Management – With increasing organizational complexity as a result of new roles and responsibilties spawned from quality control initiatives, management must delegate new roles and responsibilities throughout the organization.
4. Men – Reflect the specialization of employees required for complex and technical problems.
5. Motivation – To achieve success, quality is built into every part of the process. As a result, organizations must find new ways to incentivize employees to contribute to the organization's quality programs.
6. Materials - With more advanced technology, new materials, specifications, and applications require advanced inspection techniques in order to monitor the acceptability of quality of incoming materials.
7. Machines and mechanization – In response to increasing customer demands, companies find themselves depending more and more on automation and machinery. As a result of this dependence, the quality of material supplies has also become an important factor in maximizaing labor and machine utilization rates.

8. Modern information methods – With the proliferation of modern computing system, information technology has witnessed a great deal of growth to provide the necessary control and management support required in decision making of work systems.
9. Mounting product requirements – The results of increasing customer demands have affected the requirements necessary to insure quality of the products, its safety, as well as well as its reliability.

Given the possible sources of variation, control charts can act as a tool to provide vital information in a timely manner. Although variation can come from a variety of sources, when speaking of a specific process or machine, obviously the sources of quality mentioned above may not affect the process or machine equally. Wheeler (1993) states that although we live in the information age today, manufacturers still struggle today even with the wealth of data collected as a result of the technology. Furthermore, he states the importance of tools such as control charts to provide knowledge that can sometimes only be gained cumulatively. Thus, the monitoring of system/process characteristics and attributes over time can yield this cumulative knowledge through the use of control charts.

Justification for Control Charts

Based on historical data, control charts help determine what may happen in the future in a particular process. The advantage of control charts, a type of line graph which depicts patterns over a time-dependent measure, has been tested to some degree in a controlled laboratory setting. Based on the results of an experiment, Dickson, DeSanctis, and McBride(1986) found that subjects using line plots to identify time related patterns had better decision quality than those using tabular data. The ability of decision makers to use control charts hinges upon the skills of the decision maker to correctly identify patterns which indicate a signal within a process.

Organizations which want to understand their processes look to control charts to provide them with the information they require. Scientific methods, as well as training, and benchmarking, are some of the more popular interventions implemented by organizations looking to apply TQM theory. Control charts allow decision makers to management by fact and theory which provide a sound as well as consistent approach to management. An approach grounded on statistical theory, which enables decision makers to make more informed decisions. Sadly, it has been shown that scientific methods, like control charts, are not extensively used within an organization (Ebrahimpour, 1985; Modarress and Ansari, 1989; Schroeder, Sakakibara, Flynn, and Flynn, 1992). To motivate other organizations the Malcolm Baldrige National Quality Award (MBNQA) helps provide some extrinsic motivation for organizations to implement control charts as well as other SPC tools. This motivation is the result of the MBNQA's core values which profess management by fact which can be enabled by these tools. However, receiving this prestigious award has not guaranteed the viability or sustainability of the award recipients.

One implication of control charts on the personnel subsystem is the need for training in the use of control charts. To avoid tampering by decision makers, the need to provide adequate training is paramount for effective evaluation of the work system or process and ultimately decision making. It is imperative that decision makers understand what the possible patterns which may be indicative of a trouble sign for a particular process or machine. Although the fundamental knowledge to use such tools may be no more than a series of data pattern descriptions, ultimately users should have some understanding of the underlying math and statistics involved in the construction of these control charts.

Decision Making and Graphics

Control charts provide the vital information that employees require to make informed decisions. Before the information can be perceived by the decision maker, it must be portrayed in a format which accurately represents what is going on in the system, and provides the information to the decision maker. Without an effective information portrayal system, information can be either lost or misconstrued by the decision maker. Therefore it is imperative that the effectiveness of the portrayal be evaluated so that the decision making process is done accurately and quickly.

The importance of information portrayal and its effect on the decision making process has been researched in previous studies, many of which have concentrated on decision making performance. Wheeler (1993) suggests that tables may not be sufficient enough in convey information from data. Furthermore he explains that people are visually oriented, therefore graphs should be portrayed along with the tabular data. Benbasat and Dexter (1985) found data to support this. Under experimental conditions, they found that a report which consisted of both tabular and graphical portrayal of the same data resulted in better profit performance and was the preferred presentation method over presentations with graphs or tables. For this experiment, the problem solving task was time constrained to fifteen minutes.

A similar research study was conducted by Lee and MacLachian (1986), where they studied the effects of three dimensional imagery on the interpretation of data. Lee and MacLachian utilized stereoscopic presentations by displaying binocular disparity information. Using this method of portrayal, the information would appear to be coming out of the screen, thus adding a third dimension to the data. Scatterplots and block diagrams were used to portray the data. Lee and MacLachian found that groups using 3-D scatter plots had better decision accuracy and had faster decision times than those groups using 3-D block diagrams.

Barfield and Robless (1989) also looked at the effects of dimensionality of portrayals on the decision making performance of experienced and novice decision makers. They also attempted to identify whether the medium of portrayal had an effect on decisional characteristics. In their results, Barfield and Robless found that novice decision makers had better decision accuracy using two dimensional graphs on paper, while experienced decision makers had better accuracy using three dimensional graphs on computers. They also found that both experienced and novice decision makers had higher decision confidence when using two dimensional graphs in general. Within their publication, Barfield and Robless (1989) summarize other studies dealing with graphics and decision making in the following table:

Table 2.1: Summary of previous research studies involving decision making and graphics (Barfield and Robless, 1989)

Reference	**Variable of interest**	**Subject population**	**Task/dependent variable**	**Result**
Chervany and Dickson (1974)	Raw data vs. statistically summarized reports	22 graduate students	Time to reach solution, confidence in solution	Subjects with summarized data made higher quality decision, had less confidence in the quality of their decisions, took longer to make decisions
Benbasat and Schroeder (1977)	Tabular display vs. 2-D graphs	32 business students	Time to analyze reports, cost performance, number of reports requested	Groups with graphs had lower costs and increased effectiveness of decisions and used fewer reports
Lucas and Nielsen (1980)	CRT displayed information vs. teletype terminals; graphs vs. tabular reports	36 graduate business students, 36 practicing engineers	Learning, rate of profit increase, cumulative profit	Groups using CRTs had higher profits and a higher rate of profit increase than groups using teletypes; no superiority for graphics was shown across groups, but graphics assisted engineers in rate of profit increase
Lucas (1981)	Tabular displays vs. 2-D graphs	119 middle- and upper-level managers	Selecting quarterly reorder quantities	Tabular display group found the output more useful than graphical display group
Watson and Driver (1983)	3-D graphs vs. tabular presentations of data	29 upper-level business students	Rank ordering of graphical information for conditions of immediate and delayed recall	No difference in recall of information for 3-D graphs condition over tabular presentations
Powers, Lashley, Sanchez, and Shneiderman, (1984)	2-D tables and graphs vs. either alone	74 undergraduate computer science students	Comprehensive of information displayed by tables or graphs	Tabular displays significantly increased comprehension, graphs and tables combined produced slower but more accurate performance

Benbasat and Dexter (1985)	Tabular display vs. 2-D graphs	65 business students	Time to process data, profit performance	No difference between the tabular and graphical groups was found in terms of profit performance although, groups receiving both conditions outperformed those receiving only one; subjects with graphs took less time for decision making than subjects using tabular displays (15-min problem-solving time limit)
Dickson, DeSanctis, and McBride (1986)	2-D graphs vs. 2-D tabular reports	154 upper-level business students	Readability, interpretation, accuracy, accuracy of decision making	No difference for dependent variables between 2-D graphs and 2-D tabular displays
Dickson et. al (1986)	Line plots vs. tables	320 upper-level business students	Interpretation of data points and trends, decision quality (forecast accuracy)	No difference between line plots and tables on interpretations accuracy, decision quality better for line plot group
Dickson et al. (1986)	Tables vs. graphs	363 upper-level business students	Global understanding of data	When large amounts of data were presented (nine graphs or tables) subjects with graphs outperformed those with tables
Lee and MacLachian (1986)	Stereoscopic scattergrams vs. stereoscopic block diagrams	45 lower-level business students	Interpretation, accuracy, speed of decision making	For 3-D presentations, scattergrams led to more accurate and faster decisions than 3-D block diagrams
Benbasat , Dexter, and Todd(1986)	2-D graphs vs. tabular displays, monocolor vs. multicolor	35,65,66 undergraduate and graduate marketing students (three experiments)	Information usage in terms of number of trials, report usage, time to examine reports, profit	Subjects with tabular displays requested more reports than graphical display group; subjects with tabular displays took longer to examine reports than graphs group; graphical presentation group used slightly more trials to complete the task than the tabular group; multicolor reports lead to 10% profit improvement over monocolor, no different in several performance times between multicolor vs. monocolor

Perspective in displays has also been an area of interest for researchers as applications in cockpit design, scientific visualization, and medical imaging systems are being realized. Research dealing with perspective continues to increase as researchers realize the potential benefits of perspective in display design. Perspective is one method used to represent 3-D environments by creating depth cues along the observer's line of sight (Delucia, 1995). Traditionally, a plan view (top-down perspective with 90° viewing angle) has been used to view data. Perspective views can enhance performance depending on the specific task (Bemis, Leeds, and Winer, 1988; Ellis, McGreevy, and Hitchcock, 1987). At the present time, a general guideline, based on data from aircraft display tasks, suggests that a forty-five degree elevation angle may be an optimal angle (Hickox, and Wickens, 1997; Kim, Ellis, Tyler, Hannaford, and Stark, 1987; Yeh, and Silverstein, 1992). However, incorporating perspective may introduce some ambiguity that may not be present in two dimensional displays. In part, this ambiguity is the result of displaying three axes on a two dimensional screen and may affect performance and situational awareness (McGreevy, and Ellis, 1986; O'Brien, and Wickens, 1997)

This research will expand upon the body of knowledge of past research since the domain of this research manipulates the use of dimension or perspective to display control charts data in a work system decision making environment.

Mental Workload

Mental workload has traditionally been an important area of interest in the aerospace industry, however it has also been applied to driving tasks and other system designs. However the use of mental workload in evaluating the load placed upon individuals in human computer interaction has been extremely limited thus far. Mental workload is an important aspect in the interaction between human and machines, or human and systems. Humans have a limited amount of cognitive capacity which must be allocated to processing information and executing tasks. This idea is extremely important in jobs which place demands on individuals who ultimately have limited cognitive capacity. Bailey (1983) estimates that thirty-five percent of all computer system errors is attributable to factors such as fatigue, motivation, and the worker's cognitive processes. With this fact in mind, identifying areas of improvement based on mental workload measurement can be vital in reaching the full potential of a system. Zhao and Salvendy (1996) found that for linear tasks, graphical presentations resulted in higher mental workload than using alphanumeric presentations. Interestingly, they also found that for branching tasks, mental workload was significantly lower when using graphical presentations by about 35%. Thus, it appears that the compatibility of portrayal or presentation on mental workload is relatively dependent upon the nature of the task. Control chart decision making is similar to branching tasks since decisions based on data are conditional upon the pattern recognized.

In dealing with mental workload, the problems of an agreed upon definition as well as a clear cut methodology for assessing workload continue to face researchers. Although no agreed upon definition exists, it appears that researchers in this area have utilized a more common sense definition of mental workload which deals with the degree to which a person is mentally occupied (Sanders, 1977). Gopher and Donchin (1986) define mental workload as a construct which characterizes the amount of cognitive resources required by the operator in order to complete a specific task. Assessing mental workload continues to be an important area in research because of the advancements of technology as well as the new demands placed upon employees. For the most part, measurement of workload can be broken down into several categories (Sheridan and Stassen, 1977):

1. Physiological variables – deal with those measures such as heart rate variability, respiration rate, electroencephalogram (EEG), etc.
2. Secondary methods – utilize secondary tasks to assess how much mental load can be added or subtracted from the system of interest.
3. Task analysis – measures which are based on information theory which requires the decomposition of tasks in order to assess choice and time uncertainty.
4. Attention allocation – measures usually based on eye movement which is presumed to be an indicator of mental load
5. Subjective measures – utilize information gathered from the subject to assess the level of workload.

Of these techniques, Casali and Wierwille (1983) state that subjective rating scales are the primary instruments for identifying mental workload. The reason for this appears to be because of the distinct advantages as compared to other assessment techniques (Shively, 1986):

"1. They allow a single scale across systems and tasks, (i.e. Cooper and Harper Scale, Subjective Workload Assessment Technique and the NASA Task Load Index)
2. they provide insights into the causes of performance degradations
3. they are relatively ease and inexpensive to implement,
4. several years of laboratory investigation provides theory for interpretation, and
5. real-world applications have demonstrated the validity of the technique." (p. 908)

As a result, subjective rating scales will be the methodology of choice to assess mental workload in this research. Work system evaluation is a cognitive task that draws upon the individuals ability to process and assess information available to them. Borg, Bratfisch, and Doring (1971) found that subjective scales highly correlate with objective measures of mental workload. However, Hendy, Hamilton, and Landry (1993) have differentiated the usefulness of various subjective rating scales. Specifically, they attempted to identify the value added benefits of scales based on the composite of multi-dimensional data such as the commonly used NASA Task Load Index (TLX) (Hart and Staveland, 1988), and the Subjective Workload Assessment Technique (SWAT) (Reid and Nygren, 1988). Based on a review of prior studies, Hendy et. al (1993) concluded that univariate ratings of workload provide a more sensitive measure of workload over NASA TLX, and SWAT. However, Hendy et. al (1993) by no means discounted the ability of NASA TLX or SWAT in assessing perceived mental workload. The usefulness of NASA TLX, and SWAT remain in the ability to isolate those factors contributing to the perceived level of workload. Of the two commonly used approaches, Hill, Zaklad, Bittner, Byers, and Christ (1988) found that an overall scale such as NASA TLX was superior to SWAT in terms of measurement sensitivity to workload. The reason for this being that SWAT requires subjects to make ratings along various comparisons using a three point scale: high, medium, or low. Hart and Staveland (1988) point out that many tasks do not impose extremely high, or extremely low workload.

As a result, they state that the SWAT may not be sensitive to discriminating differences among tasks which impose only a moderate amount of workload.

One other popular rating scale used by researchers studying mental workload is the Modified Cooper-Harper scale (Wierwille and Casali, 1983). This scale utilizes a set of decision rules based on a decision tree in order to assign a global value to mental workload, however fails to point out the source of workload as NASA TLX and SWAT are able to do. This scale utilizes a ten-point absolute scale which ranges from very easy to impossible. Due to this lack of diagnosticity, the NASA TLX methodology will form the basis for mental workload assessment for this research due to its lack of intrusiveness, diagnosticity, and sensitivity.

NASA Task Load Index

The NASA TLX was developed by Hart and Staveland (1988) as a tool to assess mental workload. Since mental workload is a hypothetical construct which may have a different meaning to different subject, NASA TLX assumes that workload is multi-dimensional. That is, depending on a given situation, a series of characteristics help define the perceived workload. Using controlled experiments, Hart and Staveland (1988) derived six specific scales to describe mental workload:

1. mental demand: the amount of mental and perceptual activity required by an individual (MD)
2. physical demand: the amount of physical activity required by an individual (PD)
3. temporal demand: the amount of pressure in terms of rate or pace demanded by the task (TD)
4. (observed) performance: the amount of satisfaction perceived by the subject or the amount of perceived success in completing task requirements, experimenter goals, or subject goals (OP)
5. effort: the amount of physical and mental effort required to achieve the level of performance for the task (E)
6. frustration level: the amount of frustration, irritation, stress, or insecurity experienced by the subject in completing the task (F)

After undergoing the experimental tasks, subjects will be asked to conduct pairwise comparisons of all the factors to indicate the relative importance of the scales to overall perceived workload. This allows subjects to identify those specific factors which are most relevant to the task or information portrayal method in determining perceived workload. The Once all pairwise comparisons are completed, subjects must evaluate all six subscales along bipolar anchors(high and low), either graphically or through interval scales. If an interval scale is going to be used, Hart and Staveland (1988) suggest that the total number of intervals does not exceed twenty because of the subject's inability to discriminate workload to such a fine degree. The graphical procedure typically consists of a 12cm line to represent the interval scaling and the bipolar anchors. The line is typically quantified using a scale of 1 to 100 with 1-point increments. A separate evaluation along all six subscales must be completed for each task or treatment. Afterwards, the experimenter will determine a weighted workload score based on the pairwise comparisons and the data from each subscale along each task.

The following equation characterizes the global measure of workload determined by NASA TLX (Nygren, 1991) where t represents the workload estimate for individual n.

$$W(t_n) = w_{n1} * v(MD_{n1}) + w_{n2} * v(PD_{n2}) + w_{n3} * v(TD_{n3}) + w_{n4} * v(OP_{n4}) + w_{n5} * v(E_{n5}) + w_{n6} * v(F_{n6})$$

Visualization

Visualization tools can be used to provide composite displays to improve decision making. Visualization is one tool which can be used to change the perspective by which data is viewed by the user. Based on human limitations, at the composite displays may reduce visual search and decision time.

Within the last decade, technological advances have facilitated the growth of visualization usage as well as visualization research. These advances provide the platform for which powerful graphics intensive display and manipulation are now possible and accessible to the public. According to Wickens, Merwin, and Lin (1994), visualization "can be described both as a technique (e.g., incorporation of stereo, motion) and as a cognitive operation (forming a mental image of the data relationships to facilitate insights regarding the relationships and constraints within the data." (p. 45) Industry has also identified the need for visualization research. An example of these needs include the following from the Oil-Gas Industry (Petzet, 1996)

- "Technical and management people must believe that visualization technology is accurate and useful for timely solution of real problems.
- The beauty of a display can overshadow information significance or accuracy, with the danger that visualization can create a false reality if applied too zealously to sell a prospect or simulate a reservoir.
- The main end-use issue emerging from the conference was the need for understanding and displaying uncertainty in visualizations. Participants also noted rigorous treatment of fault topologies and stratigraphic geometrics as particularly important.
- More examples of visualization success and failure need to be presented in scientific meetings and journals."(p. 21)

Research in the area of visualization has also been growing. One of the purported benefits of visualization is the added dimensional perspective permitted in 3-D portrayals (see Figure 2.3 for example). This requires that users mentally represent data which may be in 3-D perspective on a two dimensional screen. The value of the added dimensional perspective has not been clear based on past research. It has been hypothesized that three dimensional perspective should facilitate the ability of users to integrate information more easily since relevant information is portrayed close together. Liu and Wickens (1992) found data to support this hypothesis, while Wickens and Todd (1990) found opposite results. According to Wickens et. al. (1994), the use of graphical portrayals can be summarized as the following "'No single graphical format is optimal for all purposes.'" However, researchers (Carswell, 1990; Wickens, and Carswell, 1995) have suggested some fundamental principles regarding the location of displays within a particular portrayal. One of the areas of research includes the Proximity Compatibility Principle(PCP) which states that displays that are common to a task or mental operation be clustered close together (also known as close display proximity) (Wickens, and Carswell, 1995). Carswell (1990) found that subjects performed better with integration tasks when the design of the graphical representations were homogeneous displays were used. Although visualization has been researched in a wide variety of areas such as nuclear reactor displays, and aviation displays, presently little research based on experimentation exists in the area of work system evaluation. This research will help extend the scientific knowledge base regarding the applicability of visualization tools in work system evaluation.

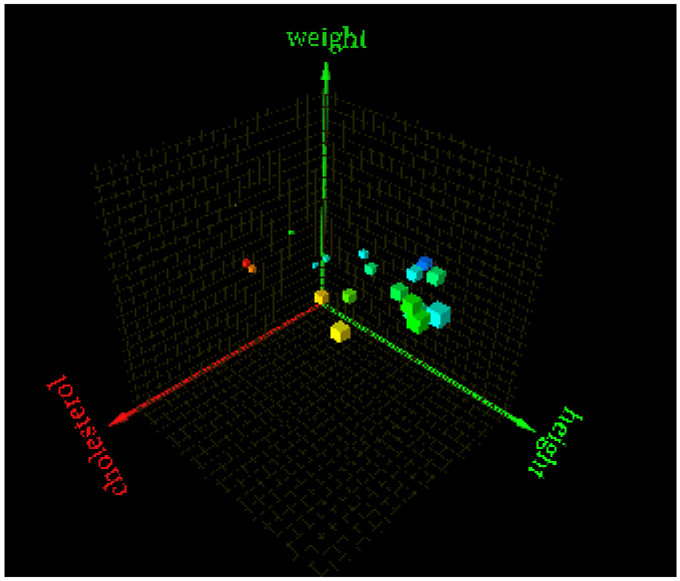


Figure 2.3: 3-D Perspective of Medical Data
(Source: SGI MineSet Page http://www.sgi.com/Products/software/MineSet/)

Spatial Ability

Given the fact that perspective introduces depth cues along the user's line of sight, a mental representation of the 3-D space is required from a 2-D screen. Several assessment tools have been developed to measure the spatial abilities of individuals, since these abilities can differ among individuals (Vandenberg, and Kuse, 1978). The Cube Comparisons test is one test that has been used to correlate spatial ability with task performance on perspective displays. Barfield and Robless(1989) have used this assessment tool to test for significance in Cube Comparisons test performance and task performance in viewing 2-D and 3-D perspective graphs. They found no correlation between the test and task performance for students, however found a significant relationship between accuracy and orientation for experienced managers.

Evaluation

In the context of work system decision making, information is processed by the decision maker(s). Once the information is processed, the decision maker(s) will need to eventually make an evaluation in order to decide whether an intervention in a work system is necessary. These evaluations can be both internal and external in nature. Internal evaluations concentrate on the relationships and states within a given work system, whereas external evaluations primarily focus on the relationships between the work system and the work system environment.

Evaluations can also be summative or formative in nature. Formative evaluations determine whether changes need to be made to a given system, and tend to concentrate on the process of that system. Summative evaluations provide information regarding whether the system should be continued at all, and concentrate on the outcome of a given system. Both of these evaluations have been in a wide variety of areas to provide information regarding programs, system design, and technology implementations. In most cases, control charts help support formative decision making since it provides information regarding the state of the work system. Whether a state is “in control” or “out of control” will dictate whether any corrective action needs to take place in the form of process or system changes.

SPC Portrayals and Decision Making

Monitoring a given work system using SPC tools, relies on some of the same abilities used in inspection such as searching for faults (recognizing out of control phenomena), classifying faults (classifying non random behavior), and taking appropriate action. Using SPC tools in a multi-process environment requires a worker to search for out of control phenomena from multiple sources of data. Although the data can be portrayed in a variety of ways, it is up to the user to recognize the proper pattern. Once the signal is recognized by the worker, they may also need to decide how to characterize the type of out of control phenomena present in the system. Following the classification procedure, the worker must take action or no action based on their classification.

Signal Detection Theory (TSD) has been used to model human decision performance (Wallack and Adams, 1969) in inspection tasks. The link between TSD and inspection lies in the fact that decisions makers must detect non-conforming items as signals from a field of conforming items (noise) (Thapa, Gramopadhye, Melloy, Grimes, 1996). The decision maker is basically a statistical hypothesis tester attempting to decide whether any item comes from a distribution of conforming items or from a distribution of non-conforming items. In making these decisions, the decision maker can also make mistakes in their assessments of the items either missing a signal, or signaling when there is only noise. According to this theory, decisions can be classified into the following categories as shown in Table 2.2:

Table 2.2: Signal Detection Theory Classification of Decision Making in Inspection

		Process Behavior	
		In Control	Out of Control
Inspector Perception	In Control	Correct Accept	Miss
	Out of Control	False Alarm	Hit

Two characteristics can be used to define decision making performance with TSD theory: sensitivity and response criterion. Sensitivity is the ability of an inspector to discriminate between conforming and non-conforming items, while response criterion is defined by the tendency of an inspector to be biased in characterizing an item as conforming or non-conforming (Thapa et. al, 1996). Previous studies have attempted to look at the impact of training strategies on these TSD parameters (Micalizzi, and Goldberg, 1989; Thapa et. al, 1996). TSD can only be applied if each decision type from Table 2.2 is committed by the observer in order to calculate the probabilities associated with TSD. These probabilities can be calculated by measuring hits and false alarms, since misses and correct accepts are complementary to these probabilities. Using these probabilities, sensitivity, also noted by d’, and the response criterion, noted by β, can be calculated using the following equations and a normal distribution table which provides z-values and ordinate values (Macmillan and Creelman, 1991):

$d' = z\,(\text{Hit Rate}) - z\,(\text{false alarm rate})$

β = ordinate of signal distribution / ordinate of non-signal distribution

Even without calculating sensitivity and response criterion, TSD still can be useful as a general model for the decision making process as traditionally applied by those interested in inspection (Drury, and Kleiner, 1984).

Training

Embrey (1979) has shown that controlled practice in visual inspection tasks improves both speed and accuracy. In conducting an experiment which partly depends on a subject's ability to recall knowledge regarding statistical quality control and to use this knowledge, training becomes an important factor in reducing subject variability. Training helps in the development of knowledge and skills which may be necessary in order to complete a task efficiently and effectively. Training can help subjects become proficient in the knowledge and skills necessary to complete the task. Training helps to close the gap between a human's capabilities and the task demands. One way to help reduce the subject variability associated with introducing tasks which may be unfamiliar to subjects is the use of training coupled with testing to measure whether a satisfactory proficiency level has been attained by the subject.

The progressive part training method is one training strategy used to improve task performance. Salvendy and Seymour (1973) successfully implemented this methodology in training people in developing industrial skills. Kleiner and Drury (1993) studied the effectiveness of this training methodology on the performance of precision metal inspectors. From their study, improvements in inspection performance were found both at the individual level and at the group level resulting in a reduced scrap rate.

A progressive part training approach would split up the domain of knowledge or skills into sections and incrementally trains the subjects to some criterion. If a task required training in four sections S1, S2, S3, and S4, the progressive part approach would be as follows (Gramopadhye, Kimbler, Kimbler, Bhagwat, and Rao, 1995; Salvendy and Seymour, 1973):

1. Train individual sections S1, S2, S3, and S4 to some criterion
2. Train S1 and S2, and S3 and S4 to some criterion
3. Train S1 and S2 and S3, and S2 and S3 and S4 to some criterion
4. Train as a whole procedure S1, S2, S3, S4 to some criterion

One of the keys to effective training design is whether the system is a closed loop system where a student receives feedback during training. This feedback, along with student trial and error make up some of the building blocks necessary for learning (Wiener, 1963). Active task training which provides task feedback has been shown to improve inspection performance. Czaja and Drury (1981) found that subjects who were given feedback with respect to search time and the correctness of answers given during training had significantly improved inspection performance. Furthermore, they found that subjects who participated in active training made fewer classification errors, search errors, and decision errors. In another study, the effects of differing feedback mechanisms during training were also studied. Specifically, feedback information and feedforward information were given to subjects in varying inspection environments. Thapa, Gramopadhye, Melloy, and Grimes (1996) found that subjects who were given feedback information improved their inspection sensitivity as well as the time to make decisions. Clearly, the role of feedback during training cannot be underestimated in any task that requires some of the same abilities as an inspection task.

Chapter 3: Methodology

This chapter will discuss how the research was conducted in an experimental setting. Based on evidence cited in the previous chapter, it appears that the use of three dimensional perspective portrayals have resulted in mixed results with respect to decision making time and accuracy as well as other task related outcomes. However, the research does appear to suggest that the more complex the data (number of dimensions or the number of graphs, etc.), the better subjects will perform on the various task related outcomes using three dimensional perspective portrayals. At the same time, portrayals also impose a mental workload as subjects attempt to perceive and process information from the tool. Little research exists in the effect of portrayal methodology or human-computer interaction on mental workload.

Overview

Subjects were given a scenario in a manufacturing setting. Subjects played the role of quality assurance engineer, in which they were required to identify problems based on portrayals of control chart data of various manufacturing processes.

Subjects

Undergraduate and graduate students from the Industrial and Systems Engineering Department were recruited for this experiment. A total of eighteen subjects were used, with six subjects used during pilot testing and twelve used for the main experiment. Students volunteered to participate in the experiment and were compensated for their participation. Compensation was based on the number of experimental phases completed. For each completed phase, subjects were given two dollars and fifty cents, totaling ten dollars for the complete experimental procedure. The experimental phases included the following: 1. Pre-experiment 2. Training 3. Experiment 4. Post-experiment. Compensation was given to subjects after the completion of the experiment. No limitations or restrictions were placed upon selection of subjects with regards to age or gender. Subjects were required to have prior Internet experience, since the experiment extensively used Netscape Navigator to collect data and portray control chart information. Only subjects with no or little SPC knowledge were allowed to participate to avoid differences due to subject expertise. A progressive part training program was administered to provide the necessary knowledge to complete the experiment and to reduce the variation attributed to differences in SPC knowledge.

Equipment

An IBM PC compatible workstation, and a Silicon Graphics O2 Workstation was used to conduct this research. Using the Silicon Graphics O2 Workstation, MineSet was used to develop the composite three dimensional perspective displays of control charts. On the IBM PC, Minitab was used to develop the multiple and composite two dimensional displays, while Adobe Photoshop was used to format all of the displays. Once the displays were formatted, they were converted into an animated Graphics Interchange Format (GIF) file in order to portray data using Netscape Navigator. The IBM PC compatible workstation was also used to collect data, as well as provide the platform for presenting training information via Microsoft PowerPoint.

Silicon Graphics O2 Workstation

A Silicon Graphics O2 workstation was used as the platform for developing composite three dimensional perspective portrayals of control chart data. This platform was necessary in order to utilize the complex features involved in visualization activities. The configuration of the workstation was the following: 180Mhz RISC chip, 96 MB RAM, 2.1 GB Hard Drive, Video Conferencing Package, and a 17" monitor.

Visualization Software

The software selected for use in the experiment is MineSet 2.0 developed by Silicon Graphics, a commercially available product. MineSet is a set of powerful data analysis and data mining tools used to analyze and portray complex sets of data. MineSet has been used in medicine, aeronautics, as well as other scientific applications to portray complex data sets in graphically representations.

IBM PC compatible workstation

A Pentium 200MHz MMX PC workstation was used for the experiment to create the animated GIF files, to provide a training platform, to provide a means of data collection, and to portray the various displays. The machine configuration included the following: 17" monitor, 3.2GB hard drive, 32 MB RAM. The display resolution was set at 1152 x 864 with a refresh rate of 70Hz.

Facilities

The experiments were conducted in the Macroergonomics and Group Decision Systems Laboratory located in 567 Whittemore Hall at the Virginia Polytechnic Institute and State University(Figure 3.1). This laboratory contains seven IBM PC Compatible workstations, one Windows NT 4.0 server, one Macintosh OS 7.5 server, an InFocus projection system, a one-way window, two whiteboards, and videotaping equipment.

Figure 3.1: Facilities in the Macroergonomics and Group Decision Systems Laboratory

For the experiment, the workstation setup used by subjects is shown in Figure 3.2:

Figure 3.2: Experimental setup used by subjects for training, data collection, and portrayal of control chart information

Experimental Design

This research used a one-way within-subjects experimental design in order to determine the effects of different portrayal methodologies on the dependent variables. In order to utilize this design, subjects were asked to identify special causes using one of the three display methodologies. With a within subjects design, fewer subjects were needed to complete all experimental conditions while at the same time increasing the statistical power of the analysis. The main independent variable represented the different display methodologies (see Table 3.1).

Table 3.1: Treatment conditions used for experimental design

Treatment Condition 1	Treatment Condition 2	Treatment Condition 3
Multiple 2-D portrayals/ w/ 8 events	**Composite 2-D portrayal w/ 8 events**	**Composite 3-D perspective portrayal w/ 8 events**

The type of non-random signal was also manipulated: 1. outside the control limits 2. run 3 trend 4. A phase related correlation. Subjects were given two examples of each type of non-random signal. As a result, subjects were required to look at a total of eight different examples or events per display. A baseline task, which consisted of two two dimensional portrayals of control chart data, was used to provide a common reference so that mental workload could be assessed. The baseline task was used as suggested by Staveland and Hart (1988) to provide subjects with a common reference task to make mental workload comparisons. This reference task was always presented before any of the treatment conditions were given to each subject(see Appendix C3).

Counterbalancing of the treatment combinations was used to reduce the possibility of ordering effects. The counterbalanced experimental design is shown in Table 3.2:

Table 3.2: Counterbalancing Scheme for Experiment

Presentation Order	First Treatment	Second Treatment	Third Treatment
A	Multiple 2-D	Composite 2-D	Composite 3-D
B	Multiple 2-D	Composite 3-D	Composite 2-D
C	Composite 2-D	Composite 3-D	Multiple 2-D
D	Composite 2-D	Multiple 2-D	Composite 3-D
E	Composite 3-D	Multiple 2-D	Composite 2-D
F	Composite 3-D	Composite 2-D	Multiple 2-D

Using twelve subjects for this experimental design resulted in two replications of the presentation order. The specific out of control events was presented randomly for any given portrayal. These out of control events were randomly picked from a pool of data sets(see Appendix G1) which contained the required non-random signal. Four data sets for each non-random signal was produced, totaling sixteen sets of non-random data. Each data set contained fifty points with a non-random signal inserted at a random time. The time at which the completed signal appeared was uniformly distributed between a range of 20 points and 45 points and was randomly assigned to each treatment for each subject.

For the trend, the incline or decline of the best fit line of the signal was kept constant at 20 degrees. For the run, the mean was shifted by ± 1.0 from the normalized mean. These constant levels set were used since pretesting subjects had some difficulty recognizing smaller angles for the trend and smaller shifts for the run. For the phase related patterns, the cross correlation between each serial process set was kept constant at .85. This number was attained by adding a constant to a sinusoidal function for five points, over which have the sine wave was completed. This was necessary since pretesting subjects had difficulty with lower cross correlation values. The size of the display, colors used among displays were also kept constant. These variables were controlled in order to provide precision and reduce confounding.

Task

Subjects completed a task set in a manufacturing environment. Subjects were assigned the role of quality assurance engineer (see Appendix B1) and this role remained constant throughout the experiment and across subjects. As quality assurance engineers, subjects were instructed to search for and identify any non-random signal in three processes using one of the three display methodologies. In order to maintain decision consistency among subjects a formalized set of rules was used to define the non-random signals. The definitions used for this experiment are as follows:

1. Outside the control limits: a data point which lies above the upper control limit or below the lower control limit(Hoyer, and Ellis, 1996)
2. Runs: Seven consecutive points above the centerline or seven consecutive points below the centerline(Hoyer, and Ellis, 1996)
3. Trends: Six consecutive points are in a monotone increasing or decreasing pattern(Hoyer, and Ellis, 1996)

4. Phase related data: Five consecutive points which exhibit similar behavior or inverse behavior (magnitude and direction) in two control charts based on some lag difference

Subjects were given control chart data, from a variety of machining operations as well as a legend to relate the data to the appropriate machining process. Subjects were given additional background information associated with incorrectly classifying the process, as well as the associated savings with the correct classification of out of control signals as shown in Appendix B2-B4. This provided explicit instructions for the subjects in order to maintain some control of the response criterion of subjects.

Using an animated GIF, data was portrayed dynamically to the subject for all portrayals. A page created in hypertext markup language (html) contained the animated GIF, a color coded legend, and a button to stop the animated GIF from providing additional data. An example of the multiple 2-D, composite 2-D, and the composite 3-D perspective can be seen in Figures 3.3, 3.4, and 3.5:

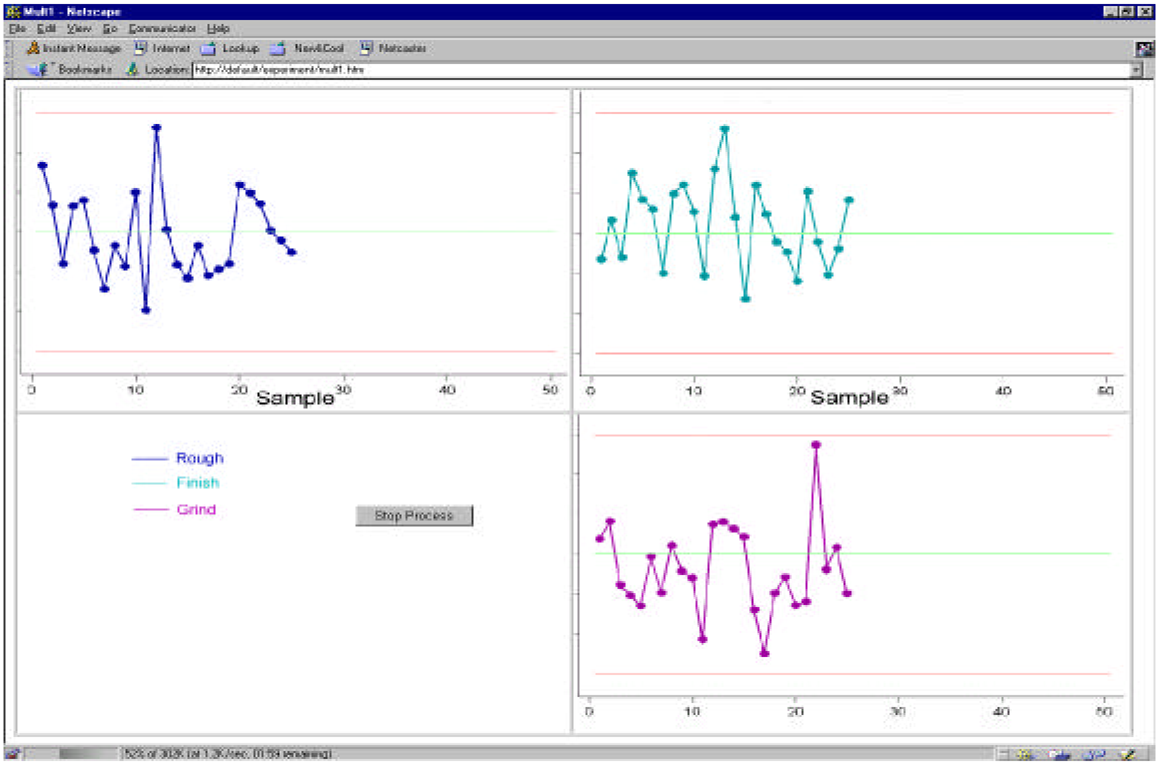


Figure 3.3: Sample Screen used for Multiple 2-D Treatment Level

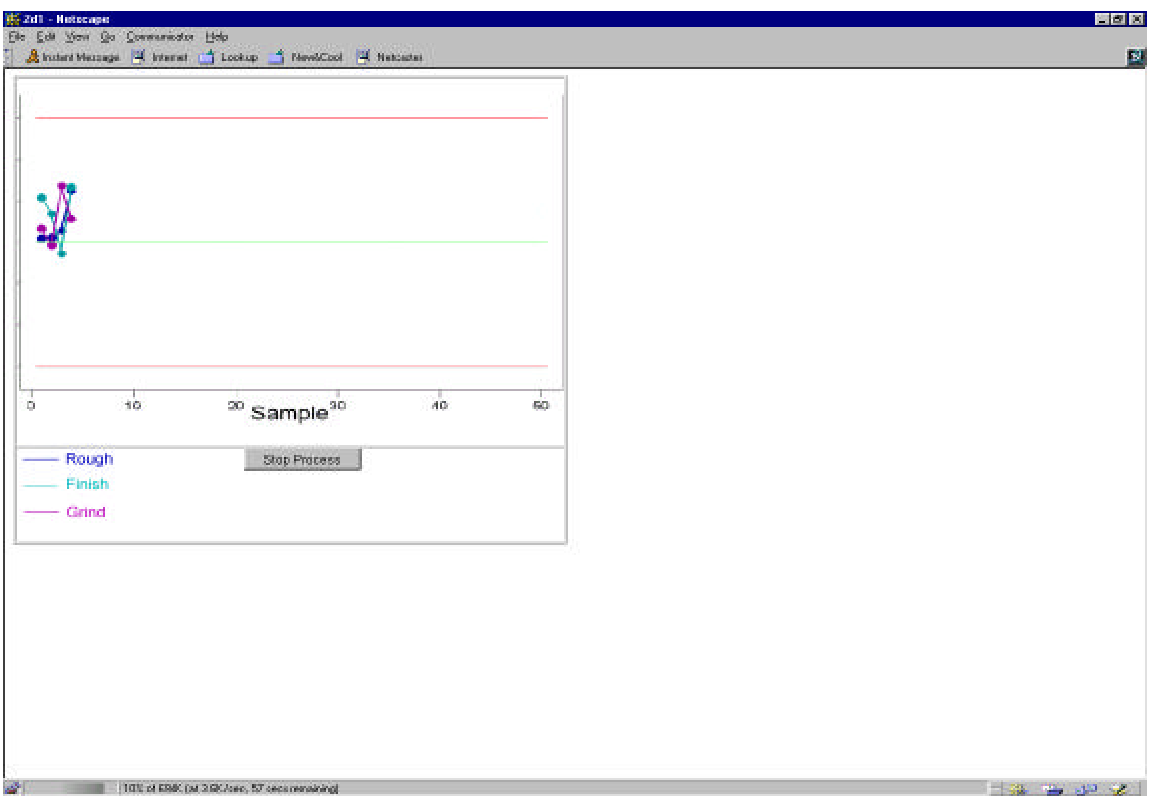


Figure 3.4: Sample Screen used for Composite 2-D Treatment Level

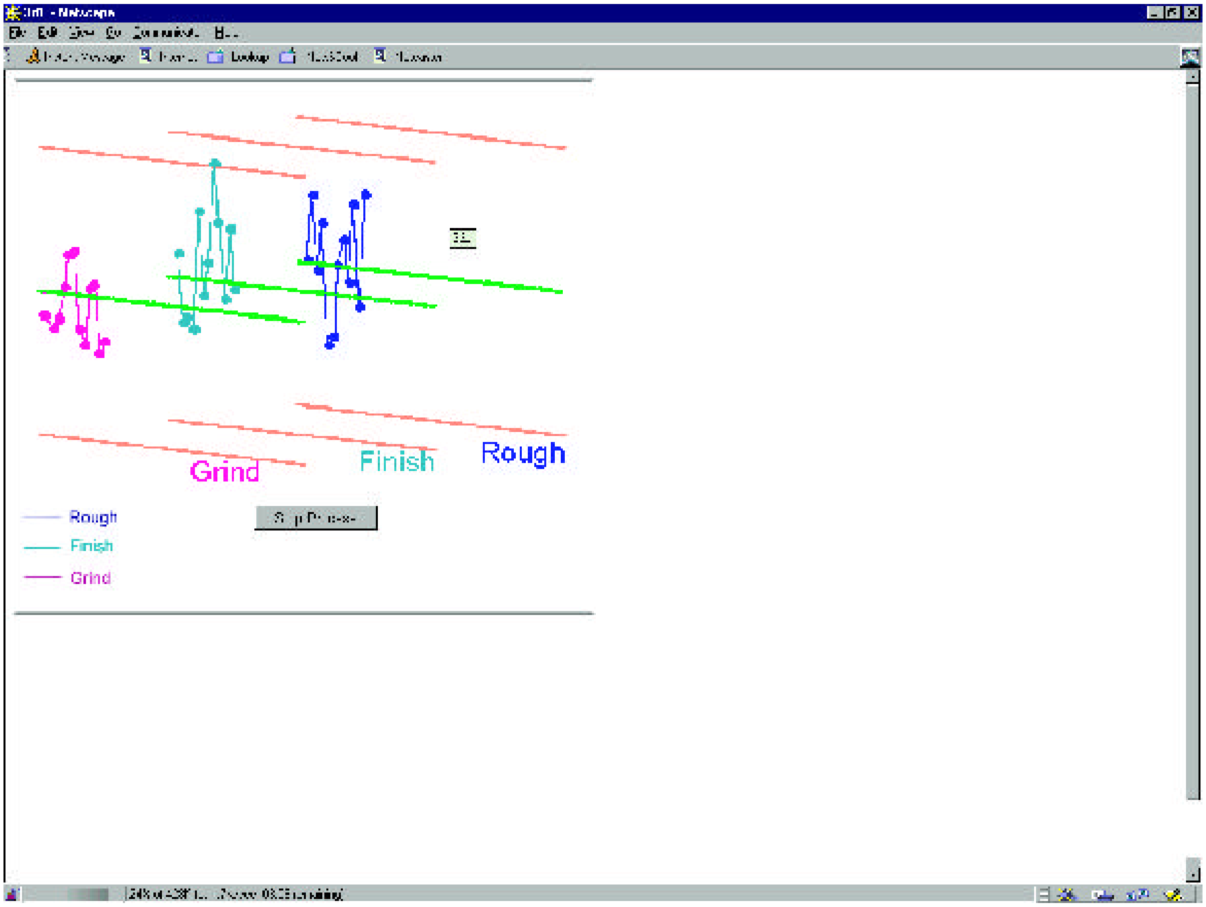


Figure 3.5: Sample Screen used for Composite 3-D Treatment Level

The animated GIF was developed using MineSet and Minitab. Manipulations in color and display size were made using Adobe Photoshop in order to keep color and display size for each graph constant among treatment combinations. Microsoft GIF Animator was used to create the animated GIF files necessary to portray data dynamically

Once a signal was recognized by the subject from the animated GIF, they were instructed to press a key on the Internet page. Although certain processes may be more likely to have different reliabilities, a more fundamental issue is the distribution associated with each manufacturing process. For this research, it was assumed that the reliabilities of the processes are constant, while the distribution of the processes was normal. Following the recognition of the signal, the portrayal was no longer displayed, and a follow up screen asked the subject to report the out of control incident (see Appendix B5). After the report was logged by the subject, another event was portrayed to the subject until all events were completed for that specific portrayal. Following all events, the NASA TLX form was given to the user to rate the specific dimensions of mental workload (see Appendix C2) for the specific portrayal. For the baseline task, they were also required to weigh the dimensions of workload (see Appendix C1), and these weights were used for the remaining portrayals. This procedure has been suggested by Hart and Staveland (1988), as long as the baseline task is similar to the remaining tasks. Subjects were asked specific information regarding the difficulty of the portrayal, and any problems associated with identifying signals with the portrayal. Data throughout the experiment was automatically collected using html form pages.

Procedure

Subjects arrived at the Macroergonomics and Group Decision Systems Laboratory, located in 567 Whittemore Hall. Subjects were asked to read and sign an informed consent form (see Appendix D1). After completing the form, subjects filled out a pre-experiment questionnaire (see Appendix E1) made up of questions relating to computer experience, statistical process control knowledge, computer visualization experience, and academic standing. This questionnaire was administered using an Internet form page. Following the completion of the questionnaire, subjects were asked to complete the Cube Comparisons test using pencil and paper. Following this test, subjects participated in the progressive part training procedure which dealt with statistical process control knowledge, using Microsoft PowerPoint slides (see Appendix F1). The training procedure consisted of ten modules. Following the completion of each module, a five question evaluation test (see Appendix F2) was administered to test their proficiency level for that specific training module. Subjects who did not meet a criterion of 80% for any evaluation test were asked to repeat the module dealing with that subject matter. The training procedure was self-paced.

Following the completion of the training questionnaires, subjects looked at a pair of two dimensional control charts containing control chart data on a computer screen. They were asked to evaluate whether an “out of control” signal exists on either graph and to identify the signal. A sample page of the possible alternatives were printed on paper for the subject’s convenience (see Appendix B5). This pre-experiment task was the same for all subjects to provide baseline measurements for mental workload. Subjects were asked to fill out a questionnaire regarding the workload variation as applied to reading the pair of two dimensional control charts. This questionnaire was based upon the NASA Task Load Index developed to assess mental workload. In this questionnaire, they were asked to weigh various dimensions of workload as defined by this workload assessment index (see Appendix C1). Definitions of each workload term were provided at the bottom of this questionnaire, and were also provided as a separate handout available to the subject.

After completing this questionnaire, subjects were given a set of instructions on the computer which provided background information for a work system decision making task (see Appendix B1). Depending on the display presentation order for the subject, subjects read one of three information sheets on the computer corresponding to a specific display format (see Appendix B2, B3, B4). These information sheets were nearly identical except for the manufacturing site and the product produced at the site. After reading the information, the experimenter answered any questions that the subject had before starting the task. Subjects were given an experimental treatment block corresponding to one of the three display formats (multiple two dimensional control charts, composite two dimensional control chart, and composite three dimensional perspective control chart) in which they must indicate the presence of "out of control" patterns on the display and identify which pattern was recognized. The sample page which provided the possible out of control alternatives was again provided to subjects as in the baseline task. Decision time and the answer for each trial was recorded automatically on the computer. Following the presentation of eight "out of control" events, subjects were asked to complete a post-treatment questionnaire to assess subject decision confidence and difficulty associated with the decision making task (see Appendix B6). A separate questionnaire was also used to provide subjective ratings from the NASA Task Load Index to assess mental workload (see Appendix C2). Definitions for the workload terms were provided as a separate handout and also provided at the bottom of the rating form. The answers to these questionnaires were collected using the computer. Then subjects were given an optional two minute break before the next treatment block.

Following the completion of the first treatment block, the subject viewed one of the two remaining displays using the exact same procedure mentioned above and the appropriate information sheet. Following the second treatment, the remaining treatment block was administered using the same procedure. After all three treatment blocks were completed, a post experiment questionnaire was given on the computer (see Appendix E2). This questionnaire was used to assess subject preferences regarding the display and any recommendations they would have regarding work system design. Once this questionnaire was completed, subjects were paid for their participation and dismissed from the experiment.

Training

Subjects were trained in statistical quality control, specifically the fundamentals of control charts, and the recognition of specific control chart patterns. Using subjects with little or no prior knowledge of SPC, training to criterion was used to minimize differences associated with subjective expertise in SPC. A PowerPoint presentation was used as the primary source of training. Based on supporting literature, the progressive part training method was used to train subjects in SPC. Specifically, the training program was focused around these areas: 1. components of control charts and random variation (S1) 2. outside control limits and trends (S2) 3. runs (S3) 4. phase related correlation. (S4)

1. Train individual sections S1, S2, S3, and S4 to an 80% criterion level
2. Train S1 and S2, and S3 and S4 to an 80% criterion level
3. Train S1 and S2 and S3, and S2 and S3 and S4 to an 80% criterion level
4. Train as a whole procedure S1, S2, S3, S4 to an 80% criterion level

PowerPoint slides were developed to train the subjects and are provided in Appendix F1. Five question tests were provided to subjects following each module of the training procedure and are provided in Appendix F2. The 80% criterion level was reached if four of five questions were answered correctly. In this case subjects were allowed to proceed to the next module. If this 80% criterion level was not reached, subjects were asked to repeat the last training module completed by the subject and to re-take the same five question test.

Data Collection

Subject related data, demographic data, decisional data, spatial ability data, and subjective data was collected throughout the experiment:

Subject related data

Once subjects enrolled to participate in this experiment, data was collected in order to facilitate the execution of the experiment in a controlled manner. Contact information such as name, phone number, email address, times of availability, and date of experiment was collected. After this data was collected, subjects were assigned a subject code. Throughout the experimental data collection procedures, this code was utilized to identify the individual in order to insure anonymity.

Demographic data

In order to be able to be to assess the general knowledge level of subjects, a pre-experiment questionnaire was administered. This questionnaire asked questions dealing with the level of prior experience with quality control, academic level, general Internet experience, and visualization experience. (see Appendix E1)

Decisional data

Search and decision time was collected via an on-line experimental data collection form used by the experimenter. Search and decision time which measures the combined time required to search for a specific signal and characterize that signal was measured for each specific event. These times were only counted when an accurate decision was made. Stopping time was computed by the total time required to complete the task for each display..

Decision accuracy (DA) was measured as a raw score which indicates the number of events correctly classified by the subject. Specifically, each event was counted as one opportunity to answer a question. This measure attempted to identify whether one portrayal has a distinct advantage over another in terms of decision accuracy.

Decision confidence (DC) was assessed by incorporating a question into a post treatment questionnaire. This question was used to provide subjective feedback on how comfortable the subject was with any particular portrayal.

Sensitivity and response criterion, both of which come from Signal Detection Theory, were also of interest in this research. In order to compute these measures, the probability of a subject indicating a signal when a signal is present (probability of a hit), and the probability of a subject indicating noise a signal when there is noise (probability of a false alarm) had to be computed. From these probabilities, the sensitivity and response criterion could be computed using the equations listed in the previous chapter. In order to calculate these probabilities, each data point was treated as a separate opportunity to indicate an answer. When subjects allowed the portrayal to continue to proceed without stopping the process, this indicated that subjects only saw noise (processes is in control). When subjects stopped the process to indicate a signal (one or more of the processes indicates non-random behavior), this indicated that subjects saw a signal. By knowing when the subject stopped the process and by knowing when the signal actually appeared, both of these probabilities can be computed for each display.

Mental workload

As indicated in the previous chapter, the NASA Task Load Index was used to measure the subject's assessment of workload. After the subject was trained, a baseline task was given using a pair of separate 2-D portrayal of control chart data. This step allowed the experimenter to collect data about the weightings of the components of workload. This provided a common reference task(see Appendix C3) which subjects can relate to in assessing the workload of the other portrayal conditions. As mentioned earlier, the baseline task required subjects to identify out of control signals using a pair of separate 2-D portrayals and to indicate the appropriate signal via an online report form.

The basic assumption is that the baseline task is somewhat related to the tasks used in the actual experiment. For this experiment, the task remained relatively constant, however the portrayal was changed between treatment combinations. After each treatment combination, the subject was asked to complete the NASA TLX for that particular treatment with respect to the baseline task. This approach is slightly different that the task originally used by Hart and Staveland (1988). However, they suggest this approach may also be useful in reducing between subject variability, since all the subjects have a common reference task to compare their experiences with.

Spatial Ability Data

Viewing 3-D perspective data on a 2-D monitor required subjects to perform mental tasks in order to perceive the data correctly. However, all individuals do not have the same spatial abilities to perform the necessary mental tasks to view the 3-D perspective data. Several tests exist to assess the spatial ability of subjects. For the purpose of this research, the Cube Comparisons test was used to assess the spatial ability of subjects. This test was also used by Barfield and Robless (1989) in a task which also used two dimensional and three dimensional perspective graphs.

Subjective Data

In order to allow the subject to provide some qualitative data which justified their task performance, a questionnaire was provided to each subject following the completion of all events for a specific portrayal. Subjects were given an opportunity to judge the effectiveness of all the portrayals, provide data dealing with difficulties associated with identifying signals and provide a rating for decision confidence, etc. This qualitative data helped to support the quantitative findings of the research. Although verbal protocol could have been used to provide similar data, the interference of thinking aloud with the task was a major concern, especially when dealing with mental workload assessment. Thus, the retrospective report was selected based on its lack of interference with the main task of interest.

MineSet 2.0

MineSet developed by Silicon Graphics Inc. is a data mining tool that has visualization capabilities to explore complex sets of data. MineSet offers a variety of customizable tools to portray data in a variety of multi-dimensional formats. For the purpose of this research, only basic tools were required to model the control chart data of various manufacturing processes. Although MineSet allows subjects to manipulate the perspective of the data, this feature was not used for this research. The perspective of the 3-D perspective portrayals remained fixed throughout the portrayal so that data could be portrayed dynamically through an animated GIF file. MineSet also has the capabilities to display data as decision trees, charts, graphs, scattergrams, etc. depending on the particular needs of the user. An graphical example of the decision tree function is provided in Figure 3.6:

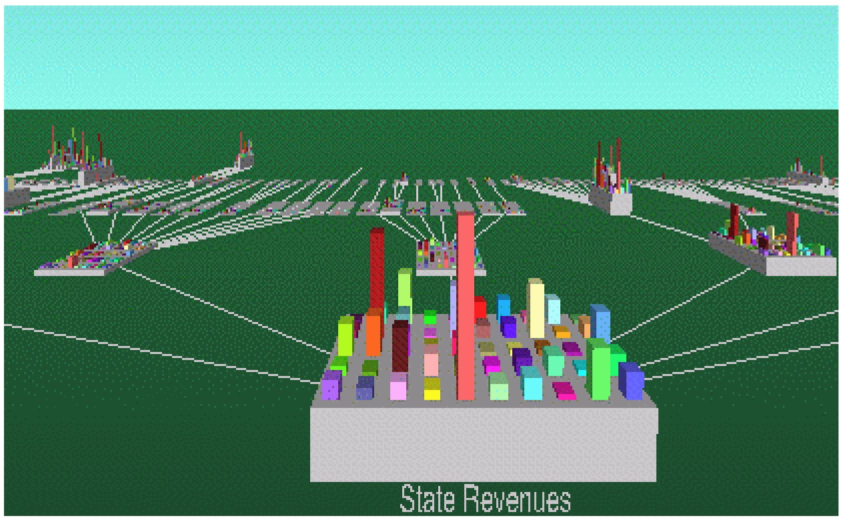


Figure 3.6: Sample Data of Tree Visualizer Tool
(from http://www.sgi.com/Products/software/MineSet/products/)

MineSet has other capabilities that could easily be used in an industrial setting. MineSet has the capabilities to gather information on-line from other sources in order to portray information in real time, using more complex networking hardware. For this research, data from a flat file was used to portray information about the manufacturing system.

Experimental Plan

The following list represented the plan which was followed to execute the experiment:

1. IRB Approval
2. Experiment development
3. Pre-testing

4. Solicit participants
5. Individual Data collection:
   a) Informed consent completion
   b) Pre-experimental questionnaire completion (demographic data collection)
   c) Cube Comparisons Test
   d) Progressive Part Training Program on Statistical Process Control via Microsoft PowerPoint
   e) Baseline treatment
   f) NASA TLX data
   g) Treatment combination 1
   h) NASA TLX data and supplemental data collection
   i) Treatment combination 2
   j) NASA TLX data and supplemental data collection
   k) Treatment combination 3
   l) NASA TLX data and supplemental data collection
   m) Post-experiment questionnaire
6. Data Analysis

Chapter 4: Results

Subjects' Demographics

Demographic information was collected via an Internet form page before the experimental session began. A questionnaire was used to collect this demographic information to be used for data analysis. This questionnaire assessed the level of education, experience with the Internet, experience with visualization applications, and experience with statistical process control (SPC) or quality control (QC). The pre-experimental questionnaire used to collect this data for this research effort is shown in Appendix E1.

Academic Status

Data pertaining to the educational level of the participants was collected through a question on the pre-experiment questionnaire. Figure 4.1 shows a histogram which represents the educational level of the 12 subjects used for the experiment. The mean educational level of the subjects was 2.00 with a standard deviation of 1.28, where 1 = sophomore, 2 = junior, and 3 = senior. It is important to note that seniors could not be recruited for the experiment. The lack of seniors in the sample population was due to the fact that seniors typically take a course in statistical process/quality control in their senior year. Masters students were also difficult to get since many had an undergraduate education which consisted of a class in quality control or statistical process control. At the same time the research was conducted, many masters students were also in the process of learning statistical process control in one of their graduate courses. Students with some prior SPC educational were not permitted to participate in the experiment since the progressive part training scheme was used to provide the necessary knowledge required by the experiment.

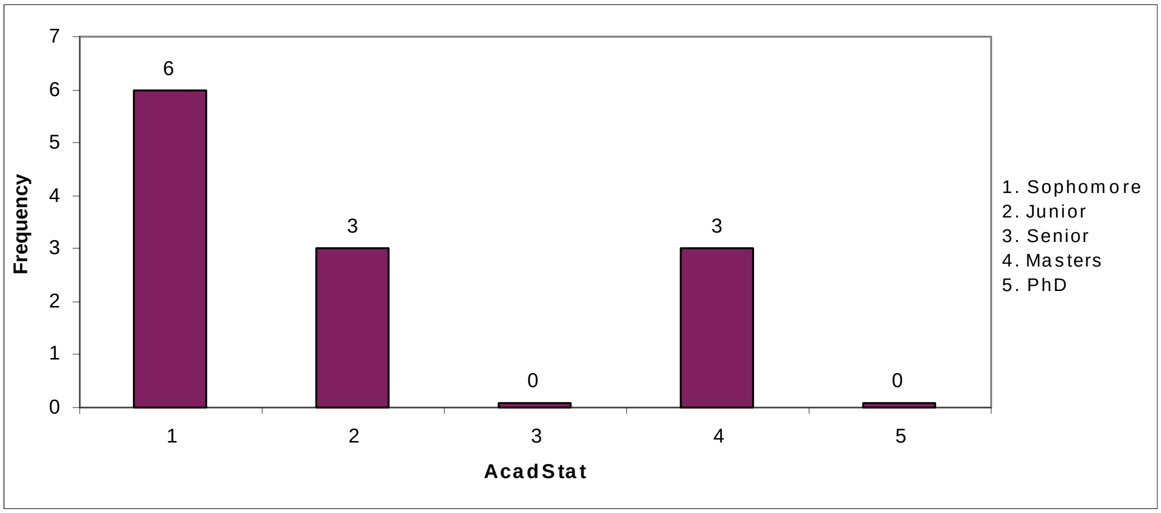


Figure 4.1: Histogram of educational level of experimental subjects

Statistical Process Control or Quality Control Knowledge

Prior knowledge in statistical process control (SPC) or quality control (QC) was assessed using a question in the pre-experimental questionnaire to insure that all subjects had little or no prior knowledge of these topics. In Figure 4.2, a histogram of prior SPC or QC knowledge is shown. Based on the data collected, the average prior knowledge in SPC or QC was 1.667 (n=12) with a standard deviation of .492, where 1 = no SPC/QC experience, and 2 = little SPC/QC experience.

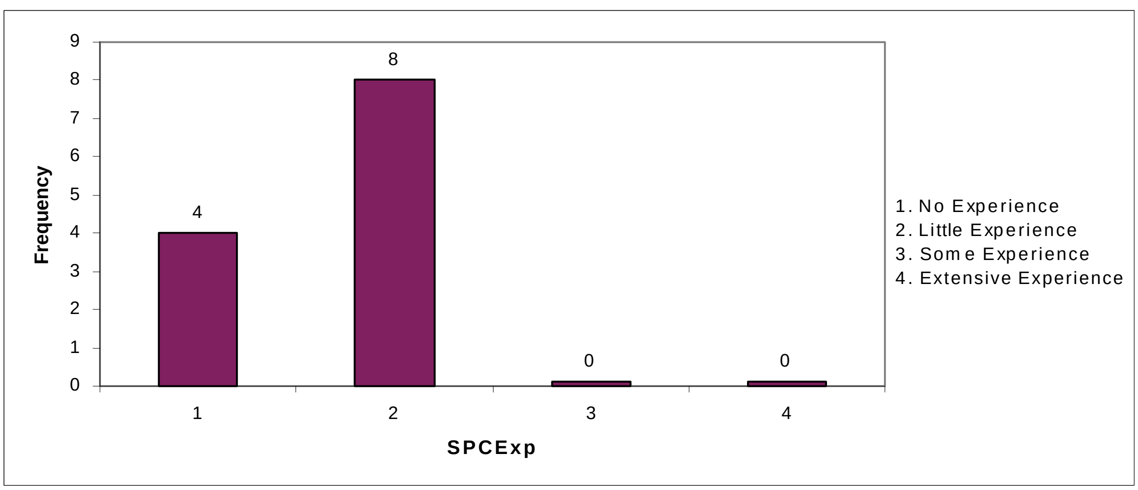


Figure 4.2: Histogram for Statistical Process Control or Quality Control Knowledge

Visualization Experience

Visualization experience of subjects was collected to understand how experience with applications which utilize 3-D cognitive abilities may influence decisional time and accuracy relating to the 3-D perspective portrayals of control chart data. The mean level of visualization experience was 2.167 (n=12) with a standard deviation of .835, where 2 = little visualization experience, and 3 = some visualization experience. A histogram of the visualization experience for the experimental subjects is shown in Figure 4.3.

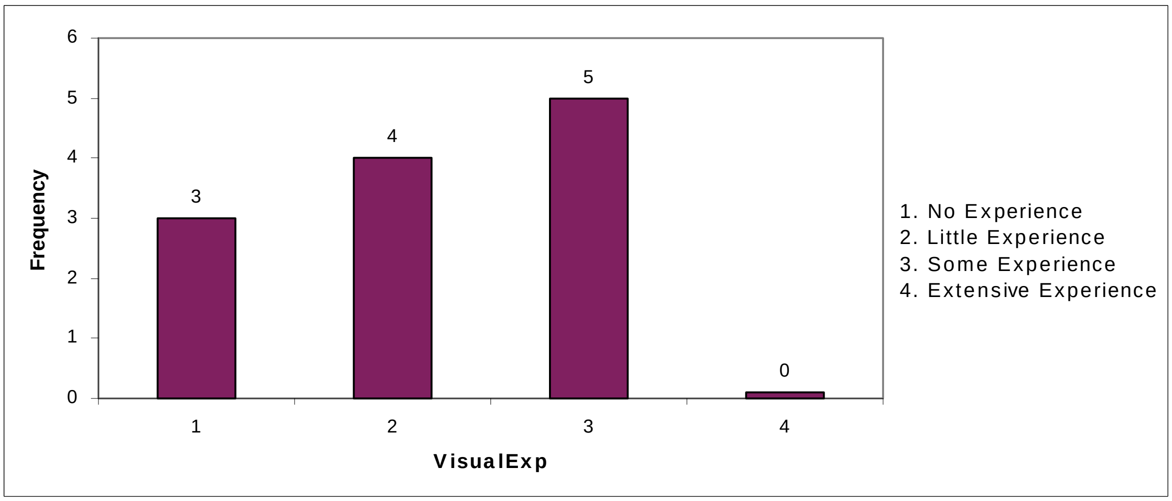


Figure 4.3: Histogram of visualization experience of experimental subjects

Internet Experience

Experience with the Internet was also collected since the experimental task depended upon the subjects ability to utilize the Internet browser to collect experimental data. This data was collected to understand how this experience may impact decision time of subjects. The average experience with visualization applications was 3.417 with a standard deviation of 0.649, where 3 = some Internet experience and 4 = extensive Internet Experience. Figure 4.4 illustrates a histogram for Internet experience for the 12 subjects who participated in this study.

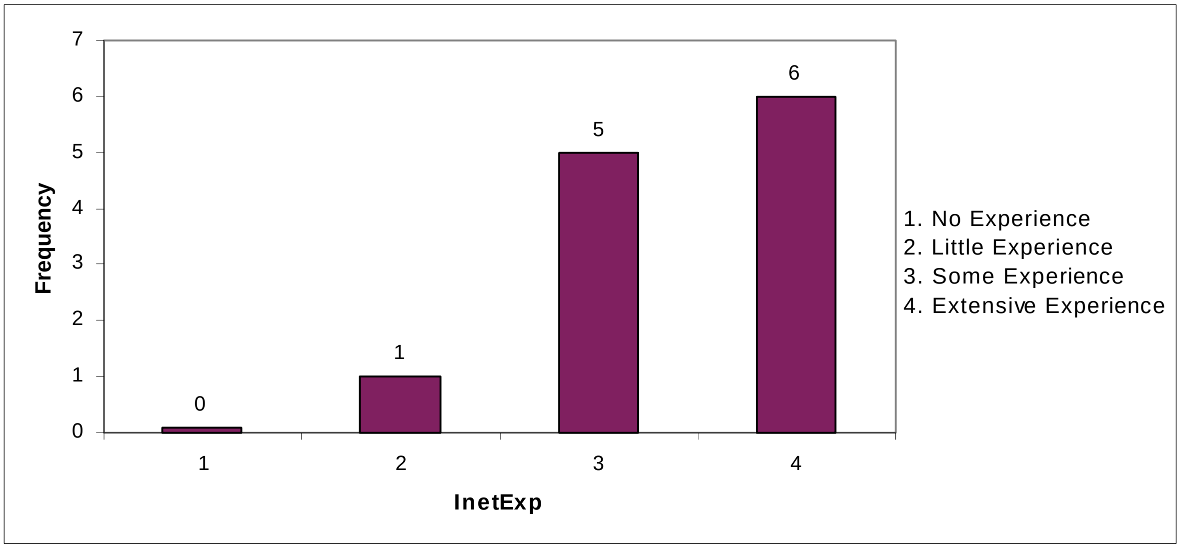


Figure 4.4: Histogram of Internet experience of experimental subjects

Tests for Ordering Effects

Since this research effort used a repeated measures design, the effect of learning was potentially an issue in experimentation. The treatment levels used for the independent variable were completely counterbalanced for this experiment to counteract the potential effects of learning. A baseline task was always given before any of the counterbalanced treatment levels were administered to obtain NASA TLX workload data. To insure the counterbalancing measures were successful, a one-way analysis of variance was conducted on the order for the dependent variables. No significant ordering effects were found for all the dependent variables in this study. Table 4.1 shows the results from this analysis.

Table 4.1: Results of ANOVA to test the effects of order on the dependent variables

Dependent Variable	F-Value	p-value	Significance
Mental Demand	$F_{(2,20)}=1.40$	0.32	n.s.
Physical Demand	$F_{(2,20)}=0.48$	0.64	n.s.
Temporal Demand	$F_{(2,20)}=1.80$	0.25	n.s.
Observed Performance	$F_{(2,20)}=0.76$	0.51	n.s.
Frustration	$F_{(2,20)}=0.15$	0.87	n.s.
Effort	$F_{(2,20)}=0.40$	0.69	n.s.
Total Workload	$F_{(2,20)}=0.84$	0.48	n.s.
Search Time	$F_{(2,20)}=0.28$	0.77	n.s.
Decision Time	$F_{(2,20)}=1.00$	0.42	n.s.
Search + Decision Time	$F_{(2,20)}=0.33$	0.73	n.s.
Accuracy	$F_{(2,20)}=1.11$	0.39	n.s.
Decision Confidence	$F_{(2,20)}=0.52$	0.62	n.s.
Decision Quickness	$F_{(2,20)}=1.50$	0.30	n.s.
Decision Accuracy	$F_{(2,20)}=0.91$	0.45	n.s.
Sensitivity	$F_{(2,20)}=2.28$	0.18	n.s.
Response Criterion	$F_{(2,20)}=0.96$	0.44	n.s.

Tests for Experimental Hypotheses

Decision Making Performance

The results in the following section represent characteristics of decision making involved with using the various displays.

Decision Accuracy

Decision accuracy was measured by the percentage of events correctly identified within five points after a signal was introduced. Eight events were introduced per subject per display. For all tests of significance for this research, the α level was set at 0.05. The Analysis of Variance (ANOVA) indicated a significant main effect of display, rejecting the null hypothesis that the means of decision accuracy across the displays are equal ($F_{(2,22)}$=8.54, p=0.002). From the Least Significant Difference test, differences were found among the displays as shown in Figure 4.5. Multiple 2-D displays were found to have significantly higher decision accuracies than the composite 2-D display. Multiple 2-D displays were found to have significantly higher decision accuracy than the composite 3-D displays.

The decision accuracy for the multiple 2-D displays, composite 2-D display, and composite 3-D display was 0.68, 0.51, and 0.30 respectively(where 1.0 equals 100% correctly answered) with standard deviations of 0.10, 0.15, and 0.20 respectively. The average decision accuracy across all subjects was 0.55 with a standard deviation of 0.20. Figure 4.5 shows the average accuracy scores for each display. Means which are not significantly different from one another (as tested using the Least Significant Difference test) are indicated by a line over the means in Figure 4.5. This convention will be used throughout the remainder of the chapter.

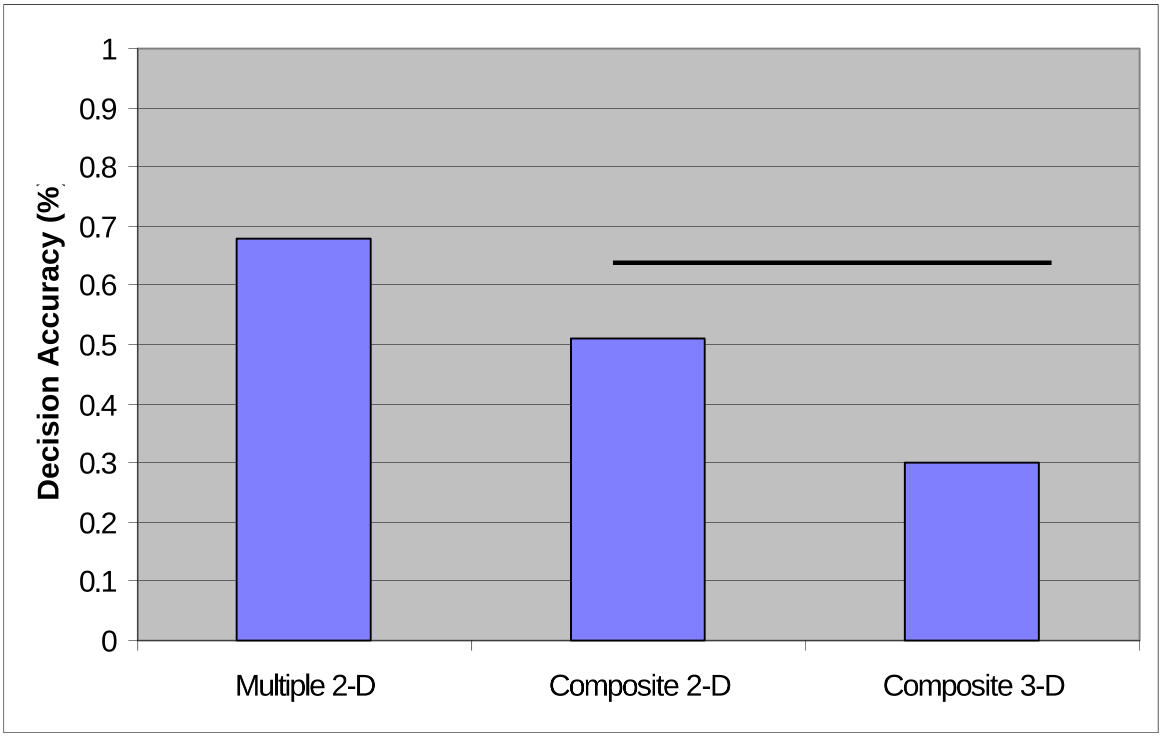


Figure 4.5: Average accuracy scores for each display

An analysis of the type of signal and its effect on accuracy was also performed. The ANOVA indicated a significant main effect of signal type, rejecting the null hypothesis that the means of decision accuracy across the signal types are equal ($F_{(3,33)}$=17.79, p<0.001). Subjects were given a total of six events for each type of non-random signal. A Least Significant Difference test was conducted to identify all possible significant comparisons. The results of this analysis indicated the following significant differences between the following comparisons: outside the control limits better than runs, outside the control limits better than trends, outside the control limits better than phase related signals, runs better than phase related signals, and trends better than phase related signals. Figure 4.6 represents the average decision accuracy for each non-random signal.

The average decision accuracy for the outside the control limits, phase related signals, runs, and trends are 0.77, 0.21, 0.58, and 0.62 (1.0 meaning 100% of signals were correctly identified) respectively with standard deviations of 0.21, 0.20, 0.23, 0.22 respectively. The average decision accuracy across subjects was 55% with a standard deviation of 20%.

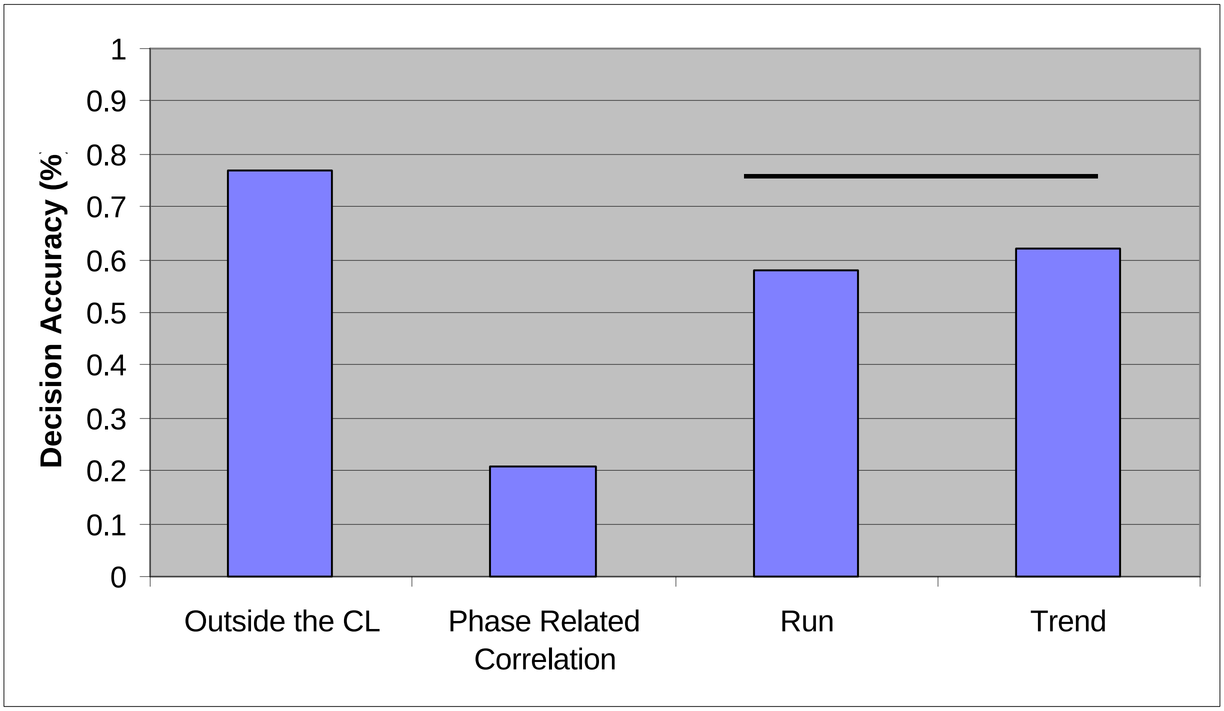


Figure 4.6: Histogram of subject accuracy scores by signal type

Search Time

For this analysis, search time was calculated by subtracting the time at which an event appeared from the time when subjects correctly stopped the process to characterize the signal. This accounted for subjects prematurely stopping the process without the correct answer, while also accounting for the presentation of a non-random signal at different times. These values were averaged for each subject to get an average search time. The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of search time across the displays are equal ($F_{(2,22)}=0.13$, $p=0.877$). The average search time for the multiple 2-D, composite 2-D, and composite 3-D displays were 7.52, 6.84, and 8.21 seconds respectively with standard deviations of 2.53, 2.61, and 3.57 respectively. The average search time across subjects was 7.54 seconds with a standard deviation of 2.91.

Decision Time

Decision time was calculated by the time required to submit a correct decision after a subject stopped the process to characterize the signal. The decision times for each subject were averaged to get an average decision time. The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of decision time across the displays are equal ($F_{(2,22)}$=0.64, p=0.536). The average decision time for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 4.54, 4.48, and 4.25 seconds respectively with standard deviations of .84, .88, and 1.00 respectively. The average decision time across subjects was 4.42 seconds with a standard deviation of 0.89.

Search and Decision Time

Search + Decision time was calculated by the sum of the decision time and the search time. Search + Decision time was used to measure the time involved in the decision making process necessary to recognize each non-random signal. An average of these times were taken for each subject. The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of search + decision time across the displays are equal ($F_{(2,22)}$=0.33, p=0.724). Based on the ANOVA results, the type of display did not significant affect search + decision time. The average search + decision time for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 12.02, 11.31, and 12.46 seconds respectively with standard deviations of 3.12, 2.81, and 3.51 respectively. The average search + decision time across subjects was 11.93 seconds with a standard deviation of 3.10.

Stopping Time

Stopping time was measured as the sum of all the search + decision times for any single display. Since signals were randomly distributed throughout the data sets, individual comparisons could not be made within subjects. Rather, the mean stopping time for a display was used to test whether any differences existed between the displays to account for the random distribution of signals. In order to do this type of analysis, a two-sample t-test was used to test any significant differences in stopping time. Using this test, the type of display did not have a significant effect on the mean stopping time of the displays. Mean stopping times for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 23.62, 24.86, and 21.87 minutes respectively, with standard deviations of 3.52, 3.01, and 5.61 respectively. The average stopping time across subjects was 23.45 minutes with a standard deviation of 4.27.

Signal Detection Measures of Performance

Using Signal Detection Theory to provide the metrics to measure decisional characteristics, the probability of correctly identifying a signal when a signal was present (probability of a hit), and the probability identifying a signal when only random variation exists (probability of a false alarm) were calculated. The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of the hit rate and false alarm rate across the displays are equal (Hit Rate: $F_{(2,22)}$=0.74, p=0.487, False Alarm Rate: $F_{(2,22)}$=1.63, p=0.219). The average hit rate for the multiple 2-D, composite 2-D, and composite 3-D displays were p=0.51, p=0.57, and p=0.51 respectively with standard deviations of 0.12, 0.21, and 0.19 respectively. The average false alarm rate for the multiple 2-D, composite 2-D, and composite 3-D displays were p=0.13, p=0.17, and p=0.16 respectively with standard deviations of 0.06, 0.07, and 0.08 respectively.

Sensitivity and response criterion were also computed using the hit rates and the false alarm rates. The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of sensitivity and response criterion across the displays are equal (d': $F_{(2,22)}$=0.21, p=0.812, β: $F_{(2,22)}$=1.82, p=0.185). The average sensitivity values for the multiple 2-D, composite 2-D, and composite 3-D displays were d'=1.58, d'=1.33, and d'=1.39 respectively with standard deviations of 0.89, 1.11, and 1.12 respectively. The average response criterion values for the multiple 2-D, composite 2-D, and composite 3-D displays were β=2.24, β=1.48, and β=1.78 respectively with standard deviations of 1.09, 0.74, and 0.86 respectively.

Receiver operating characteristic (ROC) curves are typically used in TSD to provide a description of the observers response biases. However these ROC curves could not be computed for this research since the hit probability was not a monotonically increasing function of the false-alarm rate, nor was the slope of this function monotonically decreasing (Green and Swets, 1966). Typically these ROC curves involve manipulating the response criterion to measure the effects on hit rate and false alarm rate.

Summary of Results of Decisional Characteristics

Table 4.3 summarizes the results relating to decisional characteristics measured in this research. The individual ANOVA tables can be found in Appendix A1.

Table 4.2: Summary of results of decisional characteristics

Decision Variable	F-Value (or T-Value)	p-value	Significance
Decision Accuracy	$F_{(2,22)} = 8.54$	0.002*	significant
Search Time	$F_{(2,22)} = 0.13$	0.877	n.s.
Decision Time	$F_{(2,22)} = 0.64$	0.536	n.s.
Search + Decision Time	$F_{(2,22)} = 0.33$	0.724	n.s.
Stopping Time	Mult – 2D: T = -0.93 Mult- 3D: T = 0.92 2D – 3D: T = 1.63	Mult – 2D: 0.36 Mult- 3D: 0.37 2D – 3D: 0.12	n.s.
Probability of Hit	$F_{(2,22)} = 0.74$	0.487	n.s.
Probability of False Alarm	$F_{(2,22)} = 1.63$	0.219	n.s.
Sensitivity (d')	$F_{(2,22)} = 0.21$	0.812	n.s.
Response Criterion (c)	$F_{(2,22)} = 1.82$	0.185	n.s.

* significant at p <.01

Mental Workload

Workload was computed using the NASA Task Load Index (TLX) test. This test is comprised of the following dimensions: mental demand, physical demand, temporal demand, observed performance, frustration and effort. Subjects provided weights for each of these dimensions by making pairwise comparisons between all of these workload dimensions using a baseline task as a reference. After all of these pairwise comparisons are made, the weights were obtained by taking a frequency count of the selection made by subject for each dimension. Following this procedure, subjects were given one of the three treatment levels of the research. After the experimental treatment was completed by the subjects, they were asked to assess each specific dimension using a scale of zero to twenty. The rating score for each dimension was computed by multiplying the weighting obtained from the baseline task and the rating for the specific dimension for the specific display. A rating score was obtained for each of the displays used in this experimental study. The total (mental) workload for each display was computed by the summation of all the rating scores of the workload dimensions multiplied by the appropriate weightings. Lower rating scores are better than higher rating scores since they indicate less total (mental) workload when assessing total (mental) workload. Similarly, the lower rating scores for the specific dimensions are better than higher rating scores since the contribution to the total (mental) workload is less for that specific workload dimension. All of the rating scores have no units since they are based on subjective scales which have no units.

Total (Mental) Workload

The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of mental workload ratings across the displays are equal ($F_{(2,22)}$=9.42, p=0.001). The Least Significant Difference (LSD) Test was used to determine all possible significant differences among the various displays. Using this test, it was possible to isolate where those differences existed in total mental workload scores. This test indicated that subjects using the multiple two dimensional displays had significantly lower total mental workload ratings than when using the composite three dimensional display. Similarly, it was found that subjects using the multiple two dimensional displays also had significantly lower total mental workload ratings than when using composite two dimensional display.

High workload in the NASA TLX tool was represented by a maximum rating score of 300, while low workload was represented by a minimum rating score of 0. Average total workload scores for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 154.8, 195.5, and 206.7 respectively, while the standard deviations for these displays were 40.6, 33.8, and 24.3 respectively. These average scores show that subjects rated the multiple 2-D display has imposing a fair amount of workload, while the composite displays imposed a moderately high amount of workload. Average total workload scores across subjects was 185.67 with a standard deviation of 39.66. Diagnostic information regarding the specific dimensions of total workload follows. Figure 4.7 illustrates the average total workload rating scores for each display.

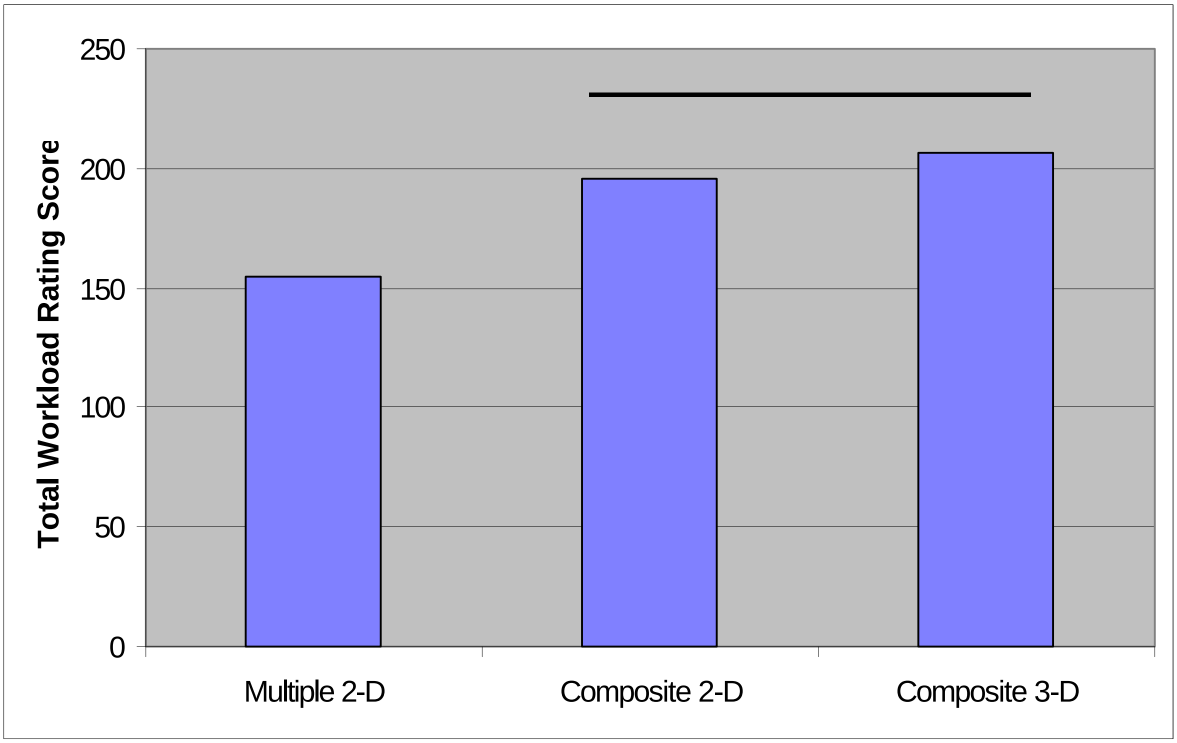


Figure 4.7: Average total workload rating scores for each display

Mental Demand

Mental demand was computed by a subjective rating of mental demand multiplied by a weight determined during the baseline task. The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of mental demand ratings across the displays are equal ($F_{(2,22)}$=7.89, p=0.003). From the Least Significant Difference test, it was found that the multiple 2-D displays had a significantly lower mental demand rating than the composite 2-D displays, and that the multiple 2-D displays had a significantly lower mental demand rating that the composite 3-D displays. Average mental demand ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 38.17, 58.92, and 55.0 respectively, while standard deviations for these displays were 24.29, 23.47, and 24.38 respectively. Average mental demand ratings across subjects was 50.69 with a standard deviation of 25.07. Figure 4.8 illustrates the average mental demand rating scores for each display.

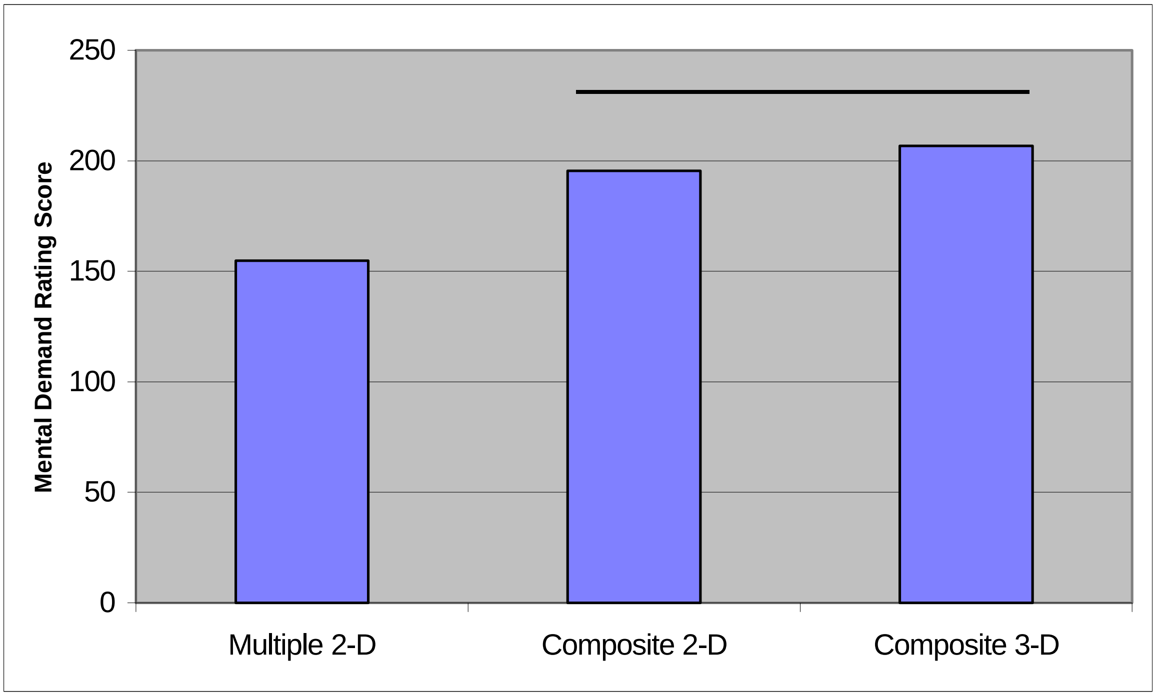


Figure 4.8: Average mental demands rating scores for each display

Physical Demand

The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of physical demand ratings across the displays are equal ($F_{(2,22)}$=5.08, p=0.015). The Least Significant Difference test indicated that physical demand ratings were significantly lower for multiple 2-D displays than for composite 2-D displays. The average physical demand rating for multiple 2-D, composite 2-D, and composite 3-D perspective displays were 17.08, 24.08, and 19.67 respectively, with standard deviations of 22.41, 32.02, and 26.12 respectively. Average physical demand ratings across subjects was 20.28 with a standard deviation of 26.52. Figure 4.9 illustrates the average physical demand rating scores for each display. Several subjects did not believe that the baseline task given to them was physically demanding. As a result, no weighting was given to this workload dimension which resulted in scores of zero for the physical demand ratings for all treatment levels.

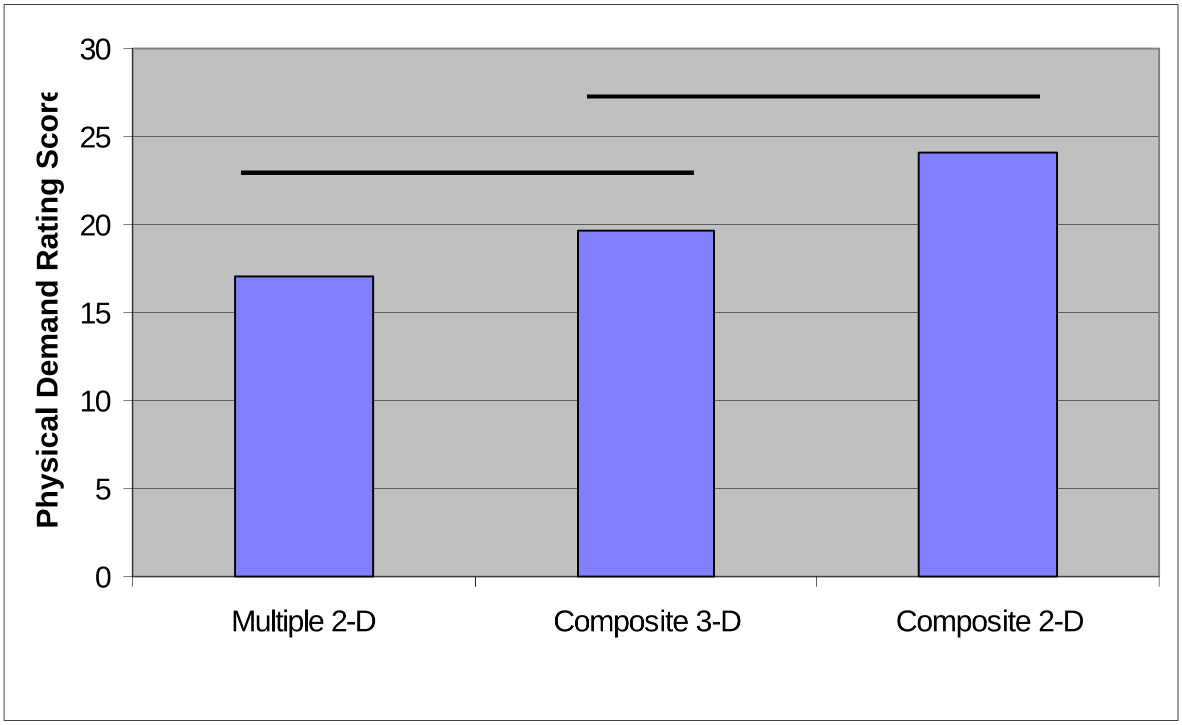


Figure 4.9: Histogram of subject physical demand rating scores for each display

Temporal Demand

The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of temporal demand ratings across the displays are equal ($F_{(2,22)}$=3.85, p=0.037). Post-hoc analysis was conducted using the Least Significant Difference test and indicated that there was a significantly lower temporal demand rating for the multiple 2-D displays than for the composite 3-D perspective display. The average temporal demand ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 33.00, 37.42, and 46.58 respectively with standard deviations of 16.71, 12.23, and 17.49 respectively. Average temporal demand ratings across subjects was 39.00 with a standard deviation of 16.24. Figure 4.10 illustrates the average temporal demand rating scores for each display.

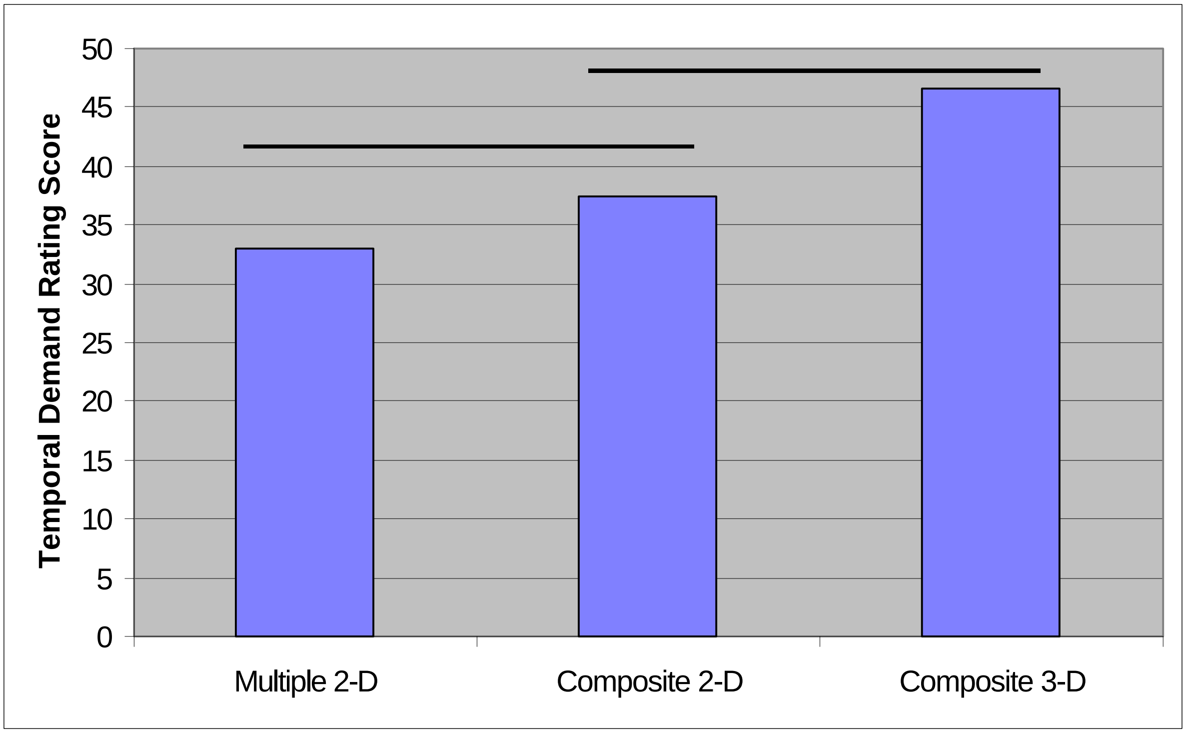


Figure 4.10: Average temporal demand rating scores for each display

Observed Performance

The ANOVA indicated no significant main effect of display, accepting the null hypothesis that the means of observed performance ratings across the displays are equal ($F_{(2,22)}$=1.03, p=0.375). Average rating scores in observed performance for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 33.85, 27.17, and 31.83 respectively with standard deviations of 12.49, 12.44, and 20.84 respectively. Average observed performance ratings across subjects was 30.92 with a standard deviation of 15.56.

Frustration

The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of frustration ratings across the displays are equal ($F_{(2,22)}$ = 4.62, p=0.021). The Least Significant Difference test indicated that subjects gave significantly lower frustration rating scores for the multiple 2-D displays, than for the composite 3-D perspective display. Average ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 11.5, 14.5, and 20.0 with standard deviations of 16.8 17.3, and 26.3 respectively. The average frustration ratings across subjects was 15.33 with a standard deviation of 20.31. Figure 4.11 illustrates the average frustration rating scores for each display. Several subjects did not believe that the baseline task given to them was frustrating. As a result, no weighting was assigned to this workload dimension which resulted in scores of zero for the frustration ratings for all treatment levels for some of the subjects.

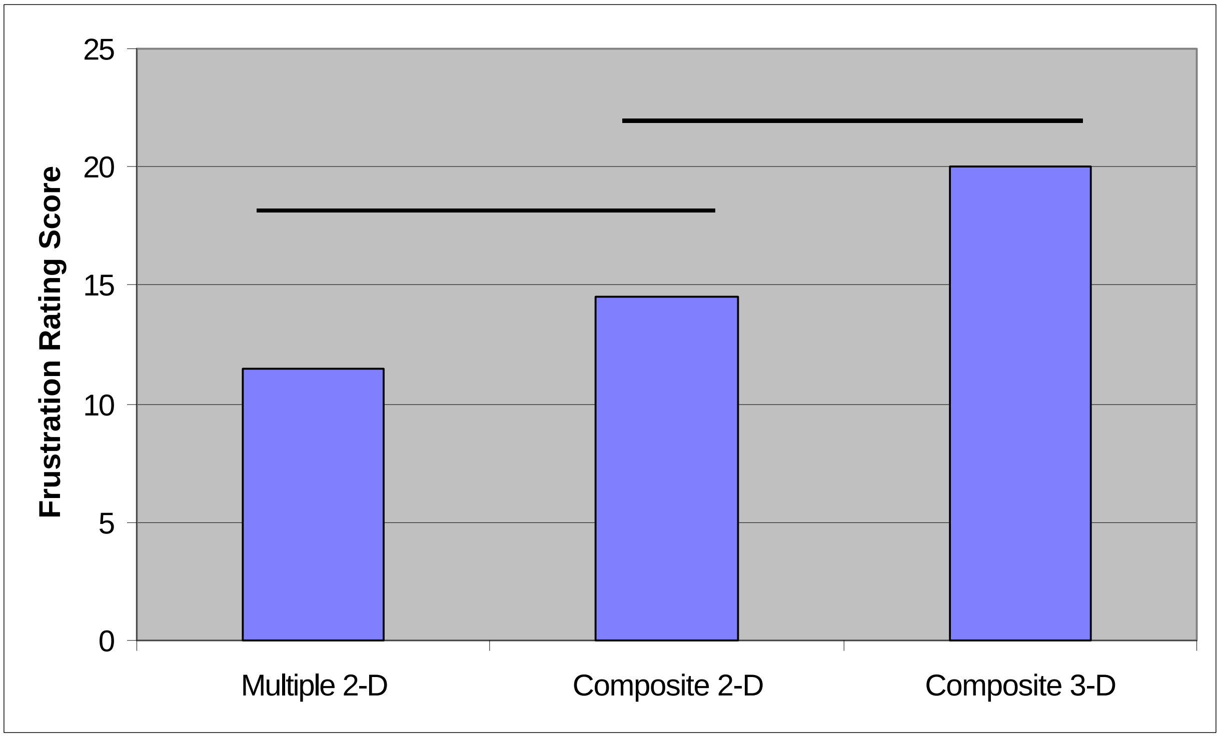


Figure 4.11: Average frustration rating scores for each display

Effort

The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of effort ratings across the displays are equal ($F_{(2,22)} = 4.29$, p=0.027). Using the Least Significant Difference test, multiple 2-D displays had significantly lower effort ratings than the 2-D composite displays. Similarly, the multiple 2-D displays also had significantly lower effort ratings than the 3-D perspective composite display. Average effort rating scores for the multiple 2-D, composite 2-D, and composite 3-D displays were 21.33, 33.42, and 33.58 respectively, with standard deviations of 11.23, 17.71, and 16.50 respectively. Average effort ratings across subjects was 29.44 with a standard deviation of 16.05. Figure 4.12 illustrates the average effort rating scores for each display.

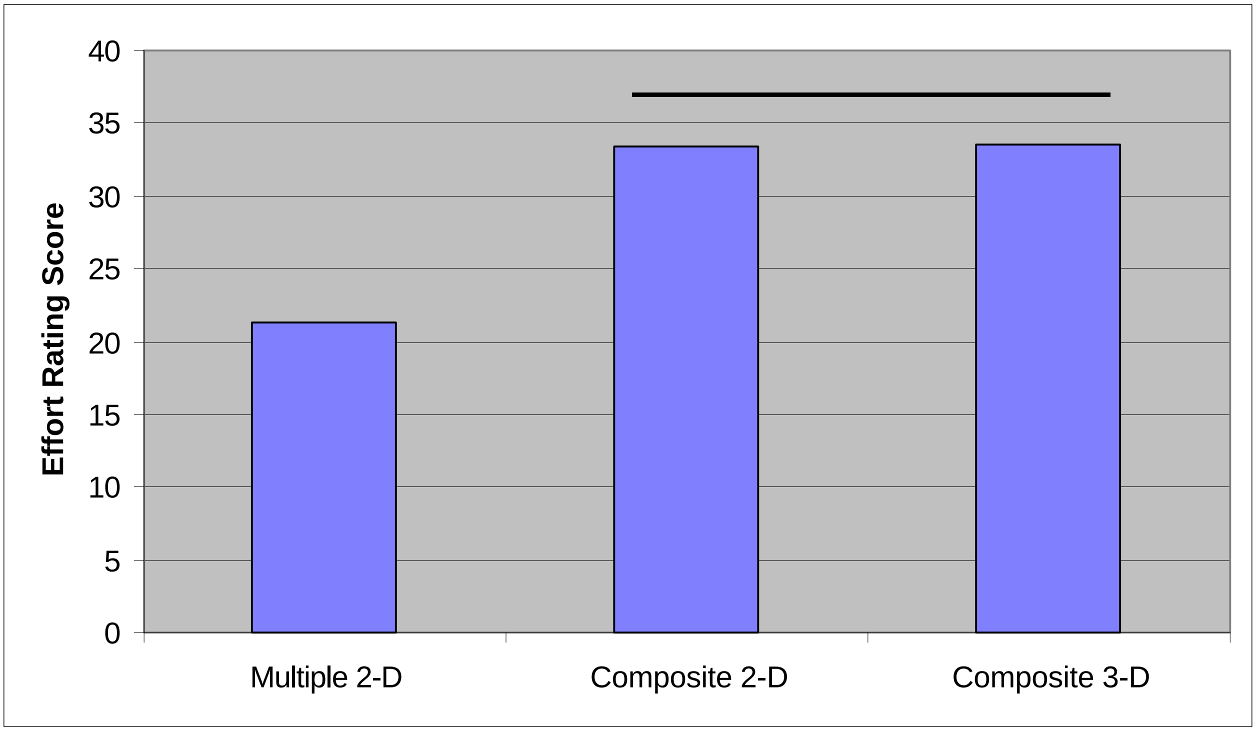


Figure 4.12: Average effort rating scores for each display

Workload Summary

Table 4.2 summarizes the NASA TLX results for this study:

Table 4.3: Summary of Analysis of Variance and Least Significant Difference Tests for NASA TLX

Rating Scale	F-Value	p-value	Significant	Significant Levels from Least Significant Difference Test
Total Workload	$F_{(2,22)}$=9.42	0.001	significant**	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Mental Demand	$F_{(2,22)}$=7.89	0.003	significant**	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Physical Demand	$F_{(2,22)}$=5.08	0.015	significant*	- Multiple 2D, composite 2D
Temporal Demand	$F_{(2,22)}$=3.85	0.037	significant*	- Multiple 2D, composite 3D
Observed Performance	$F_{(2,22)}$=1.03	0.375	n.s.	-
Frustration	$F_{(2,22)}$=4.62	0.021	significant*	- Multiple 2D, composite 3D
Effort	$F_{(2,22)}$=4.29	0.027	significant*	- Multiple 2D, composite 2D - Multiple 2D, composite 3D

* significant at p <.05
** significant at p <.01

The individual ANOVA tables for the workload measures can be found in Appendix A2.

Subjective Measures

Subjective data was collected in order to understand how the displays affected subject sentiments regarding each display. Subjective questions dealing with decision confidence, accuracy, and time were given to subjects following each treatment in order to provide insight into the objective measures collected during the experiment. These questions utilized a five point Likert-type scale to help quantify this subjective data.

Subjective Rating of Decision Confidence

Decision confidence was assessed through a post-treatment question given to subjects after completing all eight events for a display (see Question 1 in Appendix B6). Specifically, subjects were asked to rate their decision confidence on a five-point Likert-type scale, with the following categories: strongly disagree(S/D), disagree(D), undecided(U), agree(A), and strongly agree(S/A). The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of decision confidence across the displays are equal ($F_{(2,22)}$=14.44, $p < 0.001$). Figure 4.13 illustrates data for the subjective rating of decision confidence.

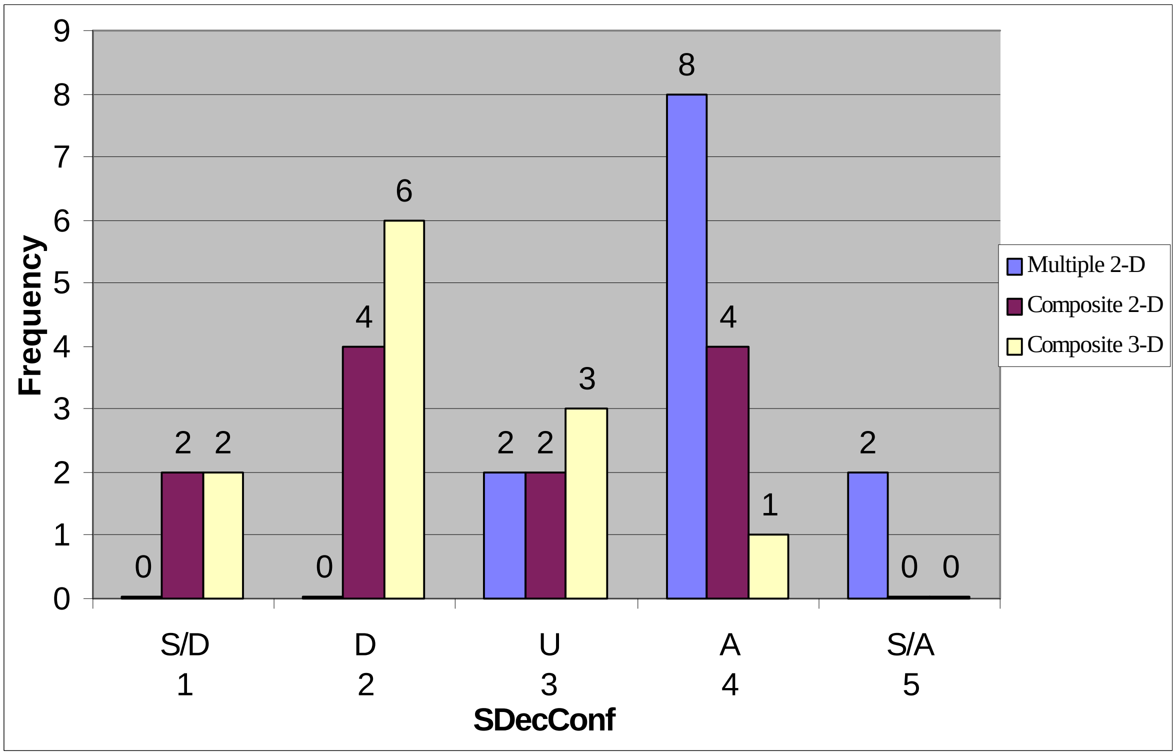


Figure 4.13: Histogram of subjective decision confidence rating for the various displays

The Least Significant Difference test indicated that multiple 2-D displays had higher decision confidence ratings than the composite 2-D display, and that multiple 2-D displays also had higher decision confidence ratings than the composite 3-D ratings as shown in Figure 4.14. These results partially support the experimental hypothesis that the subjects using composite 3-D display would have less decision confidence ratings than 2-D displays. No significant difference was found between composite 2-D, and composite 3-D displays in the subjective decision confidence ratings. Average decision confidence ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 4.00, 2.67, and 2.25 respectively, with standard deviations of .60, 1.16, and .87 respectively.

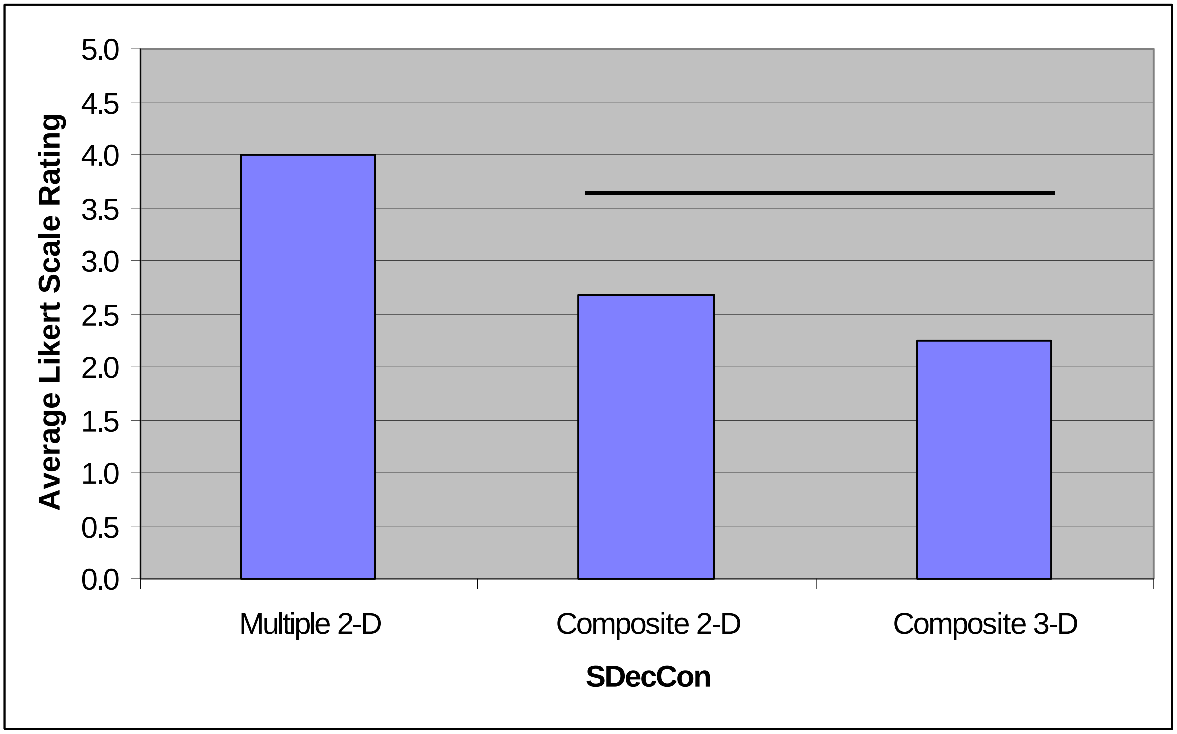


Figure 4.14: Average subjective decision confidence ratings for the various displays

Subjective Rating of Decision Time

Subjects were also given a post-treatment question asking them to assess whether the display helped subjects recognize non-random events quickly (see Question 2 in Appendix B6) using a Likert-type scale. The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of decision time ratings across the displays are equal ($F_{(2,22)}$=11.23, $p < 0.001$). Figure 4.15 illustrates data for the subjective rating of decision time.

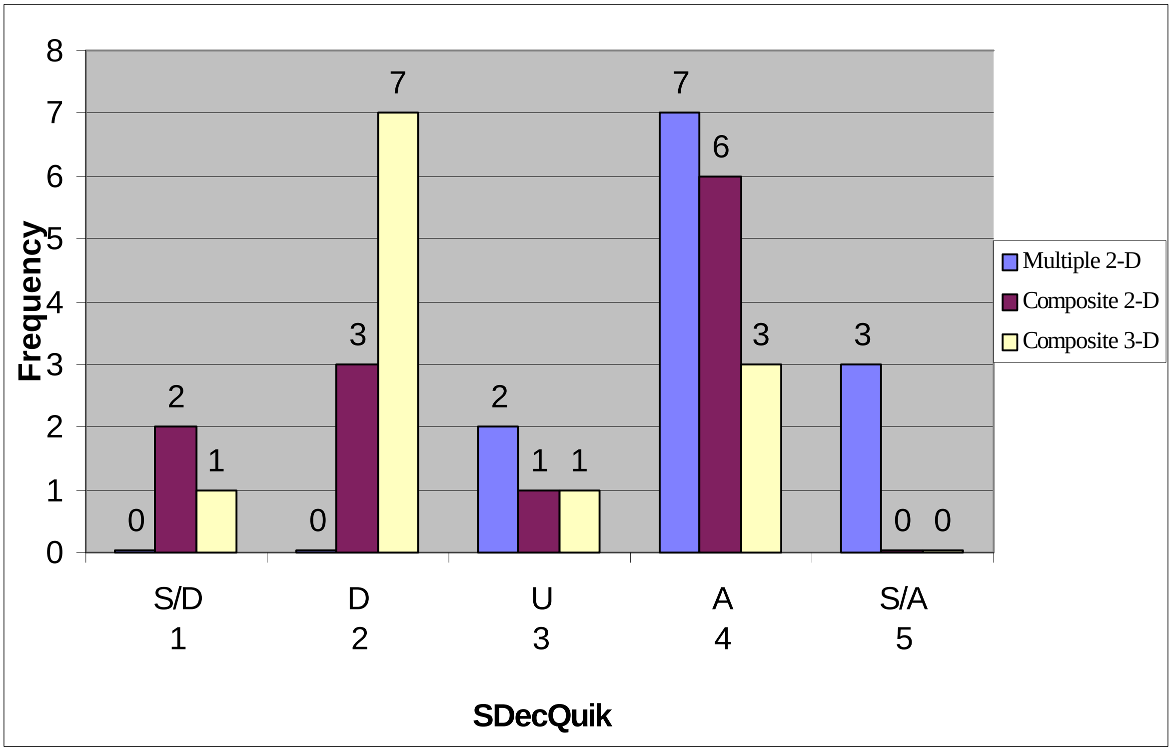


Figure 4.15: Histogram of subjective decision quickness rating for the various displays

The Least Significant Difference test indicated that multiple 2-D displays had higher decision quickness ratings than the composite 2-D display, and that multiple 2-D displays also had higher decision quickness ratings than the composite 3-D ratings as shown in Figure 4.16. No significant difference was found between composite 2-D, and composite 3-D displays in the subjective decision quickness ratings. Average decision quickness ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 4.08, 2.92, and 2.50 respectively, with standard deviations of 0.67, 1.24, and 1.00 respectively.

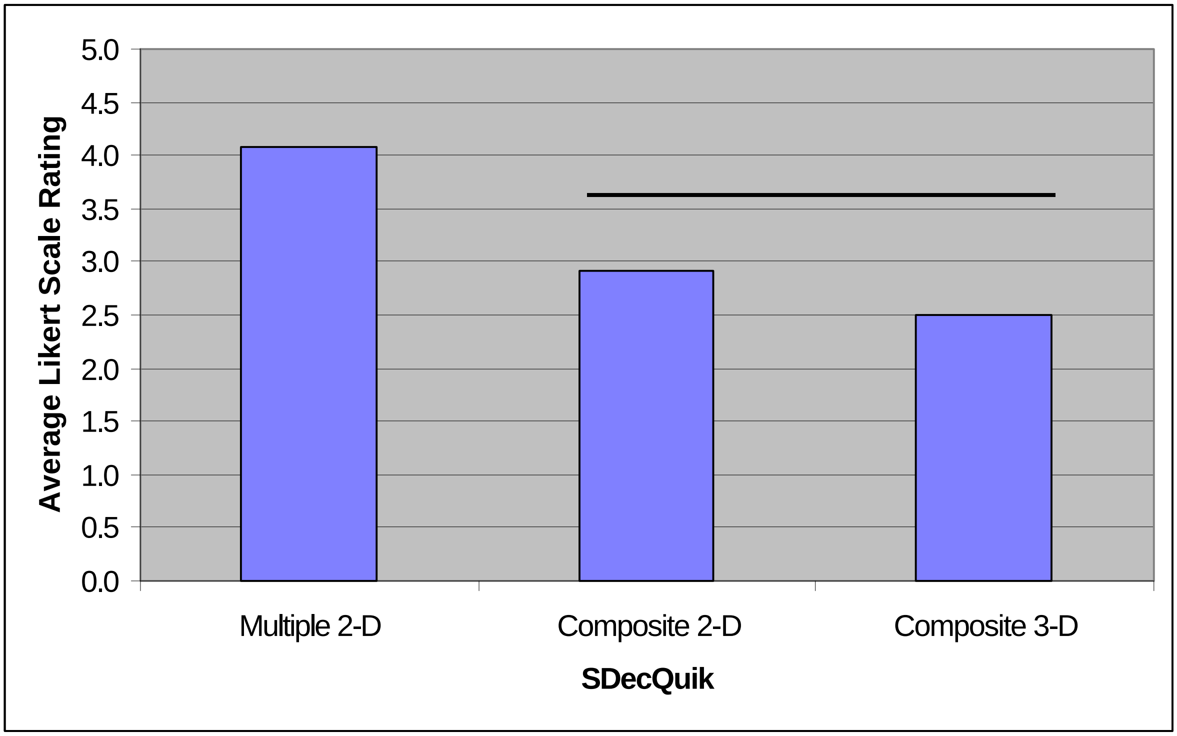


Figure 4.16: Average subjective decision quickness ratings for the various displays

Subjective Rating of Decision Accuracy

Subjective assessments of decision accuracy were collected via a post-treatment question (see Question 3 in Appendix B6) using a Likert-type scale. Specifically, subjects were asked whether the display helped their decision accuracy for the experimental task. The ANOVA indicated a significant main effect of display, rejecting the null hypothesis that the means of decision accuracy ratings across the displays are equal ($F_{(2,22)}$=12.36, $p < 0.001$). Figure 4.17 illustrates data for the subjective rating of decision accuracy.

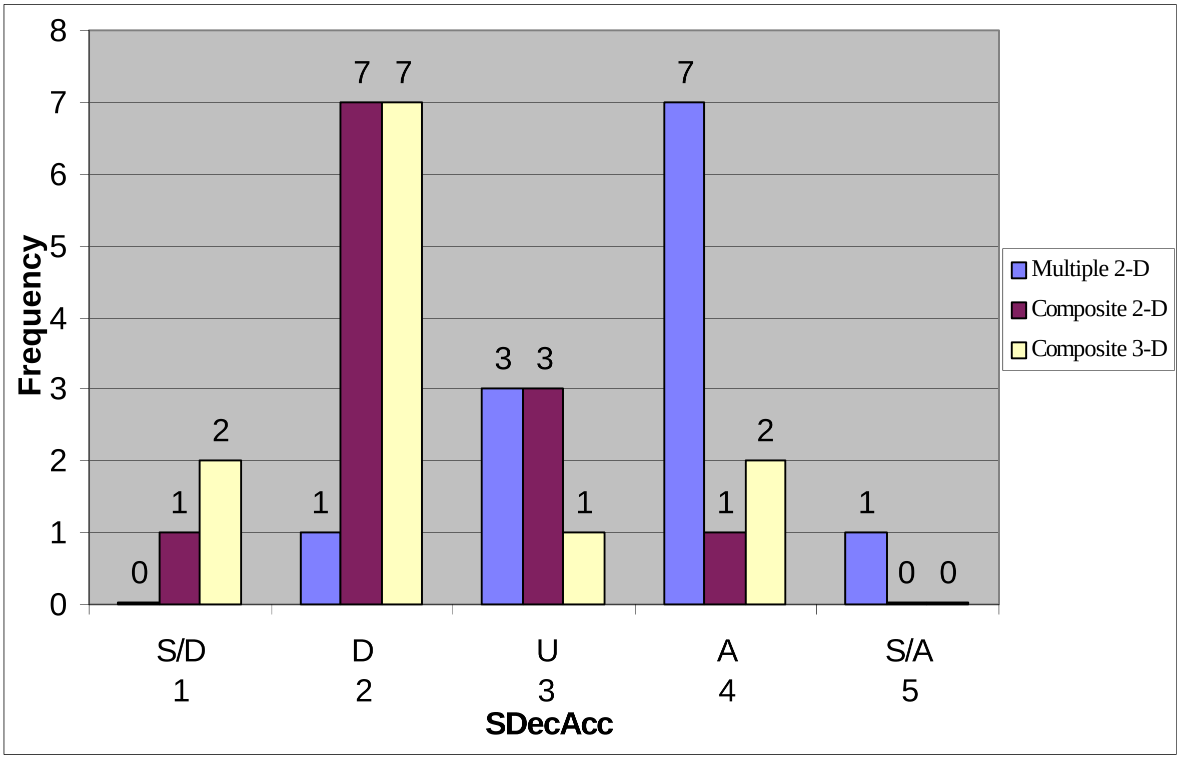


Figure 4.17: Histogram of subjective decision accuracy rating for the various displays

The Least Significant Difference test indicated that subjects significantly rated the multiple 2-D displays higher on the decision accuracy question than the composite 2-D display. Similarly, the Least Significant Difference test also indicated that subjects significantly rated the multiple 2-D displays higher on the decision accuracy question than the composite 3-D perspective display as shown in Figure 4.18. No significant difference was found in these ratings between the composite 2-D and the composite 3-D perspective displays. Average decision accuracy ratings for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 3.67, 2.33, and 2,25 respectively, with standard deviations of 0.78, 0.78, and 0.97 respectively.

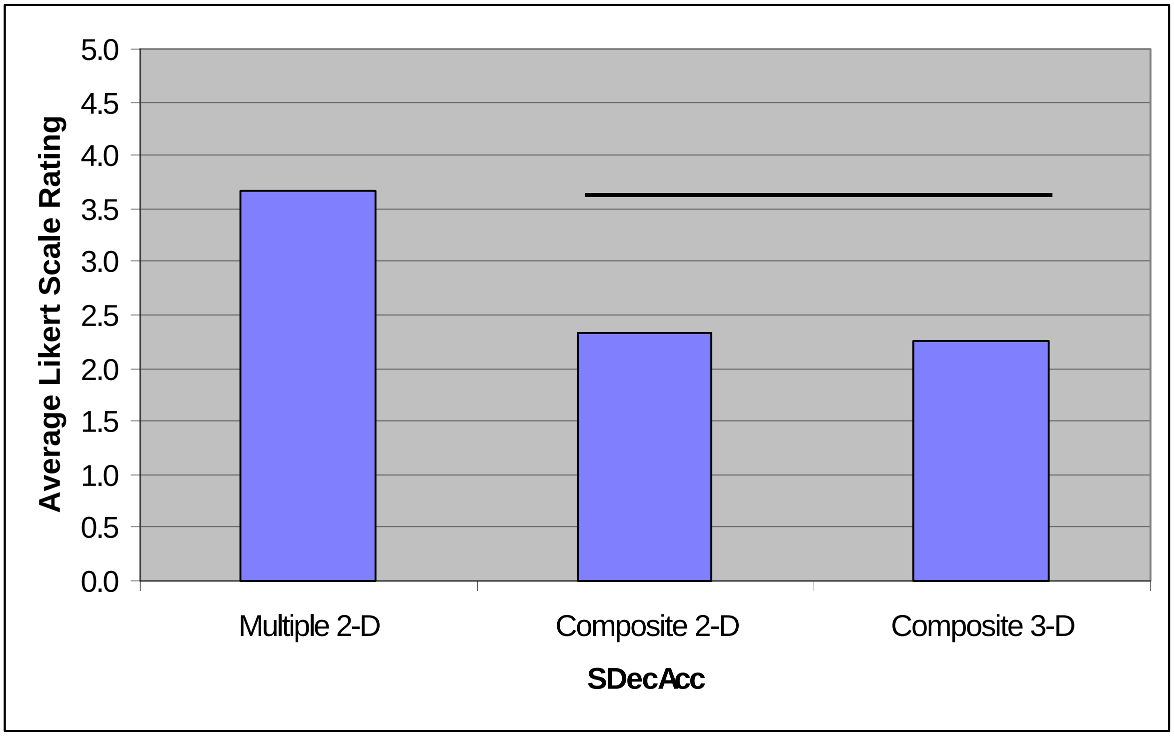


Figure 4.18: Average subjective decision accuracy ratings for the various displays

Subjective Preferences

Another subjective measure was incorporated into the post-experiment to understand subjective preferences and recommendations regarding the displays and the task in general.

Display Preferences

In a post-experiment questionnaire, subjects were asked to rank the various displays based on their overall preferences following the completion of the entire experimental task(see Appendix E2). Using the Friedman test for rank untied data, it was found that subject preferences were significantly different among the displays (k=3, n=12, X^2_{ranks} = 15.17). Figure 4.19 illustrates a histogram of the rankings for each of the displays.

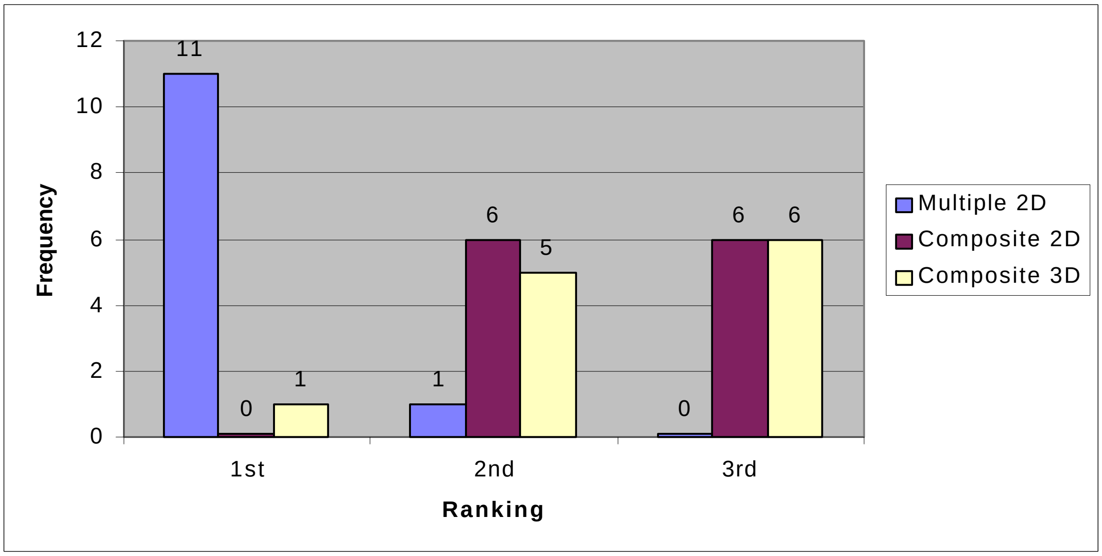


Figure 4.19: Histogram of subject overall rankings of the various displays

Post hoc paired comparisons were performed on the rank totals of each display to identify the specific differences in the ranks for the displays. From this analysis, it was found that subjects ranked the multiple 2-D displays significantly better than the composite 2-D display, and the multiple 2-D displays significantly better than the composite 3-D perspective display. No significant differences between rankings were found between the composite 2-D display, and the composite 3-D display. Average rankings (see Figure 4.20) for the multiple 2-D, composite 2-D, and composite 3-D perspective displays were 1.08, 2.50, and 2.42, with standard deviations of 2.89, 0.52, 0.67 respectively.

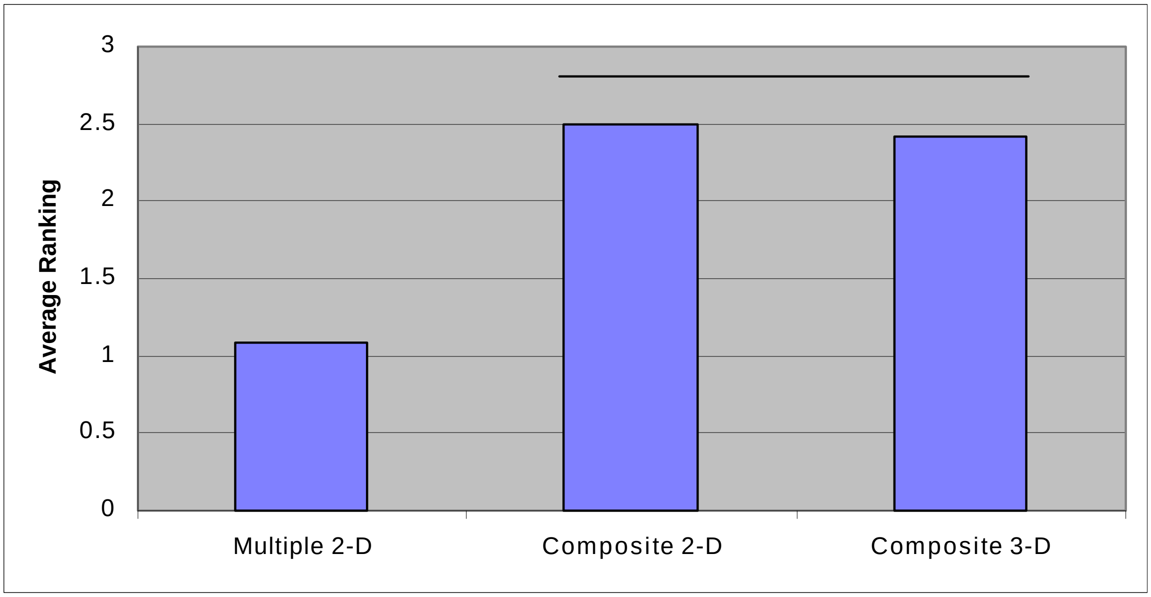


Figure 4.20: Average subject rankings of the various displays

Other Post-Hoc Analyses

Relationship between composite 3-D accuracy and Cube Comparisons Test

To test whether 3-D ability was related to the accuracy scores obtained from the composite 3-D perspective display, a Pearson-r correlation coefficient was calculated for these two variables. To test whether this correlation was significant, a t-test of significant was used to see whether this correlation coefficient was significantly different than zero. The results of this analysis showed that the correlation between these two variables ($r_{3d,cct}$=.389) was not a significant correlation.

Relationship between decision performance and mental workload

Since both of these factors were found to be significantly affected by the type of display, an analysis looking at the relationship between these two variables was conducted. To understand if decision performance is related to mental workload assessments of subjects, a Pearson-r correlation coefficient was calculated. The correlations between these variables are shown in Table 4.4.

Table 4.4: Pearson-r correlation coefficients between decision performance variables and total mental workload

Relationship	Pearson-r correlation coefficient
Total Mental Workload and Decision Accuracy	r = -0.233
Total Mental Workload and Search Time	r = -0.095
Total Mental Workload and Decision Time	r = 0.377
Total Mental Workload and Search + Decision Time	r = 0.015
Total Mental Workload and Stopping Time	r = -0.098

Relationship between decision accuracy and decision time

To test whether a tradeoff existed between decision accuracy and decision time, correlation coefficients were computed between the decision accuracy and search time($r_{da,st}$ = -0.08), decision accuracy and decision time($r_{da,dt}$ = -0.12), and decision accuracy and search + decision time($r_{da,s+dt}$ = -0.12). From these calculations, none of correlation coefficients for these relationships were found to be significant.

Relationship between subjective and objective measures of decision time

To see whether any relationship existed between the subjective and objective measures of decision time, a Pearson-r correlation coefficient was calculated as part of the post hoc analysis. The correlation between subjective decision time and the various measures of decision time are shown in Table 4.5. None of these correlations were found to be significant.

Table 4.5: Pearson-r correlation coefficients among subjective and objective measures of decision time

Relationship	Pearson-r correlation coefficient
Subjective Decision Time and Search Time	r = 0.04
Subjective Decision Time and Decision Time	r = 0.28
Subjective Decision Time and Search + Decision Time	r = 0.12
Subjective Decision Time and Stop Time	r = 0.04

Relationship between subjective measures of decision time, accuracy, and confidence

To test whether a relationship existed between the subjective measures of decision time, accuracy and confidence existed, a Pearson product moment correlation matrix was computed. The results of this correlation matrix are provided in Table 4.6 with t-values to test significance provided in parentheses. The results indicated that all of the correlation coefficients among the subjective measures were significant.

Table 4.6: Pearson product moment correlation matrix for subjective measures

	Decision Confidence	Decision Quickness
Decision Time	0.858 (t=9.74)	
Decision Accuracy	0.838 (t=8.95)	0.769 (t=7.01)

Relationship between Display type and Non-Random Signal

Given the fact that there were different types of non-random signals presented in each type of display, an analysis to see whether a relationship existed between the type of display and the number of non-random signals correctly identified was conducted. Table 4.7 illustrates the observed frequencies of accuracy for each display type. A chi-squared test was conducted on the display type and the type of signal. From this analysis, it was important to note than one of the expected frequency cells fell below five, a critical number when conducting the chi-squared test. However, Cochran (1954) and Maxwell (1961) states that this rule is too stringent, and states that it is permissible to have a few expected frequencies below five.

Table 4.7: Observed Frequencies of Accurate Classifications of Non-Random Signals

Type of Signal	Multiple – 2D	Composite 2D	Composite 3-D
Outside the Control Limits	19	22	14
Phase Related Correlation	10	3	2
Run	17	11	14
Trend	19	13	13

From this analysis, it was found that the type of display was not significantly related to the accuracies relating to the non-random signals ($\chi^2_{(6)} = 6.44$).

Summary of Results

Table 4.8 summarizes the results of the statistical analyses conducted of the collected data.

Table 4.8: Summary of statistical analyses relating to the experimental hypotheses

Variable	Test-Statistic	p-value	Significant	Significant Levels using Least Significant Difference Test
Workload Dimensions				
Total Workload	$F_{(2,22)} = 9.42$	0.001**	significant	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Mental Demand	$F_{(2,22)} = 7.89$	0.003**	significant	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Physical Demand	$F_{(2,22)} = 5.08$	0.015*	significant	- Multiple 2D, composite 2D
Temporal Demand	$F_{(2,22)} = 3.85$	0.037*	significant	- Multiple 2D, composite 3D
Observed Performance	$F_{(2,22)} = 1.03$	0.375	n.s.	-
Frustration	$F_{(2,22)} = 4.62$	0.021*	significant	- Multiple 2D, composite 3D
Effort	$F_{(2,22)} = 4.29$	0.027*	significant	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Decisional Variables				
Decision Accuracy	$F_{(2,22)} = 11.52$	< 0.002**	significant	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Decision Confidence	$F_{(2,22)} = 14.44$	< 0.001***	significant	- Multiple 2D, composite 2D - Multiple 2D, composite 3D
Search Time	$F_{(2,22)} = 0.13$	0.877	n.s.	-
Decision Time	$F_{(2,22)} = 0.64$	0.536	n.s.	-
Search + Decision Time	$F_{(2,22)} = 0.33$	0.724	n.s.	-
Stopping Time	Mult – 2D: T = -0.93 Mult- 3D: T = 0.92 2D – 3D: T = 1.63	Mult – 2D: 0.36 Mult- 3D: 0.37 2D – 3D: 0.12	n.s.	-
Probability of Hit	$F_{(2,22)} = 0.74$	0.487	n.s.	-
Probability of False Alarm	$F_{(2,22)} = 1.63$	0.219	n.s.	-

* significant p <.05
** significant p <.01
*** significant p <.001

Chapter 5: Discussion

Subjects' Demographics

The data gathered from the pre-experimental questionnaire provides some insight regarding the population used for this research effort. The sample population involved with this research was taken from a well educated population, the average subject having been a junior. On average, this sample population can be described as one with some Internet experience, little visualization experience, and little statistical process control(SPC) or quality control (QC) experience. It was desirable to have subjects with some Internet experience, since the majority of the experiment was conducted using Netscape Navigator 4.03 as its primary graphical user interface. The fact that subjects on average had little statistical process control or quality control knowledge was extremely important for this research, since subjects were ultimately trained through the use of a progressive part training method in SPC or QC. Testing individuals several times and using an 80% criterion for passing subjects allowed the experimenter to control some of the differences in knowledge in SPC or QC necessary to complete the experiment. One quarter of the subjects failed at least one evaluation test following one of the training modules. No restrictions were placed upon the level of experience relating to visualization applications.

Effect of order on dependent variables

The results indicated that order did not have a significant effect on all the dependent variables. This indicates that the potential learning effect was not an issue when considering all of the dependent variables. The lack of significance is extremely important since measurement of workload relies on the basic premise that a minimal amount of learning is present. A minimal amount of learning is necessary since workload focuses on the demands associated with the task assuming the human is already familiar with the task(Gopher and Donchin, 1986).

Effect of display type on decisional characteristics

Decision accuracy and display type

The type of display was found to have a significant effect on the decision accuracy of subjects. Further analysis showed that decision accuracy was significantly higher for multiple two dimensional displays than for composite two dimensional displays. Similarly, this analysis also showed that decision accuracy was significantly higher for the multiple two dimensional displays than for composite three dimensional perspective displays. These results are partially opposite to the experimental hypothesis that individuals using three dimensional perspective control charts will have better decision accuracy than when using two dimensional control charts.

As mentioned in the previous chapter, the average decision accuracy for the multiple 2-D display, composite 2-D display, and composite 3-D display was 0.68, 0.51, and 0.30 respectively (where 1.0 equals 100% correctly answered). To shed some light on these results, other studies which share common elements with this research were also reviewed. Gramopadhye, Drury, and Sharit (1993) found that subjects who were passively trained (using a progressive part training scheme) in aircraft inspection were able to attain an average decision accuracy of 0.735 using a single 2-D computer display. In a different study, Schoonard and Gould (1973) presented photographs of integrated chips to inspectors. These inspectors had a restricted field of view, but were able to attain an overall accuracy rate of 0.73 when given sixteen seconds to inspect the photographs. This research allowed subjects to view the data for five seconds before introducing a new set of data points. These accuracy results are similar to the multiple 2-D display results computed in this research.

The Proximity Compatibility Principle (PCP) provides some insight into these results. First, the use of the composite displays may help to reduce the amount of visual search time as well as memory resources required to shift between data sets. This term within the PCP context is known as the *information access cost*. Second, the presence of overlap between data sets as a result of the close spatial proximity among data sets is known within the PCP principles as *clutter, overlap, and confusion*. Theoretically, by placing data sets in close proximity to one another, the information access costs can be reduced, since less eye movements are required of subjects to move from data set to data set. As a result, this should increase performance, thus improving accuracy. However, the close spatial proximity also had negative consequences as a result of integrating the graphs into a single display. The close spatial proximity caused data points and the lines between them to frequently overlap. When this overlapping occurs, PCP predicts that performance would decline. The decline could be associated with the additional processing demands required to segregate the overlapping data. These additional processing requirements negated any gains obtained from reducing the visual search time, and imposed additional requirements necessary to separate overlapping data sets. Subjects supported this notion, only complaining of overlapping data for the composite displays.

The objective of these control chart displays is to function as decision tools which convert data into information. The ability of decision makers to accurately perceive this information to make the correct decision is an important element in system design. For this research, the decision making task was specifically constrained within the statistical quality control or process control domain. Users of control charts represent detectors of non-random signals using one of the displays (decision tools). As detectors, the results indicated that subjects were more successful identifying a non-random signal using the multiple 2-D displays than either composite display. At the same time, the results also indicated that there were no significant differences in decision accuracy between the composite displays. The implications of these results suggest that the use of composite displays may not be an effective way of portraying information for the decision maker to make an accurate assessment of work system or process performance.

Decision accuracy and signal type

For the specific signals, the results indicated that there were significant differences in the ability of subjects to identify the various non-random signals. Subjects were significantly more accurate in identifying outside the control limits signals than the runs, trends, and phase related correlation. Subjects also were significantly more accurate in identifying runs and trends than phase related correlation. One possible explanation for these results could be the number of points required to make a positive identification for any given signal. For example, in order to identify an out of control signal, only one point needs to be outside the control limits. Similarly, in order to identify a trend, six consecutive points in a monotonically increasing or decreasing pattern must appear. For the run, seven consecutive points above or below the mean must be present. For the phase related correlation, five points on two separate data sets must appear to be correlated. In order to identify phase related correlation the user must track a total of ten points between the two graphs in order to positively identify the phase related signal. Previous research (Miller, 1956; Baddley, 1994) has found that humans have difficulty with remembering and handling more than seven plus or minus two chunks of information into short term memory. Each data point may represent a chunk of information. In order to identify a non-random pattern, the relative location of each point must be stored into short term memory. Also, the number of points which meet the criteria defined by the definitions of the non-random signals must also be stored. In other words, for four points which have monotonically increased on the control chart, the subject must remember that it is the fourth point of a trend. The problems associated with recognizing the phase related signals may be the result of storing too many chunks of information into short term memory, since it requires ten points between two graphs for a positive identification.

At the same time, outside the control limit signals, runs, and trends require the subject to focus on one data set in order to evaluate whether the process in random or not. Phase related correlation requires that subjects focus their concentration or attention along two separate data sets. PCP suggests that when subjects are required to shift their attention between data sets whether across displays or within a display, a decrease in performance can be expected due to information access costs (Van der Heijden, 1992; Wickens and Carswell, 1995). The shift in attention between sets of data occurs when subjects are required to look across two data sets to identify phase related signals. Therefore subjects are engaged in storing a large amount of information into working memory, while shifting attention between data sets in order to recognize phase related correlation. Furthermore, since phase related correlation depends on the behavior of two processes, overlap may play a critical role in identifying the presence or lack of presence of a signal.

The results also indicated that there was no significant correlation between the type of display used and the type of signal in terms of decision accuracy. This indicates that subjects were not better at looking at specific signals using one specific type of display. Rather, the results showed that in terms of accuracy the type of display and type of signal are independent. This result is logical since signals are based on a specific type of pattern, whereas the display is the means by which a subject views the patterns on the screen.

Although little research appears to exist dealing with human detection of these non-random signals, much research has been conducted using computers to automatically detect non-random signals. However, the type of non-random signals used in other studies do not correspond to the signals used for this research study. Therefore, no direct comparisons can be made between other studies and this research effort. Three computerized approaches used by other researchers include back propagation (Pham and Oztemel, 1992), learning vector quantization (Pham and Oztemel, 1994), and multilayer perceptron network Cheng (1997). These systems are based on neural net algorithms which are programs which attempt to "learn" how to recognize signals and patterns based on a set of specified rules. The different algorithms represent different methods which the computer uses to "learn" the process of identifying the non-random signals. All of these approaches were able to correctly classify various signals 95% of the time. These computerized approaches appear to have better correct classification rates that human operators, however any direct comparisons between the human detector and an automated detection system is premature, especially considering the fact that subjects had difficulty identifying phase related signals.

Signal Detection Theory Measures

The type of display did not have a significant effect on sensitivity or response criterion in this research. Similarly, the components used to compute these signal detection theory (TSD) values, the hit rate and false alarm rate, also were not significantly affected by the type of display. At first glance, this may appear to conflict those results mentioned earlier in the decision accuracy section. The difference in results can be explained by how each of these measures was computed. Decision accuracy was computed using a holistic perspective by treating each set of data as one event. For the TSD measures, each individual data point within that data set was treated as an opportunity for the subject to classify the data as a signal (out of control) or noise (in control). The lack of significant differences in sensitivity may be the result of how subjects appeared to process the data. When subjects classified an out of control process as in control, subjects would allow the process to continue waiting for a recognizable signal. They may not have looked back to previous points which they may have missed to reevaluate the entire process. Furthermore, subjects appeared to process a limited window of data points since new points were constantly being introduced. This can be indirectly supported by the fact that subjects rarely correctly identified the processes beyond the five point window imposed by the task. As a result, the miss rate would have increased, which would have decreased the hit rate. This may have caused the sensitivity measure to be insensitive to the effects of the display used.

The lack of significance in the response criterion as a result of display is not a surprise. Subjects were given the same cost and benefits (see cost/benefit structure in Appendix B2, B3, and B4) associated with the four possible outcomes associated with TSD. The costs and benefits help provide information for a decision strategy to meet the objectives of the task. The goals of this experiment were to maximize overall profit of each facility. These costs and benefits were shared among the displays which ultimately help keep the response criterion relatively constant.

Decision Making Time

Based on the results of the experiment, it was found that the type of display did not have a significant effect on the search time, decision time, search + decision time, and stopping time. These results do not support the hypothesis that individuals using composite three dimensional perspective control charts will require more decision time than individuals using two dimensional control charts. Although visual search time can be theoretically decreased with a composite display, the problems associated with depth ambiguity imposed by the perspective employed in the three dimensional display and the presence of overlapping may have cancelled out any gains associated with the close proximity of the data sets. Depth ambiguity is not present in the composite two dimensional display. However, the overlapping of data points was more severe than in the three dimensional perspective display, and as a result could have negated any gains associated with minimizing visual scanning time. In a study conducted by Wickens, Merwin, and Lin (1994), dimensionality did not have a significant effect on solution times for similar tasks which require a low to moderate amount of integration of information across data sets.

Decision time was not significantly affected by the type of display used. Decision time involves the subjects ability to correctly characterize a signal once a signal was detected. One possible explanation for the lack of significance in decision time due to the display type could be the way subjects searched for signals. Subjects only indicated a signal when they were able to characterize it. Thus the decision time may have only measured the subjects ability to select the appropriate signal from the menu of signals available to them on the screen. As a result, decision time should not be affected by the type of display, since the characterization of the signal may have taken place during the visual search for non-random signals.

Since search time and decision time were not significantly affected by the type of display used, it would be logical that both search + decision time, and stopping time were also not significant. This lack of significance is expected since the components of search + decision time and stopping time are linear combinations of the search time and decision time.

Decision accuracy and decision time tradeoff

One area of interest in utilizing different display formats is to test whether any tradeoffs exist between decision time (measured by search time, decision time, and search + decision time) and decision accuracy. The results indicated no significant relationship between these variables. The lack of significance between the decision and search times and decision accuracy provides some vital information regarding these displays. Since the various measures of search and decision times were not significantly affected by the type of display used, none of the displays has a clear cut advantage over the other displays in terms of decision time. However, the fact that subjects had better decision accuracy on multiple two dimensional displays than either composite displays suggests that the best display among the displays tested would be the multiple two dimensional display.

Effect of display type on Mental Workload

The type of display had a significant effect on total mental workload. The results indicated that mental workload was significantly less when using the multiple 2-D display than either composite display. Furthermore, there was no significant difference between the composite displays in any of the workload dimensions nor the total workload rating. This finding partially supports the experimental hypothesis that subjects using 3-D perspective control charts will exhibit more workload than individuals using two dimensional charts. Given the possible range of the mental workload score (0 = low mental workload, 300 = high workload), the average mental workload ratings of the multiple 2-D display (154.8) indicates that subjects felt that the display imposed a fair amount of workload. The average workload rating for the composite 2-D display (195.5) and composite 3-D display (206.7) indicates that subjects felt that these displays imposed a moderately high amount of mental workload.

Since total mental workload is computed by the summation of the weighted rating scores of several dimensions, an explanation of these results would be better served by looking at the individual dimensions which provide some diagnostic information relating to these results. The results of the workload analysis will be discussed and explained using concepts and guidelines derived from the Proximity Compatibility Principle.

Mental Demand Ratings

The type of display showed a significant effect in mental demand rating scores between multiple 2-D displays and composite 2-D displays, and multiple 2-D displays and composite 3-D perspective displays. No significant difference was found between the composite displays. The presence of overlapping with the composite displays could be a mediating factor in explaining the significant differences among the multiple and composite displays. The presence of overlap requires subjects to focus additional cognitive resources in segregating or separating data. Even without overlapping, the close proximity of two data sets can have a negative impact on the processing of each data set independently (Eriksen and Eriksen, 1974; Kramer and Jacobsen; 1991). The presence of overlapping points was a popular source of criticism of the composite displays by subjects. Given the fact that there was less overlapping within the three dimensional perspective displays than the two dimensional display, one would expect some significant difference in mental demand ratings to exist. One possible explanation for this lack of significant difference may lie in the fact that viewing three dimensional perspective data introduces some degree of depth ambiguity which requires additional cognitive resources to process the control chart data.

Physical Demand Ratings

The data indicated that there was a significant difference in physical demand ratings between the multiple two dimensional display and the composite two dimensional perspective display. The results showed that the composite two dimensional displays received significantly higher physical demand ratings than the multiple two dimensional displays. One possible explanation of these results is the requirements placed upon the user when overlapping of data occurs in a composite display. In order to perceptually separate overlapping data, the user may need to expend more resources focusing on slight changes in the data to be able to better separate the data sets from one another. Although overlapping did occur in the three dimensional perspective displays, users may have been better able to perceptually segregate the data since recent data points overlapped points which were introduced much earlier. These points which were being obscured had already been analyzed by the user and were not necessary to use to identify the presence of a signal. In the two dimensional composite displays, the overlapping of data points continuously interfered with the users ability to identify patterns. This may explain why users did not rate the multiple two dimensional and composite three dimensional perspective displays significantly different.

Temporal Demand Ratings

The temporal demand ratings of subjects was significantly higher for the composite three dimensional displays than the multiple two dimensional displays. This suggests that subjects felt significantly more stress associated with the time in completing the task when using composite three dimensional perspective displays. A possible reason for the higher rating for the composite three dimensional perspective display could be the presence of depth ambiguity in the composite three dimensional perspective display. This display contained three sets of axes corresponding to the three sets of data used for the task. However, unlike the multiple two dimensional displays, the three dimensional displays used perspective in an attempt to decrease the visual search time due to the close proximity of the axes. Therefore in order to recognize a signal, subjects may have been required to expend additional cognitive resources to properly associate a data set with the appropriate set of axes.

Observed Performance Ratings

No significantly difference was noted in the observed performance ratings between the displays. Observed performance is defined by how successful the subject feels they were in accomplishing the goals of the task or how satisfied they were with their performance. One possible explanation for this lack of significance could be that subjects were comfortable with their performance for any of the display formats. From the training program, subjects were required to identify the various signals, and were tested on their ability to recognize these signals as well as general knowledge relating to SPC. Given the fact that subjects were successful in passing the evaluation tests after each training module to a criterion of 80%, subjects may have expected similar performance on their ability to recognize signals when using any of the displays. The multiple two dimensional displays which had the best average accuracy at 68% is quite low with respect to the 80% criterion which was the expectation during training. Therefore, their expectations in performance may have been inflated as a result of the success experienced by subjects during the training portion of the experiment.

Frustration Ratings

Multiple two dimensional displays received significantly lower frustration ratings than composite three dimensional perspective displays. No significant differences were found between the multiple two dimensional displays and the composite two dimensional display, as well as the composite two dimensional displays and the composite three dimensional perspective displays. Frustration as defined by the NASA TLX measure refers to the insecurity, discouragement, or stress associated with the task while using one of the displays. An explanation for these results could be the depth ambiguity associated with the use of perspective when using the three dimensional display. Many subjects believed that the three dimensional perspective display was confusing to look which may be attributed to the depth ambiguity introduced by the perspective display. The presence of depth ambiguity may have outweighed any reductions in frustration ratings associated with the reduction in visual search time in the composite three dimensional perspective display. However, the fact that no difference was found between the multiple two dimensional display and the composite two dimensional display is somewhat surprising since the presence of continuous overlapping could create an frustrating environment for the users of the display. A possible explanation for this lack of significant difference could be the fact that any increase in ratings associated with overlapping may have been negated by a decrease in ratings associated with a reduction in visual search time.

Effort Rating

The type of display showed significant differences in effort rating between the multiple two dimensional display and either composite display. No significant difference was found between the composite displays in effort rating. The composite displays had significant higher effort ratings than the multiple two dimensional display. Effort refers to the amount of mental and physical work required to accomplish the level of performance achieved for the particular display. The presence of overlapping in both displays provide one possible explanation for the high effort rating scores for the composite displays. Since the composite three dimensional display had less overlapping within the display, it would be expected that the composite three dimensional perspective display would have a lower effort rating than the composite two dimensional display. The presence of depth ambiguity due to the use of perspective may have increased the amount of effort required by subjects to recognize the signals on the three dimensional perspective displays. This may account for the lack of significant difference between the composite displays.

Total Mental Workload

The data indicated that the type of display had a significant effect on the total mental workload ratings. Subjects rated the composite displays significantly higher than the multiple two dimensional displays in total mental workload. Based on these results it could be suggested that composite two dimensional displays should not be considered as a viable option for use in control chart decision making based on the total mental workload demands imposed by the composite formats.

However, future research still needs to be conducted before any such generalization is made regarding the use of these displays. Specifically, the moderating effect of display size could have reduced the amount of overlap, confusion, and clutter inherent in each of the composite displays. Given the fact that overlap has a negative effect on decision performance, display size which was not manipulated for this experiment should be further investigated.

One purpose of workload research is to test whether any tradeoff exists between increased workload for a specific task and decision performance. The data indicated that no significant relationship existed between workload and decision accuracy. This result is surprising considering the similar differences between means were found in total workload and decision accuracy. No significant correlations were found among the various measures of decision time and total mental workload except for decision time and total mental workload.

Subjective Measures

Decision Accuracy

Subjects rated how accurate they believed they were in using each specific display for the SPC task. Subjects felt they were significantly more accurate with the multiple two dimensional display than either composite display on their perception of accuracy of the task.

No significant difference was found between the composite displays. These results support the objective measures of decision accuracy.

Decision Confidence

The results indicated that the type of display had a significant effect on decision confidence ratings between the multiple two dimensional displays and the composite displays. No significant difference was found between the composite displays. The results showed that subjects were more confident using the multiple 2-D displays than either composite displays. This result partially supports the hypothesis that individuals using composite three dimensional perspective control charts will have less decision confidence in their evaluations than individuals using two dimensional charts. These subjective results also support the objective measures of decision accuracy in a similar fashion. The objective measure of decision accuracy also indicated significantly better accuracy on the multiple two dimensional display than either composite display. An explanation for these results could reflect the inherent characteristics of the composite displays mentioned earlier. The presence of overlap in both composite displays, and the presence of depth ambiguity in the composite three dimensional perspective display may have impacted their confidence in recognizing signals. The clutter and confusion associated with the ambiguity and the overlap may result in uncertainty in performing the cognitive operations required by the task (Wickens and Carswell, 1995).

Decision Time

Display type had a significant effect on the decision time rating of the display. Subjects indicated that the multiple two dimensional displays helped them identify signals significantly more quickly than either composite display. Again, no difference was found among the composite displays. These results contradict those results found among the objective measures of decision time where no significant differences were found among the displays. One possible reason for this contradiction could be the subjective perception in decision accuracy. Subjects were required to correctly classify a signal within five points after the original signal was introduced. Thus, a time related element was inherent built into the experiment when considering accuracy. For those displays which subjects perceived they were more accurate, they may have also perceived that the same display may have helped them make quicker decisions, since accuracy is somewhat tied to decision quickness. The relationship between decision quickness and the subjective measure of decision accuracy was found to be significant, which supports this explanation.

Spatial orientation and decision accuracy

Decision accuracy on three dimensional perspective displays could have been impacted by the three dimensional cognitive abilities of subjects. As a result, subjects were given the Cube Comparisons Test to see whether any relationship existed between the results of this test and decision accuracy of the task. The results indicated that there was no significant relationship between the three dimensional cognitive scores and decision accuracy for the three dimensional perspective displays. One possible explanation of this could be the fact that subjects were on average had little experience with three dimensional applications as found in the pre-experimental questionnaire. The lack of significance between these two variables are similar to the results found by Barfield and Robless (1989).

Implications for Work System Design and Display Design

This study was motivated by how the integration of displays could affect decision making performance as well as mental workload in a cellular manufacturing context. Changes in the way information is displayed can have an impact on decision making which can affect the effectiveness of a work system. Specifically, changes were made to the technological subsystem in the sociotechnical systems theoretical model. The results showed that the type of display had a clear effect on the human in the personnel subsystem through changes in mental workload as well as an overall effect on decision performance. Display dimensionality can be used as a viable option to create more effective work systems, however this effectiveness has been found to be dependent upon the type of task and the type of information used. This research represents a preliminary effort to provide general insight on improving work system performance as applied to the use of control charts and the decisions associated with using these types of decision tools. The results of this study were generally supported by the Proximity Compatibility Principle.

Although the implications of this study are somewhat limited by the display size, display resolution, and perspective used for this research, they do provide a reference point for future research and point out some opportunities which may ultimately help work system design for those who use control charts. One fundamental issue that truly needs to be assessed is how much automation is required in these control chart decision making tasks. Several subjects believed that the task was tedious and believed that a computer which recognized signals automatically may be better suited for the searching portion of the task. There is still a lack of research which has specifically addressed this problem. This research assumed that the human would continue to take an active role in quality control including the use of control charts as its fundamental decision tool. Therefore, little automation exists since the human operator is still involved in the recognition of the non-random signals. As mentioned earlier, research does currently exist where the computer is able to recognize non-random signals with little assistance from the human operator. Additionally, Lall and Stanislao(1992) have developed an expert system which functions to help users to select the appropriate control chart and help users to identify the possible causes of an out of control process. The implementation of expert system technology represents another level of automation, where technology is more involved in the decision making loop relating to quality control. Technology is used to assist the human operator to select the proper graph in order to recognize non-random signals and to provide assistance in taking corrective action if necessary. Finally, the use of neural networks represents another level of automation, where the human operator has a limited role in the quality control decision making loop. Rather than test the specific tools which may or may not improve overall effectiveness, the complete work system including these various technological solutions needs to be empirically investigated. Previous research can help filter which technological solutions may best represent each level of automation.

This research also explored how display complexity affected decision making performance and workload. Previous research helped to explain much of the results which indicated the relatively poor decision performance on the composite displays. System designers now have at their disposal a wide range of options when displaying data. In this case, control chart data was displayed on a computer generated two dimensional screen.

Technology though need not be limited to these types of displays. Stereoscopic images as well as virtual environments also provide the system design with viable options which may or may not improve the performance of the user. These additional alternatives provide the user with a somewhat immersive environment in which they may be able to interact with the data. However, any further investigation of these display alternatives should continue to look at not only decision performance but also mental workload. In this research, mental workload did provide additional information regarding the display alternatives which helped to support the results of decision performance.

This research required subjects to view control chart data in a variety of display formats. It was assumed that control charts were an effective way of conveying information to the decision maker. This assumption has been supported by the fact that subjects using line graphs had better decision performance than those who used tabular data. Control charts are essentially line graphs which convey bits of information through its use of lines, colors, white space, and other graphical elements. These elements of the graph were kept relatively constant throughout the experiment to avoid possible confounding with the manipulation of the display. However, changes in the bit rate within the display by manipulating the colors, the relative size of the elements of the display, etc. may affect decision performance. Christ (1990) found that the use of color could improve user performance during visual search tasks. Delucia (1995) found that in certain situations differences in the relative sizes among graph elements can result in misinterpretation of spatial relationships for perspective displays. As a result, designers should be aware of the possible impact of these design elements on decision performance.

Future research

This research tested how displays affected various dependent variables. Generally speaking, the results indicated that multiple two dimensional control charts had a distinct advantage over the composite two dimensional and composite three dimensional perspective displays. However, this research can in no way provide conclusive evidence that multiple two dimensional displays should be used for control chart decision making tasks rather than composite displays due to the display size, display resolution, and perspective used in this research.

From a technological standpoint, the manipulation of display size combined with the use of composite displays should be investigated since it appeared to be a major criticism of both composite displays. By increasing the display size and minimizing the overlapping and clutter present in such a display, more insight may be gained of the viability of composite displays in a real world environment. Increasing display size may also provide the researcher with more design options in manipulating the perspective used in the three dimensional display as well as provide additional design options with the two dimensional composite display. In a broader context, future research need not be confined to the use of two dimensional representations viewed on computer monitors. Use of stereoscopic images and virtual reality may also provide some viable alternatives to improving decision performance. Although the use of perspective was not an effective way of portraying information in this experiment, the use of stereoscopic images and virtual reality may help provide depth cues which could improve decision making performance.

This research also attempted to look at the potential effects of perspective on various decision making characteristics and workload. From past literature, the effectiveness of perspective is task dependent. The task in this research specifically looked at how control chart decision making could be affected by the use of perspective. However, the nature of the outside the control limits, runs, and trends were distinctly different from the phase related correlation. The first three non-random signals relied on the ability of the decision maker to focus attention on one process at a time. However, the phase related correlation required subjects to focus attention across processes which were temporally displaced. Based on the Proximity Compatibility Principle, decision making on phase related signals could be improved by minimizing the temporal displacement through the manipulation of perspective. Changes in perspective could reduce or even eliminate any perceptual temporal displacement by changing the line of sight which affects how the users view the data. This research did not minimize the temporal displacement since the display size constrained the manipulation of perspective. When the temporal displacement was minimized using perspective with the available display size, the resolution of the display made it nearly impossible to understand the data on the display. Future research should attempt to directly address how the manipulation of perspective can improve decision making characteristics with respect to phase related correlation understanding that display size could be a limiting factor.

Another display issue involved the possibility of providing an interactive environment to view the control chart data. This research required subjects to view the control chart data in fixed perspectives, which may have limited the interaction between the subject and the display. The limited amount of interaction between the display and the subjects may have contributed to the perceived tediousness of the searching task. By allowing subjects to manipulate the perspective and to explore the data, subjects may be more engaged in the searching task, and be less likely to feel that the type of searching task is tedious. Furthermore, with a more engaging task combined with the freedom to "fly-through" the data, decision performance and workload may be affected. Therefore, experiments dealing with a more interactive system should be explored given its potential impact on decision performance and workload.

One possible source of this tediousness could have been the fact that subjects were only responsible for recognizing non-random signals from the control charts. This research did not require subjects to simulate a manufacturing task while monitoring the state of the process. Given the importance of workload and its potential effect on decision performance, a multiple task environment which requires subjects to carry out simulated manufacturing operations while monitoring the state of the process should also be investigated. This type of environment may help reduce the tediousness of the task since it requires subjects to perform a variety of tasks which breaks up the monotony associated with only reading control chart data. In post experiment feedback, subjects generally believed that they would be comfortable executing tasks required from two manufacturing processes while considering data from only one display. In order to have a better understanding of how this impacts decision making performance in using control charts as well as task performance in carrying out the manufacturing operations, a multiple task environment experiment should be performed. Ultimately, a work system's effectiveness should be measured not only by how well operators are able to detect and act on non-random signals, but also their ability to perform their manufacturing tasks.

Mental workload may not only affect decision making performance in using control charts in a multiple task environment but may also affect task performance relating to the requirements of the manufacturing process. High levels of mental workload have been associated with decrements of performance (Wempe and Baty, 1968). This research showed that the single task of identifying non-random signals using various displays had a fair to moderately high amount of workload. Requiring operators to perform additional manufacturing tasks and to read control charts simultaneously may increase the overall level of workload such that performance is affected. Research in this area may help clear up some of the issues dealing with the human's role in the human machine loop of the manufacturing line.

By investigating function allocation in a simulated manufacturing task, a better understanding of the technological as well as personnel subsystem implications can be obtained, while gaining insight into the impact on work system effectiveness. This is important since work system effectiveness cannot be judged solely on the basis of the recognition of special causes but also the ability of operators to perform their daily manufacturing operations. A wide variety of technological solutions exist to help operators make decisions relating to processes. In order to maximize the effectiveness of the work system, research addressing function allocation through the use of computer generated control charts, expert systems, automated computer recognition systems, etc. should be performed.

Conclusions

This research attempted to explore the effects of displays on decision making and mental workload. The empirical results provided some insight on the effectiveness of integrated displays in a control chart decision making task. These results were mixed supporting the notion that three dimensional displays impose additional workload requirements upon the user. However, the hypothesized advantage of three dimensional perspective displays in terms of decision accuracy and time were not supported. This result has broadened some of the current knowledge which exists in the visualization literature which suggests that improvement as a result of display dimensionality in decision performance is rather task dependent.

Based on the results of this study, multiple two dimensional displays had a distinct advantage over composite two dimensional displays and composite three dimensional perspective displays in terms of decision accuracy, mental workload, and decision confidence. These results were supported using the principles of the Proximity Compatibility Principle which suggest that depth ambiguity and overlapping of data may have a negative impact on decision performance and workload. However, these conclusions should not be seen as a generalization, rather a starting point for additional research in the alternative displays for control chart decision making.

References

Adhire, S. (1996). TQM age versus quality: An empirical investigation. Production and Inventory Journal, 37(1), 18-22.

Baddley, A. (1994). The magical number of seven: still magic after all these years? Psychological Review, 101, 353-356.

Bailey, R. W. (1983). Human Error in Computer Systems, Englewood Cliffs, NJ: Prentice Hall.

Barfield, W., and Robless, R. (1989). The effects of two- or three-dimensional graphics on the problem-solving performance of experienced and novice decision makers. Behaviour and Information Technology, 8(5), 369-385.

Bemis, S.V., Leeds, J. L., and Winer, E.A. (1988). Operator performance as a function of type of display: Conventional versus perspective. Human Factors, 30, 163-169.

Benbasat, I., and Dexter, A. (1985). An experimental evaluation of graphical and colour-enhanced information presentation. Management Science, 31, 1348-1364.

Benbasat, I., Dexter, A., and Todd, P. (1986). An experimental program investigating colour-enhanced and graphic information presentation: An integration of the findings. Communications of the ACM, 39(11), 1094-1105.

Benbasat, I., and Schroeder, R. (1977). An experimental investigation of some MIS design variables. Management Information Systems Quarterly, 1(1), 39-50.

Borg, G., Bratfisch, O., and Doring, S. (1971). On the problem of perceived difficulty. Scandinavian Journal of Psychology, 12, 249-260.

Carswell, C. (1990). Graphical information processing: The effects of proximity compatibility. In *Proceedings of the 34th Annual meeting of the Human Factors Society* (pp. 1494-1498). Santa Monica, CA: Human Factors and Ergonomics Society.

Casali, J. G., and Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary task workload estimation techniques in simulated flight task emphasizing communications load. Human Factors, 25, 623-641.

Cheng, C.-S. (1997). A neural network approach for the analysis of control chart patterns. International Journal of Production Research, 35(3), 667-697.

Cherns, A. (1987). Principles of Sociotechnical Design Revisited. Human Relations, 40(3), 153-162.

Chervany, N., and Dickson, G. (1974). An experimental information overload in a production environment. Management Science, 20, 1335-1344.

Chih, Wen-Hai and Rollier, Dwayne. (1990). A methodology of pattern recognition schemes for two variables in SPC. International Journal of Quality & Reliability, 12(3), 86-107.

Christ, R. E. (1990). Review and analysis of color coding research for visual displays. In M. Venturino (ed.), Selected Readings in Human Factors (89-118). Santa Monica, CA: Human Factors Society.

Czaja, S., and Drury, C. (1981). Training programs for inspection. Human Factors, 23(4), 473-484.

Cochran, W. G. (1954). Some methods of strengthening the common X^2 tests. Biometrics, 10, 417-51.

Delucia, P. (1995). Effects of pictorial relative size and ground-intercept information on judgments about potential collisions in perspective displays. Human Factors, 37(3), 528-538.

Deming, W. E. (1986). Out of the Crisis. Cambridge, MA: Massachusetts Institute of Technology.

Dickson, G., DeSanctis, G., and McBride, D. (1986). Understanding the effectiveness of computer graphics for decision support: a cumulative experimental approach. Communications of the ACM, 29(1), 40-47.

Drury, C. G. and Kleiner, B. M. (1984). Comparison of Blink Aided and Manual Inspection Using Laboratory and Plant Subjects. In Proceedings of the Human Factors Society 28th Annual Meeting (pp. 670-674). Santa Monica, CA, Human Factors Society,.

Ebrahimpour, M. (1985). An examination of quality management in Japan: Implications for management in the United States. Journal of Operations management, 5(4), 419-431.

Eick, S., and Fyock, D. (1996). Visualizing corporate data. AT&T Technical Journal, 75(1) 74-86.

Ellis, S. R., McGreevy, M.W., and Hitchcock, R. J. (1987). Perspective traffic display format and airline pilot traffic avoidance. Human Factors, 29, 371-382.

Embrey, D. E. (1979). Approaches to training for industrial inspection. Applied Ergonomics, 10, 139-144.

Emery, F. E., Trist, E. L. (1965). The causal texture of organizational environments. Human Relations, 18, 21-32.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. Perception and Psychophysics, 16, 143-149.

Feigenbaum, A. V. (1983). Total Quality Control. 3rd edition. New York, New York: McGraw-Hill Book Company.

Gibson, J., Ivancevich, J., Donnelly, J. (1994). Organizations. Burr Ridge, IL: Irwin.

Gopher, D., and Donchin, E. (1986). Workload – An examination of the concept. In K. R. Boff, L. Kaufman & J.P Thomas (eds.), Handbook of perception and human performance, Volume 2. New York: John Wiley.

Gramopadhye, A., Kimbler, D., Kimbler, E., Bhagwat, S., and Rao, P. (1995). Application of advanced technology to training for visual inspection. In Proceedings of the Human Factors Society 39th Annual Meeting (pp. 1299-1303). Santa Monica, CA: Human Factors and Ergonomics Society.

Gramopadhye, A., Drury, C., and Sharit, J. (1993). Training for decision making in aircraft inspection. In Proceedings of the Human Factors Society 37th Annual Meeting (pp. 1267-1271). Santa Monica, CA: Human Factors and Ergonomics Society.

Green, D., and Swets, J. (1966). Signal detection theory and psychophysics. New York: John Wiley.

Hammer, M. (1996). Beyond reengineering. New York, NY: HarperCollins Publishers, Inc.

Hart, S. G., and Staveland, L. E. (1988). Development of a NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.S. Hancock and N. Meshkati (Eds.), Human mental workload (pp. 139-183). Amsterdam: Elsevier.

Haskell, I., and Wickens, C.D. (1993). A theoretical and empirical comparison of two- and three-dimensional displays for aviation. International Journal of Aviation Psychology, 3, 87-109.

Hendrick, H. W. (1991). Human factors in organizational design and management. Ergonomics, 34, 743-756.

Hendy, K. C., Campbell, E. L., Landry, L. N., and Hamilton, K. M. (1989). Review of workload: Measurement and prediction. In Proceedings of the 22nd Annual Conference of the Human Factors Association of Canada (pp. 31-39). Mississauga, Ontario, Canada: Human Factors Association of Canada.

Hickox, J., and Wickens, C. (1997). 3-D electronic maps, design implications for the effects of elevation angle disparity, complexity, and feature type. In Proceedings of the Human Factors Society 41st Annual Meeting (pp. 23-27). Santa Monica, CA: Human Factors and Ergonomics Society.

Hill, S. G., Zaklad, A. L., Bittner, A. C., Jr., Byers, j. C., and Christ, R. E. (1988).Workload assessment of a mobile air defense missile system. In Proceedings of the Human Factors

Society 32nd Annual Meeting (pp. 1068-1072). Santa Monica, CA: Human Factors and Ergonomics Society.

Hoyer, R.W., and Ellis, W.C. (1996). A graphical exploration of SPC. Quality Progress, 29(6), 57-64.

Juran, J. M. (1988). Juran on Planning for Quality. New York, New York: The Free Press.

Kim, W.S., Ellis, S. R., Tyler, M.E., Hannaford, B., and Stark, L.W. (1987). Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. IEEE Transactions on Systems, Man, and Cybernetics, 17(1), 61-72.

Kleiner, B. M. and Drury, C. G. (1993). Design and Evaluation of an Inspection Training Program. Applied Ergonomics, 24(2), 75-82.

Kleiner, B. M., and Hertweck, B. (1996). By which method: Total quality management, reengineering, or deengineering? Engineering Management Journal, 1.

Kleiner, B. M. (1997a). *HFES Macroergonomics Technical Group Web Site*, (http://mgdsl.ise.vt.edu/odam/), MGDSL Lab, Virginia Polytechnic Institute and State University.

Kleiner, B.M. (1997b). Sociotechnical systems theory as an organizing framework for both ergonomics and TQM. In Proceedings of the International Ergonomics Association 13th Triennial (pp. 211-210). Tempere, Finland.

Kramer, A. F., and Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. Perception and Psychophysics, 50, 267-284.

Lall, V., and Stanislao, J. (1992). Applying a coupled expert system to quality control charts. Computers and Industrial Engineering, 23, 401-404.

Lee, J., and MacLachian, J. (1986). the effects of 3D imagery on managerial data interpretation. Management Information Systems Quarterly, 10(3), 256-269.

Liu, Y., and Wickens, C. (1992). Use of computer graphics and cluster analysis in aiding relational judgement. Human Factors, 34, 165-178.

Lucas, H. (1981). An experimental investigation of the use of computer-based graphics in decision making. Management Science, 27(7), 757-768.

Lucas, H., and Nielsen, N. (1980). The impact of the model of information presentation on learning and performance. Management Science, 26(10), 982-993.

Macmillan, N. A., and Creelman, C. D. (1991). Detection theory: A user's guide. Cambridge University Press, New York, New York.

Maxwell, A. E. (1961). Analysing Qualitative Data. John Wiley and Sons, Inc. New York, New York.

McGreevy, M.W., and Ellis, S. R. (1986). The effect of perspective geometry on judged direction in spatial information instruments. Human Factors, 28, 439-456.

Micalizzi, J., and Goldberg, J. H. (1989). Knowledge of results in visual inspection decision: sensitivity or criterion effect? International Journal of Industrial Ergonomics, 4, 225-235.

Miller, G. A. (1956). The magical number seven, plus or minus two. Psychological Review, 63, 81-97.

Mintzberg, H. (1975). Impediments to the use of management information. New York, NY: National Association of Accountants.

Modarress, B., and Ansari, A. (1989). Quality control techniques in U.S. firms: A survey. Production and Inventory Management Journal, 30(2), 58-62.

Morgan, B., Hershcler, D., Wiener, E., and Salas, E. (1993). Implications of automation technology for aircrew coordination and team performance. Advances in Man-Machine Systems Research, 6.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement technique: Implications for their use in the assessment of perceived mental workload. Human Factors, 33(1), 17-33.

O'Brien, J., and Wickens, C. (1997). Free flight cockpit displays of traffic and weather: Effects of dimensionality and data base integration. In Proceedings of the Human Factors Society 41st Annual Meeting (pp. 18-22). Santa Monica, CA: Human Factors and Ergonomics Society.

O'Reilly, C.A. (1980). Individuals and information overload in organizations: Is more necessarily better? Academy Management Journal, 23, 684-695.

Pasmore, W. (1988). Designing effective organizations. New York, NY: John Wiley & Sons, Inc.

Perrow, C. (1967). A framework for the comparative analysis of organizations. American Sociological Review , 32, 194-208.

Petzet, A. (1996). Communicating among disciplines. (multidisciplinary communication in the petroleum industry can yield economic benefits in decision-making). The Oil and Gas Journal, 94(21), 21.

Pham, D.T. and Oztemel, E. (1992) Control chart pattern recognition using neural networks. Journal of System Engineering, 2, 256-262.

Pham, D.T. and Oztemel, E. (1994) Control chart pattern recognition using learning vector quantization networks. International Journal of Production Research, 32(3), 721-729.

Powers, M., Lashley, C., Sanchez, P., and Shneiderman, B. (1984). An experimental comparison of tabular and graphic data representations. International Journal of Man-Machine Studies, 20(6), 565-566.

Reid, G. B., and Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A scaling procedure for measuring mental workload. In P. A. Hancock and N. Meshkati (Eds.), Human mental workload (pp. 185-218). Amsterdam: Elsevier.

Salvendy, G., and Seymour, D. W. (1973). Prediction and development of industrial work performance. Canada: John Wiley and Sons.

Sanders, A. T. (1977). Some remarks on Mental Load. In Mental Workload, ed. Neville Moray. New York, New York, Plenum Press.

Schoonard, J. W., and Gould, J. D. (1973). Studies of visual inspection. Ergonomics, 16, 365-379.

Schroeder, R., Sakakibara, S., Flynn, E., and Flynn, B. (1992). Japanese plants in U.S.: How good are they? Business Horizons, 35(4), 66-72.

Sheridan, T., and Stassen H. (1977). Definitions, Models and measures of human workload. In Mental Workload, ed. Neville Moray. New York, New York, Plenum Press.

Shewhart, W. A. (1931). Economic control of quality of manufacturing product. New York, NY: D. Van Nostrand.

Shively, R. (1986). Application of Mental Workload Methodology to Human-Computer Interaction. In Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics, 907-911.

Simon, H. A. (1968). The future of information processing technology. Management Science, 14(5), 619-624.

Simon, H. A. (1990). In Y. Ijiri & S. Sunders (Eds.), Information technologies and organizations. The Accounting Review, 65(3), 658-667.

Taveira, A., and Smith, M. (1996). The design of an interpretive inquiry for quality management assessment; Procedures for developing a grounded theory. Human Factors in Organizational Design and Management – V, 505-510.

Taylor, J., Felten, D. (1992). Performance by design. New Jersey: Prentice Hall.

Thapa, V., Gramopadhye, A, Melloy, B., and Grimes, L. (1996). Evaluation of different training strategies to improve decision-making performance in inspection. International Journal of Human Factors in Manufacturing, 6(3), 243-261.

Van der Heijden, A. (1992). Selective attention in vision. New York: Routledge.

Vandenberg, S., and Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. Perceptual and Motor Skills, 47, 599-604.;

Wallack, P. M., and Adams, S. K. (1969). The utility of signal detection theory in the analysis of industrial inspection accuracy. AIIE Transactions, 1, 33-44.

Wapole, R., and Myers, R. (1993). Probability and Statistics for Engineers and Scientists. MacMillan Publishing Company, New York, NY, 1993.

Watson, C., and Driver, R. (1983). The influence of computer graphics on the recall of information, Management Information Systems Quarterly, 7(1), 45-53.

Wempe, T. E., and Baty, D. L. (1968). Human information processing rates during certain multiaxis tracking tasks with a concurrent auditory task. IEEE Transactions on Man-Machine Systems, 9, 129-138.

Wheeler, D. J. (1993). Understanding variation: The key to managing chaos. Knoxville, TN: SPC Press

Wickens, C. D., and Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. Human Factors, 37(3), 473-494.

Wickens, C. D., and Todd, S. (1990). Three-dimensional display technology for aerospace and visualization. In Proceedings of the Human Factors Society 34th Annual Meeting (1479-1483). Santa Monica, CA: Human Factors and Ergonomics Society.

Wickens, C. D. (1992). Virtual reality and education. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (842-847). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Wickens, C. D., Merwin, D. H., and Lin, E. L. (1994). Implications of graphics enhancements for the visualization of scientific data: Dimensional integrality, stereopsis, motion, and mesh. Human Factors, 36(1), 44-61.

Wiener, E. L. (1963). Knowledge of results and signal rate in monitoring: A transfer of training, approach. Journal of Applied Psychology, 45, 214-222.

Wierwille, W. W. and Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. In Proceedings of the Human Factors Society 20th Annual Meeting, (pp. 129-133). Santa Monica, CA: Human Factors and Ergonomics Society.

Yeh, Y-Y, and Silverstein, L.D. (1992). Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. Human Factors, 34, 583-600.

Zhao, B., and Salvendy, G. (1996). Compatibility of task presentation and task structure in human-computer interaction. Perceptual and Motor Skills, 83(1), 163-175.

Appendix A1 Decision Making Performance ANOVA Tables

ANOVA Summary Table for Decision Accuracy by Display Type

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	0.608941	0.055358		
Within-Subject					
Display (D)	2	0.336806	0.168403	8.54	0.002
D x S	22	0.434028	0.019729		
Total	35	1.379774			

ANOVA Summary Table for Decision Accuracy by Signal Type

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	0.811921	0.073811		
Within-Subject					
Signal (L)	3	2.029514	0.676505	17.79	0.000
L x S	33	1.255208	0.038037		
Total	47	4.096644			

ANOVA Summary Table for Search Time

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	67.9559	6.1778		
Within-Subject					
Display (D)	2	2.7126	1.3563	0.13	0.877
D x S	22	226.0365	10.2744		
Total	35	296.7050			

ANOVA Summary Table for Decision Time

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	7.45308	0.67755		
Within-Subject					
Display (D)	2	1.12659	0.56329	0.64	0.536
D x S	22	19.30010	0.87728		
Total	35	27.87976			

ANOVA Summary Table for Search + Decision Time

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	71.4964	6.4997		
Within-Subject					
Display (D)	2	7.6746	3.8373	0.33	0.724
D x S	22	258.1307	11.7332		
Total	35	337.3018			

ANOVA Summary Table for Stopping Time

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	54.358	27.179		
Within-Subject					
Display (D)	2	257.754	23.432	1.84	0.182
D x S	22	324.878	14.767		
Total	35	636.990			

ANOVA Summary Table for Probability of a Hit (Hit Rate)

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	0.655496	0.059591		
Within-Subject					
Display (D)	2	0.026999	0.013499	0.74	0.487
D x S	22	0.398991	0.018136		
Total	35	1.081485			

ANOVA Summary Table for Probability of a False Alarm (False Alarm Rate)

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	0.0693377	0.0063034		
Within-Subject					
Display (D)	2	0.0116850	0.0058425	1.63	0.219
D x S	22	0.0789534	0.0035888		
Total	35	0.1599762			

ANOVA Summary Table for Probability of Sensitivity (d')

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	15.2502	1.3864		
Within-Subject					
Display (D)	2	0.3948	0.1974	0.21	0.812
D x S	22	20.6090	0.9368		
Total	35	36.2540			

ANOVA Summary Table for Probability of Response Criterion (b)

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	5.94770	0.54070		
Within-Subject					
Display (D)	2	3.54621	1.77310	1.82	0.185
D x S	22	21.37966	0.97180		
Total	35	30.87356			

Appendix A2 NASA Task Load Index ANOVA Tables

ANOVA Summary Table for Mental Demand Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	15022.97	1365.72		
Within-Subject					
Display (D)	2	2917.06	1458.53	7.89	0.003
D x S	22	4065.61	184.80		
Total	35	22005.64			

ANOVA Summary Table for Physical Demand Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	23657.89	2150.72		
Within-Subject					
Display (D)	2	300.72	150.36	5.08	0.015
D x S	22	650.61	29.57		
Total	35	24609.22			

ANOVA Summary Table for Temporal Demand Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	4790.67	435.52		
Within-Subject					
Display (D)	2	1152.17	576.08	3.85	0.037
D x S	22	3289.17	149.51		
Total	35	9232.00			

ANOVA Summary Table for Observed Performance Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	5244.08	476.73		
Within-Subject					
Display (D)	2	275.17	137.58	1.03	0.375
D x S	22	2951.50	134.16		
Total	35	8470.75			

ANOVA Summary Table for Frustration Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	12931.33	1175.58		
Within-Subject					
Display (D)	2	446.00	223.00	4.62	0.021
D x S	22	1062.67	48.30		
Total	35	14440.00			

ANOVA Summary Table for Effort Ratings

Source	df	SS	MS	F test statistic	p-value
Between-Subjects					
Subjects (S)	11	4790.22	435.47		
Within-Subject					
Display (D)	2	1184.39	592.19	4.29	0.027
D x S	22	3040.28	138.19		
Total	35	9014.89			

ANOVA Summary Table for Total Mental Workload Ratings

Source	**df**	**SS**	**MS**	**F test statistic**	**p-value**
Between-Subjects					
Subjects (S)	11	16344.7	1485.9		
Within-Subject					
Display (D)	2	17860.7	8930.3	9.42	0.001
D x S	22	20852.7	947.8		
Total	35	55058.0			

Appendix B1 Background Information Sheet

Background

You are a recent graduate from an industrial engineering program, working at NFX Corporation, a steel parts manufacturing company. Since your career began at NFX, approximately six months ago, you have taken the role of quality assurance engineer. NFX has been a growing corporation over the last fifteen years. Over the last five years, NFX has slowly implemented Total Quality Management initiatives into its manufacturing facilities. Recently, the company has mandated that control charts be used on the manufacturing floor. Presently, NFX Corporation, owns three plants which mainly produce drive shafts for different automotive customers. However, during the past year there has been a big push to improve quality while cutting costs due to increasing competition in the automotive parts industry. As a result, the Vice President of Operations has made you responsible for identifying problems associated with their manufacturing processes at each of these facilities.

In order to assess these problems, control charts have been used at the various plants for nearly five years. However, each plant has developed their own format for displaying control chart information to help workers and management have a better understanding of their manufacturing processes and capabilities. NFX has also invested heavily into a new information technology system at NFX. NFX now has an online measurement system which displays control chart data from each manufacturing facility in real time. Your job as quality assurance engineer is to provide a log of the process using this online measurement system.

Appendix B2 Paramus Facility Information Sheet

Paramus Manufacturing Facility

The Paramus Manufacturing facility produces drive shafts for small cars for various automotive customers. The basic operations for manufacturing these shafts in order are as follows:

- Rough turning operation (within ± .005" and with a surface roughness of 200m IN.)
- Finish turning operation (within ± .001" and with a surface roughness of 70m IN.)
- Grinding operation (within ± .0005" and with a surface roughness of 35m IN.)

The machining processes required to make the shaft always occur in the order mentioned above. Between the machining operations there is a queue of two shafts (lag = 2 data points). Assume that any phase related correlation can only occur between consecutive processes (Rough Turning + Finish Turning or Finish Turning + Grinding).
These three operations are necessary in order to meet the customer specifications. Raw material for these shafts is provided by another NFX site. Based on previous estimates conducted by engineers at NFX, the following cost/benefit structure exists at the manufacturing facility:

- $5,000 is the sale value of the shafts per day by letting the good shafts go through the process
- $7,500 is the savings per day associated with catching the process when it is affected by a special cause, and making appropriate process changes (modifying the machining parameters)
- $11,000 is the cost per day associated with shutting down the process and bringing it back online without changing the process
- $25,000 is the cost per day associated with producing bad shafts when the system is subject to a special cause

As the quality assurance engineer, you must provide a log of out of control events of the machining processes, keeping in mind that you want to maximize profit for the facility each day. In order to do this, control chart data will be displayed on the computer screen for you to make informed decisions. The control chart limits have been set at the ± 3s range. Control chart data will be displayed in simulated real time (5 real seconds = 9.5 simulated minutes), where approximately 50 readings are taken per day. After you have recognized a signal which may be due to a special cause, you must stop the process using the button below the control chart display. Following the stoppage of the process, you will be prompted to identify the signal which suggests a special cause. You must stop the process within five data points of where the signal of special cause occurred in order to save the company money. If you should recognize a signal after five points have elapsed, you should still stop the process so the company can record this information into its log. Once you have recognized and identified the signal, the control chart display will reset and measurements will continue until sufficient data has been gathered. The control chart data will indicate the measurements taken from the diameter after each machining operation.

Press the button to the right to being monitoring control chart data.

Appendix B3 Homdel Facility Information Sheet

Homdel Manufacturing Facility

The Homdel Manufacturing facility produces drive shafts for light pickup trucks for various automotive customers. The basic operations for manufacturing these shafts in order are as follows:

- Rough turning operation (within ± .005" and with a surface roughness of 200m IN.)
- Finish turning operation (within ± .001" and with a surface roughness of 70m IN.)
- Grinding operation (within ± .0005" and with a surface roughness of 35m IN.)

The machining processes required to make the shaft always occur in the order mentioned above. Between the machining operations there is a queue of two shafts (lag = 2 data points). Assume that any phase related correlation can only occur between consecutive processes (Rough Turning + Finish Turning or Finish Turning + Grinding). These three operations are necessary in order to meet the customer specifications. Raw material for these shafts is provided by another NFX site. Based on previous estimates conducted by engineers at NFX, the following cost/benefit structure exists at the manufacturing facility:

- $5,000 is the sale value of the shafts per day by letting the good shafts go through the process
- $7,500 is the savings per day associated with catching the process when it is affected by a special cause, and making appropriate process changes (modifying the machining parameters)
- $11,000 is the cost per day associated with shutting down the process and bringing it back online without changing the process
- $25,000 is the cost per day associated with producing bad shafts when the system is subject to a special cause

As the quality assurance engineer, you must provide a log of out of control events of the machining processes, keeping in mind that you want to maximize profit for the facility each day. In order to do this, control chart data will be displayed on the computer screen for you to make informed decisions. The control chart limits have been set at the ± 3s range. Control chart data will be displayed in simulated real time (5 real seconds = 9.5 simulated minutes), where approximately 50 readings are taken per day. After you have recognized a signal which may be due to a special cause, you must stop the process using the button below the control chart display. Following the stoppage of the process, you will be prompted to identify the signal which suggests a special cause. You must stop the process within five data points of where the signal of special cause occurred in order to save the company money. If you should recognize a signal after five points have elapsed, you should still stop the process so the company can record this information into its log. Once you have recognized and identified the signal, the control chart display will reset and measurements will continue until sufficient data has been gathered. The control chart data will indicate the measurements taken from the diameter after each machining operation.

Press the button to the right to being monitoring control chart data.

Appendix B4 Harrison Facility Information Sheet

Harrison Manufacturing Facility

The Harrison Manufacturing facility produces drive shafts for Sport Utility Vehicles for various automotive customers. The basic operations for manufacturing these shafts in order are as follows:

- Rough turning operation (within ± .005" and with a surface roughness of 200m IN.)
- Finish turning operation (within ± .001" and with a surface roughness of 70m IN.)
- Grinding operation (within ± .0005" and with a surface roughness of 35m IN.)

The machining processes required to make the shaft always occur in the order mentioned above. Between the machining operations there is a queue of two shafts (lag = 2 data points). Assume that any phase related correlation can only occur between consecutive processes (Rough Turning + Finish Turning or Finish Turning + Grinding). These three operations are necessary in order to meet the customer specifications. Raw material for these shafts is provided by another NFX site. Based on previous estimates conducted by engineers at NFX, the following cost/benefit structure exists at the manufacturing facility:

- $5,000 is the sale value of the shafts per day by letting the good shafts go through the process
- $7,500 is the savings per day associated with catching the process when it is affected by a special cause, and making appropriate process changes (modifying the machining parameters)
- $11,000 is the cost per day associated with shutting down the process and bringing it back online without changing the process
- $25,000 is the cost per day associated with producing bad shafts when the system is subject to a special cause

As the quality assurance engineer, you must provide a log of out of control events of the machining processes, keeping in mind that you want to maximize profit for the facility each day. In order to do this, control chart data will be displayed on the computer screen for you to make informed decisions. The control chart limits have been set at the ± 3 s range. Control chart data will be displayed in simulated real time (5 real seconds = 9.5 simulated minutes), where approximately 50 readings are taken per day. After you have recognized a signal which may be due to a special cause, you must stop the process using the button below the control chart display. Following the stoppage of the process, you will be prompted to identify the signal which suggests a special cause. You must stop the process within five data points of where the signal of special cause occurred in order to save the company money. If you should recognize a signal after five points have elapsed, you should still stop the process so the company can record this information into its log. Once you have recognized and identified the signal, the control chart display will reset and measurements will continue until sufficient data has been gathered. The control chart data will indicate the measurements taken from the diameter after each machining operation.

Press the button to the right to being monitoring control chart data.

Appendix B5 Out of Control Signal Screen

Report Log

Based on the data you just viewed, please indicate the type of signal which could indicate a special cause.

- ☐ Outside the Control Limits
- ☐ Phase Related Correlation
- ☐ Run
- ☐ Trend
- ☐ End of day (Control chart display went blank, only saw random variation for the day)

Appendix B6 Post Treatment Questionnaire

Post-Treatment Questionnaire

1. Using this display, I was confident in my decisions.

☐ Strongly Disagree ☐ Disagree ☐ Undecided ☐ Agree ☐ Strongly Agree

2. For this display, I believe I recognized signals quickly after they appeared.

☐ Strongly Disagree ☐ Disagree ☐ Undecided ☐ Agree ☐ Strongly Agree

3. I believe that the display used for this task helped me make accurate decisions.

☐ Strongly Disagree ☐ Disagree ☐ Undecided ☐ Agree ☐ Strongly Agree

4. For this display, I had the most problems with identifying the following signal:

☐ Outside the Control Limit ☐ Run ☐ Trend ☐ Phase Related Correlation

5. The reason I had the most problems with the signal indicated in Question 4 is:

6. For this task, I had the least problems with identifying the following signal:

☐ Outside the Control Limit ☐ Run ☐ Trend ☐ Phase Related Correlation

7. The reason I had the least problems with the signal indicated in Question 6 is:

8. Please provide any additional comments in the lines below.

Appendix C1 Weighted Workload Dimension Page

Workload Questionnaire

Instructions: Listed below are a number of terms which can be used to describe various dimensions of workload. Please select one of the dimensions from each pair to indicate which dimension contributed most to the variation in workload for the task you just completed. A definition of each term is provided at the bottom of the page.

1.	☐ Mental Demand	☐ Physical Demand
2.	☐ Temporal Demand	☐ Physical Demand
3.	☐ Temporal Demand	☐ Frustration Level
4.	☐ Temporal Demand	☐ Mental Demand
5.	☐ Performance	☐ Physical Demand
6.	☐ Temporal Demand	☐ Effort
7.	☐ Performance	☐ Mental Demand
8.	☐ Frustration	☐ Physical Demand
9.	☐ Performance	☐ Frustration
10.	☐ Frustration	☐ Mental Demand
11.	☐ Effort	☐ Physical Demand
12.	☐ Performance	☐ Effort
13.	☐ Effort	☐ Mental Demand
14.	☐ Temporal Demand	☐ Performance
15.	☐ Effort	☐ Frustration

Definition of Terms:

Mental Demand: How much thinking, deciding, remembering, looking, searching, etc. was required the task? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical Demand: How much eye movement, pushing, pulling turning, controlling, etc. was required by the task? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand: How much pressure did you feel associated with the pace or rate at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance: How successful do you think you were in accomplishing the goals of the task? How satisfied were you with your performance in achieving these goals?

Frustration Level: How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content relaxed and complacent did you feel during the task?

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Appendix C2 Post Treatment Workload Page

Questionnaire

Part I.

Instructions:
Using a scale from 0-20, please rate the magnitude of each term for the task you just completed (definitions of each term can be found at the bottom of the page):

Rating Scale for the Completed Task Rating Score

Low Mental Demand 0 <--------> 20 High Mental Demand

Low Physical Demand 0 <--------> 20 High Physical Demand

Low Temporal Demand 0 <--------> 20 High Temporal Demand

Low Perceived Performance 0 <--------> 20 High Perceived Performance Demand

Low Amount of Frustration 0 <--------> 20 High Amount of Frustration Demand

Low Amount of Effort Required 0 <--------> 20 High Amount of Effort Required

Please verify your answers are within the range of 0-20 and click submit to proceed:

Definition of Terms:

Mental Demand: How much thinking, deciding, remembering, looking, searching, etc. was required the task? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical Demand: How much eye movement, pushing, pulling turning, controlling, etc. was required by the task? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand: How much pressure did you feel associated with the pace or rate at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance: How successful do you think you were in accomplishing the goals of the task? How satisfied were you with your performance in achieving these goals?

Frustration Level: How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content relaxed and complacent did you feel during the task?

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Appendix C3 Baseline Task Instructions

Task Instructions:

Assume you are a quality assurance engineer responsible for two machines at a major steel making corporation. You are responsible for monitoring control chart displays for these two machines. Control chart data is displayed in real time. Your job is to alert management of any special causes which may appear in the control charts for either machine. However, it is important that you do not alert management when there is no special cause. In order to alert management of a special cause, you must click on the "Stop Process" button which will appear below the control charts.

Click here to begin the task.

Appendix D1 Informed Consent Form

Informed Consent for Participants of Investigative Projects

Title of Project:	Use of Integrated Displays in Work System Design
Research Investigators:	Somchart Thepvongs (Industrial and Systems Engineering graduate student) Dr. Brian M. Kleiner (Faculty Advisor and Industrial and Systems Engineering Associate Professor)

I. The Purpose of this Research

The purpose of this research is to investigate how different displays can effect the decision making process of individuals and the amount mental effort required to use these displays. The data obtained from this research will help provide recommendations for the design of displays to help support decision making.

II. Procedures

You will be required to undergo a training program which provides fundamental knowledge dealing with statistical process control. You will be tested on this knowledge following each section of the training program. The testing will require that all sections be completed and that the subject answer 80% of the questions correctly for each section. You will be given a spatial ability test (Cube Comparison Test). This pencil and paper test requires you to match cubes which have similar characteristics. Following this test, you will view some preliminary control chart data, and identify the existence of statistical process control patterns in the displayed data. Then, you will be given a manufacturing task to complete using different displays to identify various control chart patterns. This experiment is expected to take approximately two hours.

III. Risks

There are minimal risks associated with participation in this research study

IV. Benefits of this Project

Subjects will be given monetary compensation for their participation in the experiment. Subjects may also benefit from the knowledge learned from the experiment through their training in basic statistical process control topics. However, the experimenter does not guarantee any benefits from the statistical process control training.

This research will provide empirical evidence to support recommendations in work system design when using control charts to make decisions.

V. Extent of Anonymity and Confidentiality

Since this research involves individuals, anonymity and confidentiality must be considered. During all times throughout the experiment, anonymity with be maintained through the codification of individuals. Information regarding the participating subjects will only be known to the researcher, and only be used for research purposes.

VI. Compensation

Subjects will be compensated based on the phase of the experiment completed. The phases of the experiment include the following: 1. pre-experiment 2. training 3. experiment 4. post-experiment. For each completed phase, subjects will receive $2.50, totaling $10.00 for the completion of all four phases. Subjects will be paid following the completion of the experiment or subject withdrawal from the experiment.

VII. Freedom to Withdraw

You are free to withdraw from this study at any time without penalty and will be paid for the phase competed.

VIII. Approval of Research

This research project has been approved, as required by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

IX. Subject's Responsibilities

By signing this informed consent form, you are agreeing to participate in this study. As a subject I agree to do the following:

- I agree to follow the procedures of the experiment to the best of my ability.
- I also agree to not discuss any aspect of the research with others, after the conclusion of the research.

X. Subject's Permission

I have read and understood all points and conditions explained on the Informed Consent Form. I hereby acknowledge the above points and conditions and give my voluntary consent for participation in this project. I understand that I am free to withdraw at any time without penalty.

________________________________ ______________________________
Signature Date

Should I have any questions about this research or its conduct, I may contact:

Somchart Thepvongs (Research Investigator)	540-953-0216	sthepvon@vt.edu
Dr. Brian M. Kleiner (Faculty Advisor)	540-231-4926	bkleiner@vt.edu
Mr. H. T. Hurd (Chair, Institutional Review Board)	540-231-5281	hurd@vt.edu

Appendix D2 Protocol for IRB Request

Experimental Protocol

1. Justification of Research Project

Purpose of the Research Project

The purpose of this empirical research project is to investigate the influence of the level of integration of control chart displays on work system decision making (decisional variables) and individual mental workload. The level of integration deals with how displays are arranged and viewed by a user. This arrangement could include separate displays for multiple graphs, or a single graph which integrates information from multiple graphs. The graphs which will be used for this research will be control charts. Control charts are line graphs used by those in industry to better understand variation associated with characteristics of a process such as machining time, drilling diameter, servicing time, etc. These control charts are used by people to help guide decisions regarding a work system. A work system is comprised of several components: suppliers, inputs, processes, outputs, and customers. As decision makers, people attempt to make decisions using control charts to meet the goals and objectives of the work system.
Computer displays that will be tested will include separate two-dimensional displays of control chart data, a composite two-dimensional display of control chart data, and a composite three-dimensional perspective display of control chart data. All subjects will use each of the displays at some point throughout the experiment. This research will measure the time required to make decisions, the accuracy of the decisions made, and the confidence of the decisions made based on the display methodology used (separate two-dimensional displays, a composite two-dimensional display, and a composite three-dimensional perspective display of control chart data). For this research, accuracy of the decision is defined by whether a subject correctly recognizes a pattern of data on the control chart, and correctly identifies the pattern. Decision confidence is a subjective measure of how certain a subject feels regarding the correctness of the identified signal. All experimental tasks consist of a simulated manufacturing environment in which subjects are required to look at control chart data displayed in one of the three display formats. In this simulated manufacturing environment, a subject plays the role of a quality assurance engineer. As the quality assurance engineer, the subject is given control chart data. Based on the information displayed from the computer screen, subjects are required evaluate the information and make a decision based on the state of the information displayed. Subjects will indicate their decision from a list of problems only when they believe the system indicates special cause variation. A subjective questionnaire is administered to evaluate mental workload, the amount of mental resources required to complete a specific task. A subjective questionnaire will also be used to evaluate the decision confidence for the display format used.

Contribution of the Findings

Findings from this research may provide evidence to support the benefits of integrating control chart data into a specific display format. An understanding of how integration of control chart data may impact decision accuracy, decision time, decision confidence, and mental workload may be gained. As a result, this research uniquely contributes to the quality control body of knowledge by looking at the impact of various displays on these variables. Furthermore, it is an extension of the research conducted in the integrated displays or visualization domain which has seen somewhat mixed results based on the task environment. This research utilizes integrated displays in a control chart decision making environment which has not yet been looked at extensively by researchers. Based on previous research it is hypothesized that three-dimensional perspective displays will result in better decision accuracy, more mental workload, more decision time, and less decision confidence. Combining the areas from the various domains of research will help provide empirical evidence to support work system design recommendations for those using control charts for decision making.

2. Experimental Procedures

Subject Population

Twelve subjects from the Industrial and Systems Engineering graduate and undergraduate program will participate in this experiment. Subjects will be recruited from various Industrial and Systems Engineering graduate and undergraduate level classes. Subjects will participate by volunteering for the study. Each subject will attend one experimental session which will last approximately two hours. They will be given freedom to withdraw at any time throughout the experiment and will be compensated by the experimental phase completed. The phases include 1. Pre-Experiment 2. Training 3. Experiment 4. Post-Experiment. For each completed phase, subjects will be paid $2.50, totaling $10.00 for completing all four phases. Payment will be made following the completion of the experiment or following subject withdrawal from the experiment. Only subjects who have no or little knowledge in statistical quality control will be allowed to participate due to the nature of the experimental task.

Experimental Task

The experiment used in this research has subjects utilizing control chart displays in order to make decisions regarding a manufacturing process. Subjects will use three different control chart displays: multiple two-dimensional control charts, a composite two-dimensional control chart, and a composite three-dimensional perspective control chart. Subjects will use these displays one at a time. Subjects are asked to play the role of quality assurance engineer for which they must evaluate and categorize the state of the manufacturing process. Their evaluation and categorization will be based on data which is being portrayed dynamically on a computer monitor. The rate at which new points on the control chart are displayed will remain fixed. Once a certain pattern is recognized, the subject is instructed to press a button below the control chart display. Decision time will be measured from the time between the start of an "out of control" signal is presented to the time the button below the display is pressed. This time is recorded by the computer. After pressing this button, the control chart data disappears, and subjects are required to indicate which type of "out of control" pattern was present in the data. The four types of patterns to be recognized by subjects include: 1. outside the control limits 2. runs 3. trends 4. phase related correlation. Subjects will be given eight trials for each display, for which two of each type of pattern will be displayed. Decision accuracy as a raw score for each display type will be measured as the number of correctly identified patterns. The computer will only record the answer provided by the subject for each observation. The order in which the patterns are presented will be random. The order in which subjects view the different displays are counterbalanced, and will be assigned randomly to subjects until all presentation orders have been completed. The patterns for each display will be randomly drawn from a pool of data sets which have been created. This data set contains four sets of each type of "out of control" pattern. For the outside the control limit and the run pattern, the location of the signal with respect to the control limits will be controlled. For the trend, the incline or decline of the best fit line of the signal will be controlled. For the cross correlated patterns, the correlation between each serial process sets will be kept constant. Furthermore, similar colors, and the display size of each chart will be kept constant.

Experimental Protocol

As mentioned earlier, the subject is only required to participate in one experimental session, lasting approximately two hours. Subjects will arrive at the Macroergonomics and Group Decision Systems Laboratory, located in 567 Whittemore Hall. Subjects will be asked to read and sign an informed consent form (see attached Informed Consent Form). After completing the form, subjects will fill out a pre-experiment questionnaire made up of questions relating to computer experience, statistical process control knowledge, gender, and academic standing. This questionnaire will be administered using a computer. Following the completion of the questionnaire, subjects will be asked to complete a spatial ability test (Cube Comparisons test) using pencil and paper. Following this test, subjects will undergo a training procedure which deals with statistical process control knowledge, using Microsoft PowerPoint slides. Questionnaires will be administered to test their proficiency level in the general statistical process control knowledge. Subjects who do not meet a criterion of 80% for any questionnaire will be asked to repeat the slide sections dealing with that subject matter. The training procedure will be self-paced.

Following the completion of the training questionnaires, subjects will look at a pair of two-dimensional control charts containing control chart data on a computer screen. They will be asked to evaluate whether an "out of control" signal exists on either graph and identify the signal. This pre-experiment task will be the same for all subjects to provide baseline measurements for mental workload. Subjects will be asked to fill out a questionnaire regarding the level of mental workload experienced in reading the pair of two-dimensional control charts. This questionnaire will be based upon the NASA Task Load Index developed to assess mental workload. In this questionnaire, they will be asked to weigh various dimensions of workload as defined by this workload assessment index.

After completing this questionnaire, subjects will be given a set of instructions on the computer which provide background information for a work system decision making task

(see attached Background). Depending on the display presentation order for the subject, subjects will read one of three information sheets on the computer (see attached Paramus, Homdel, and Harrison Manufacturing Facility sheets), corresponding to a specific display format. After reading the information, the experimenter will answer any questions which the subject may have before starting the task. Subjects will be given an experimental treatment block corresponding to one of the three display formats (multiple two-dimensional control charts, composite two-dimensional control chart, and composite three-dimensional perspective control chart) in which they must indicate the presence of "out of control" patterns on the display and identify which pattern was recognized. Decision time and the answer for each trial will be recorded automatically on the computer. Following the presentation of eight "out of control" events, subjects will be asked to complete a post-treatment questionnaire to assess subject decision confidence and difficulty associated with the decision making task. This questionnaire will also include additional subjective ratings from the NASA Task Load Index to assess mental workload. The answers to these questionnaires will be collected using the computer. Then subjects will be given an optional two minute break before the next treatment block.
Following the completion of the first treatment block, the subject will view one of the two remaining displays using the exact same procedure mentioned above and the appropriate information sheet. Following the second treatment, the remaining treatment block will be administered using the same procedure. After all three treatment blocks have been completed, a post experiment questionnaire will be given on the computer. This questionnaire will be used to assess subject preferences regarding the display and the overall task. Once this questionnaire has been completed, subjects will be paid for their participation and dismissed from the experiment.

3. Risks and Benefits

There are no more risks associated with this research than with general computer use. Subjects will be given rest breaks during the experimental session to alleviate any general discomfort associated with general computer use for this research. Furthermore, the experimenter will be present at all times to monitor any subject discomfort, and will ask the subject to stop or to withdraw from the experiment if they feel the discomfort is too great.
Subjects may receive some benefits associated with participation in this research project. Subjects may benefit from educational training received during the experiment as well as compensation for this research. The direct benefits which <u>may</u> be received by the subject for their participation in this research include the following:
Monetary compensation for their participation, depending upon the phase of experiment completed
Training in fundamental topics in statistical process control
The monetary compensation will be guaranteed by the experimenter and will be payable upon completion of the experiment or subject withdrawal from the experiment. The experimenter does not guarantee benefits associated with the training program.

Participation by subjects will allow the experimenter to make future work system design recommendations associated with control chart decision making. These design recommendations may have an impact on decision time, decision accuracy, decision confidence, and user mental workload.

4. Confidentiality/Anonymity

The data gathered in this experiment will be kept confidential. Before subjects begin the experiment, they will be assigned a subject number in order to keep track of the data. However, no linkages or relationships will be made between the subject number and the identity of the subject to maintain anonymity at all times.

5. Informed Consent
Please see the attached sheets labeled Informed Consent for Participants of Investigative Projects.

6. Biographical Sketches

Brian M. Kleiner
Dr. Brian M. Kleiner is an associate professor, and the director of the Macroergonomics and Group Decision Systems Laboratory in the Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University (Virginia Tech).

Dr. Kleiner's publications in professional journals include articles on inspection, human factors in manufacturing, sociotechnical systems, job design, process control, human performance, and TQM. These publications have been used to provide theory and applications necessary to support this research effort. The articles relating to inspection, human factors in manufacturing, and human performance have been used to provide a model and metrics used to explain decision making for this research. Articles dealing with sociotechnical systems and job design provide the guiding theory to justify this research effort and to support changes in job design. Process control and TQM provide the theory necessary to support the use of control charts used in this research.

Education:
Ph.D. Industrial Engineering, State University of New York, 1990
M.S. Industrial Engineering, State University of New York, 1983
B.A. Psychology, State University of New York, 1981

Somchart Thepvongs
Somchart Thepvongs is presently a graduate student in the Industrial and Systems Engineering Department. In the previous year, he has been a Graduate Teaching Assistant for an introductory manufacturing class and a Graduate Research Assistant for a project titled “Analysis of Innovation Best Practices” at NASA. In June 1998, Somchart expects to complete both his thesis work and course work required for the M.S. in Management Systems Engineering at Virginia Tech. Somchart is also a member of the Penn State Alumni Association, the Society of Engineering and Management Systems, and the Institute of Industrial Engineers and the recipient of the Pratt Fellowship for the Fall of 1997.

Education:
B.S. Industrial Engineering, Pennsylvania State University, 1995

Appendix E1 Pre-Experiment Questionnaire

Pre-Experiment Questionnaire

1. Describe your current student status:

- ☐ Sophomore
- ☐ Junior
- ☐ Senior
- ☐ Graduate Student - Masters
- ☐ Graduate Student - Ph.D.

2. Describe your level of knowledge with Statistical Process Control(SPC) or Quality Control (QC):

- ☐ No Experience. I have never heard of SPC or QC.
- ☐ Little Experience. I have read about SPC or QC, but do not have a basic understanding of the skills required to solve SPC/QC problems.
- ☐ Some Experience. I have a basic understanding of SPC/QC, and have a general idea of how to solve these SPC/QC problems.
- ☐ Extensive Experience. I have had a great deal of experience using SPC/QC, and can solve SPC/QC problems.

3. Describe your experience with computer visualization applications which use three-dimensional or perspective displays (e.g. Virtual Reality, Computer Aided Design Applications, and video games):

- ☐ No Experience. I have never learned about or used computer aided visualization applications.
- ☐ Little Experience. I have learned about computer aided visualization applications, but have never used them.
- ☐ Some Experience. I have learned about computer aided visualization applications in the past but have only limited experience with them.
- ☐ Extensive Experience. I have learned and used computer aided visualization applications extensively in the past.

4. Describe your previous experience with the Internet:

- ☐ No Experience. I have never used the Internet.
- ☐ Little Experience. I have used the Internet only a few times to browse various sites.
- ☐ Some Experience. I use the Internet to browse sites occasionally.
- ☐ Extensive Experience. I use the Internet to browse sites on a daily basis.

Appendix E2 Post-Experiment Questionnaire

Post Experiment Questionnaire

Please answer the following questions. Examples of the referenced displays appear at the bottom of the page.

1. Overall, I believe the following was the best display to use was:

☐ Multiple 2-D displays ☐ Composite 2-D display ☐ Composite 3-D display

2. The reason why I think my answer in question 1 is the best display is because:

3. Overall, I believe the following was the second best display to use was:

☐ Multiple 2-D displays ☐ Composite 2-D display ☐ Composite 3-D display

4. The reason why I think my answer in question 3 is the second best display is because:

5. Overall, I believe the following was the third best display to use was:

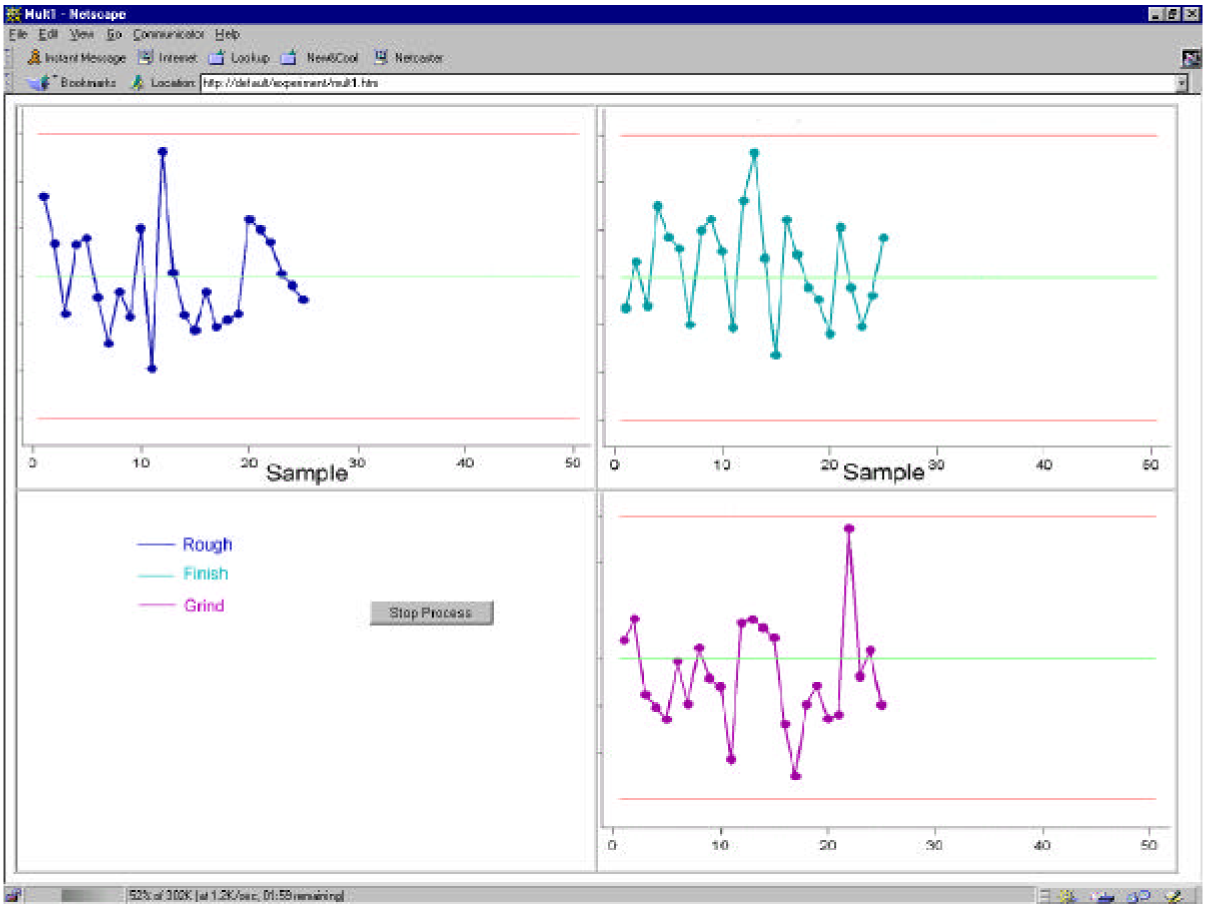
☐ Multiple 2-D displays ☐ Composite 2-D display ☐ Composite 3-D display

6. The reason why I think my answer in question 3 is the third best display is because:

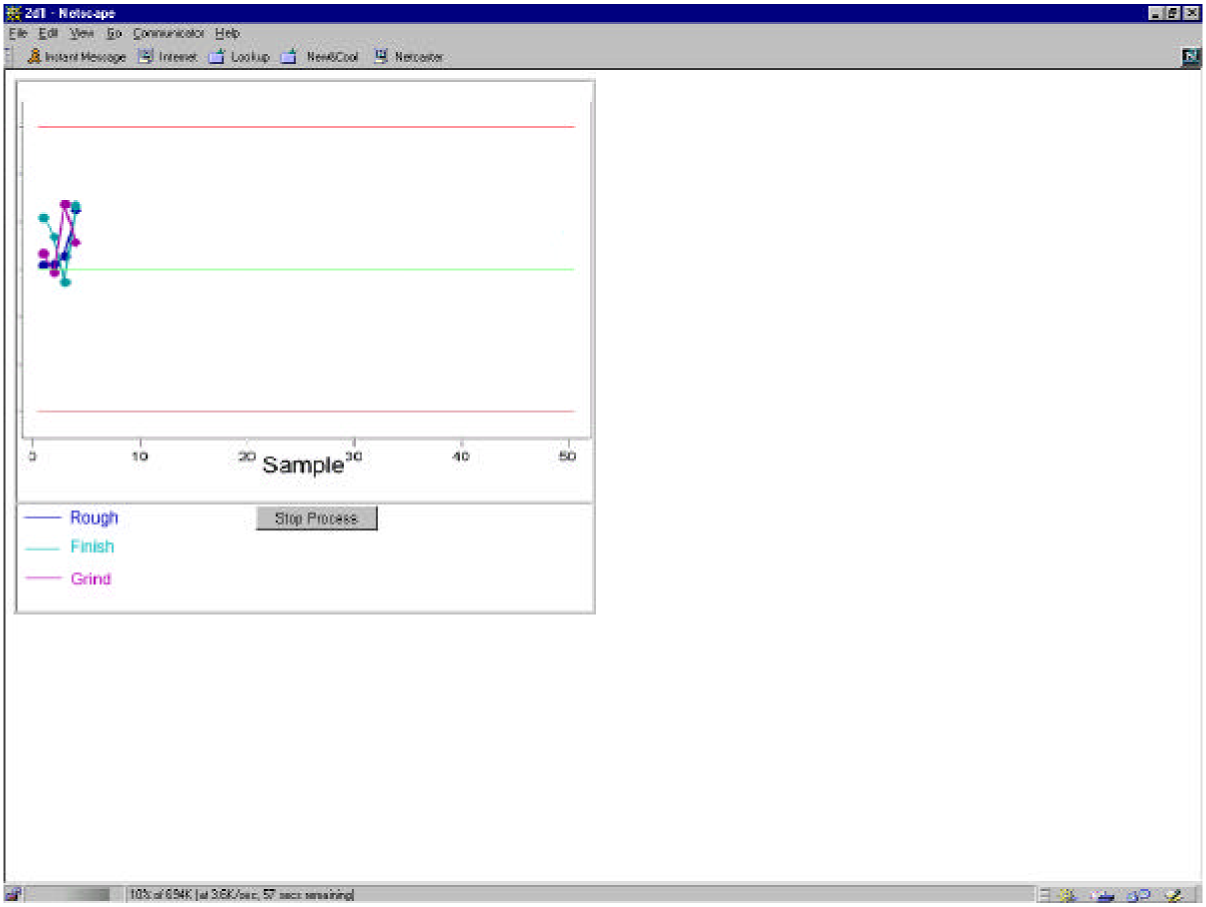
7. Given your participation in this experiment, if you could design a manufacturing line in which you could choose the number of displays and machines one person was responsible for, what would the manufacturing line consist of? Why would you design the manufacturing line this way?

Examples of Displays:

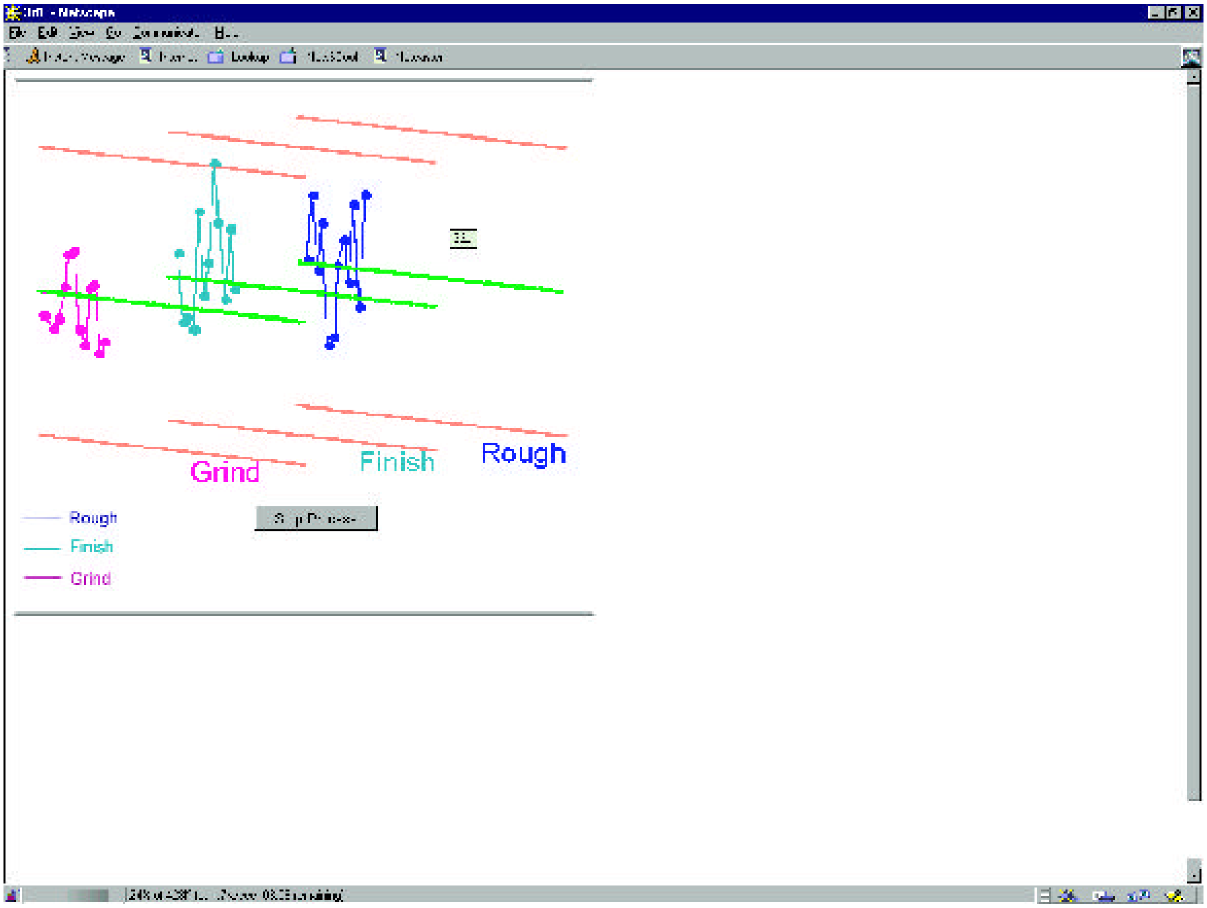
Multiple 2-D displays:



Composite 2-D display:



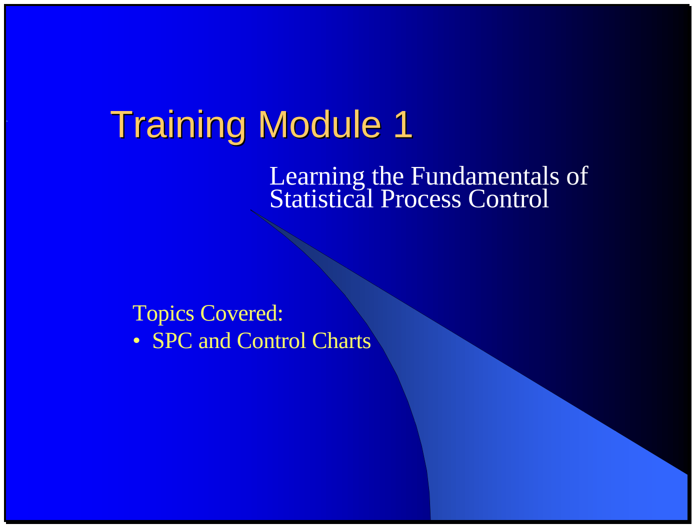
Composite 3-D display:



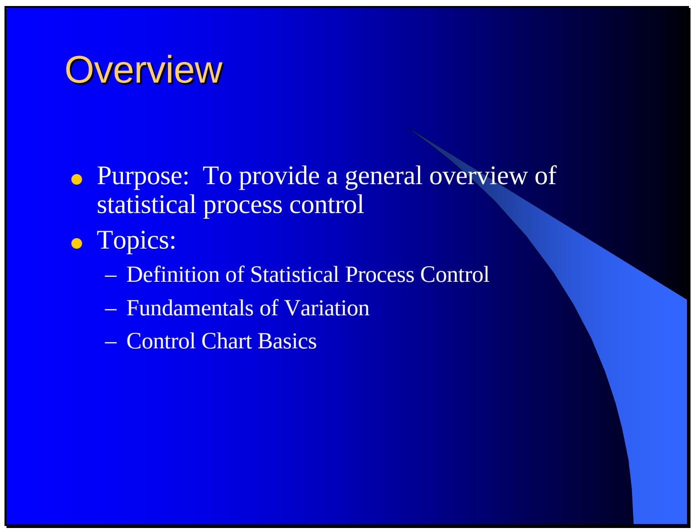
Appendix F1 Training

This section contains the Microsoft PowerPoint slides used for training subjects in statistical process control utilizing the progressive part training method.

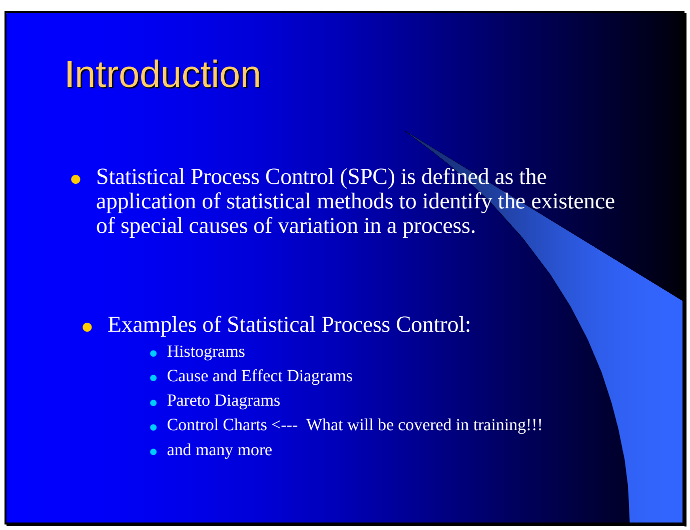
Slide 1



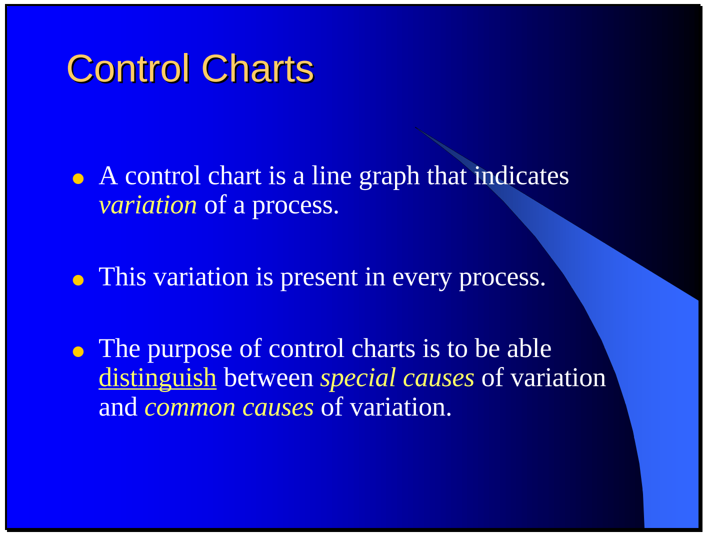
Slide 2



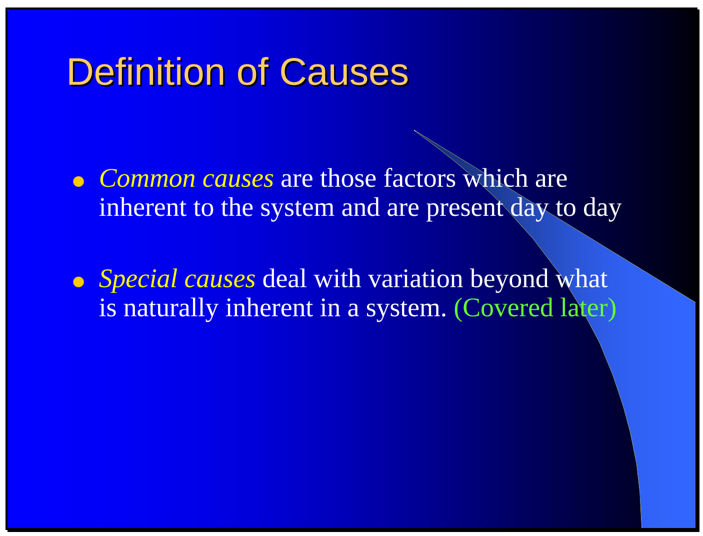
Slide 3



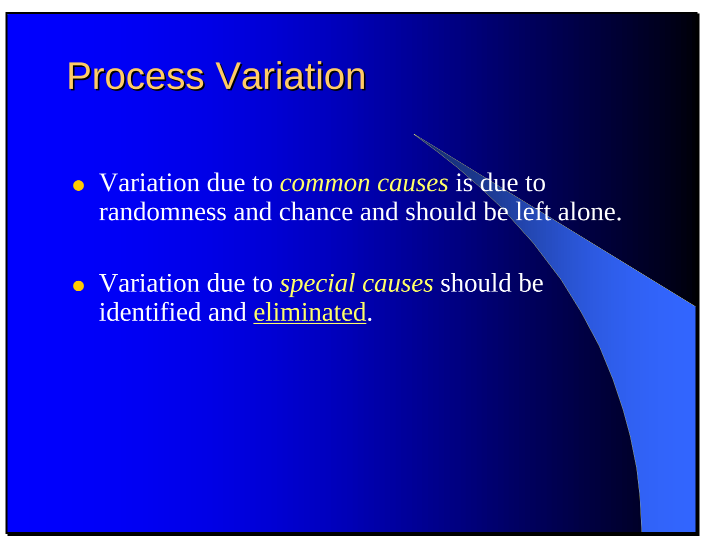
Slide 4



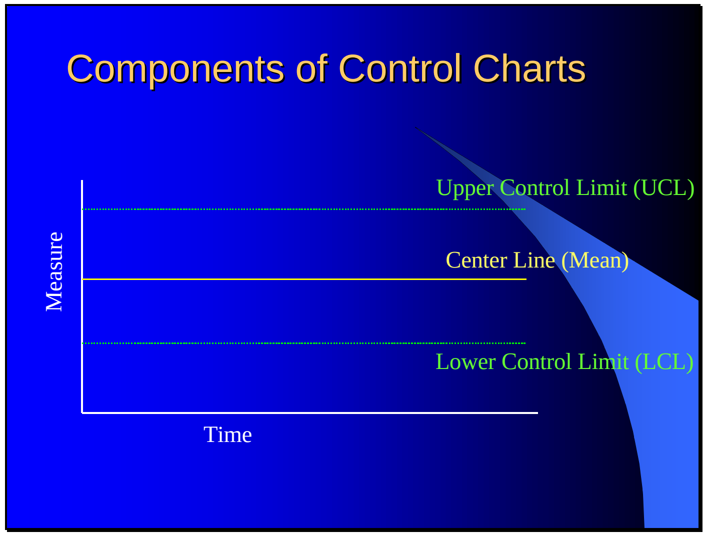
Slide 5



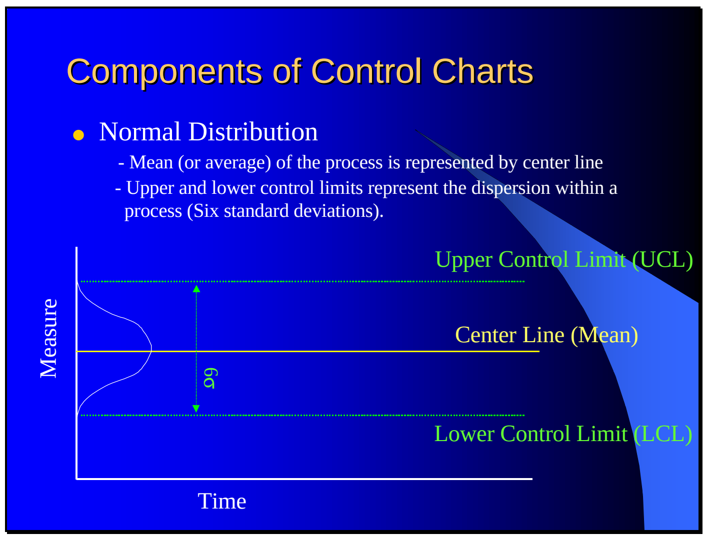
Slide 6



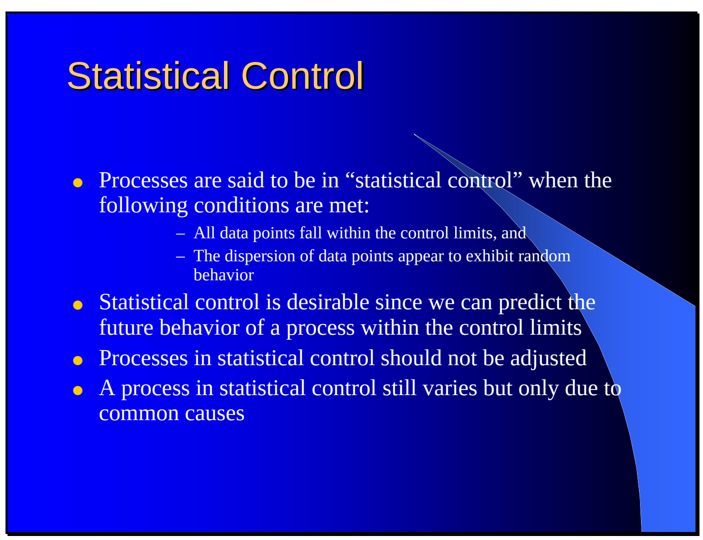
Slide 7



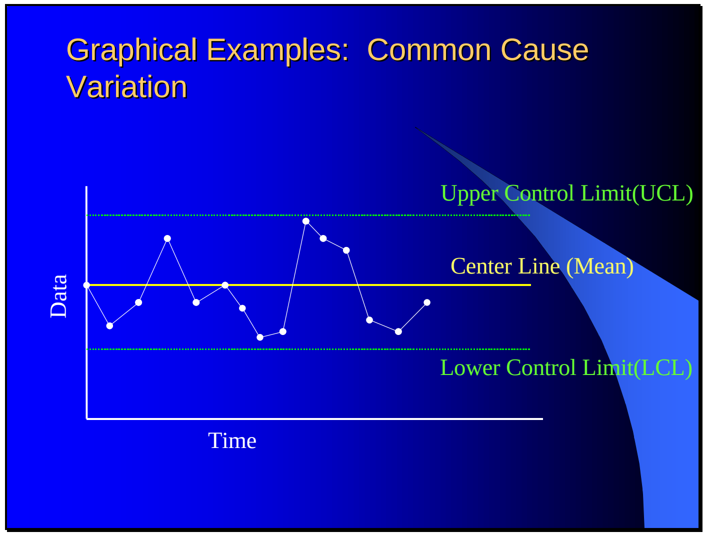
Slide 8



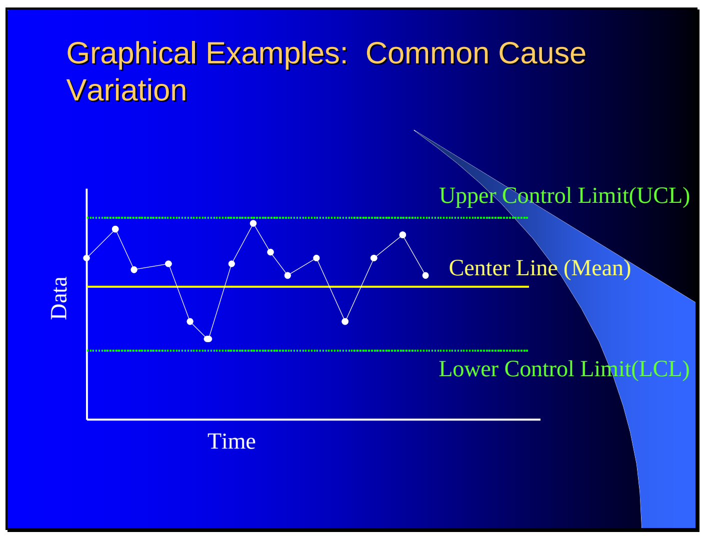
Slide 9



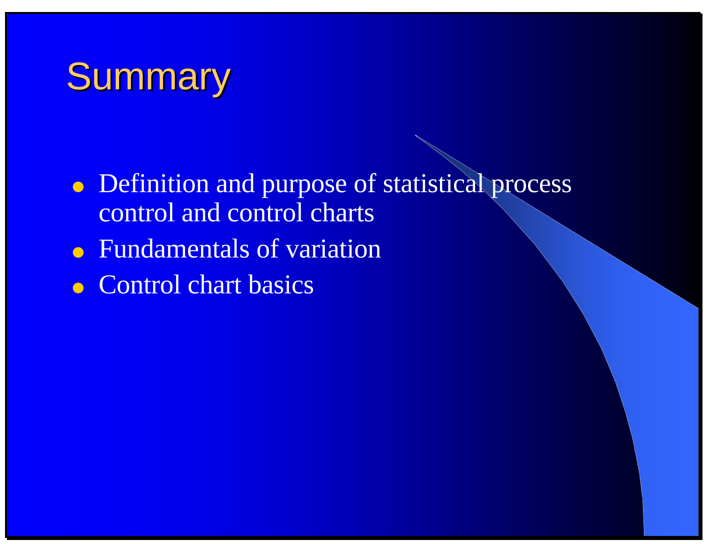
Slide 10



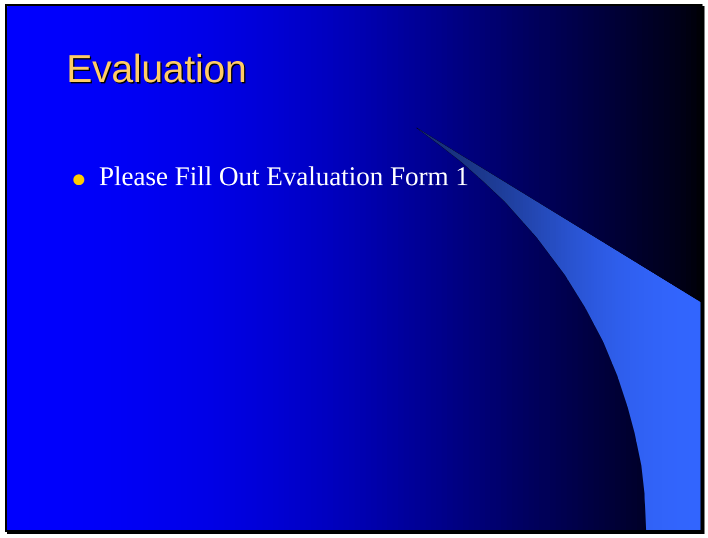
Slide 11



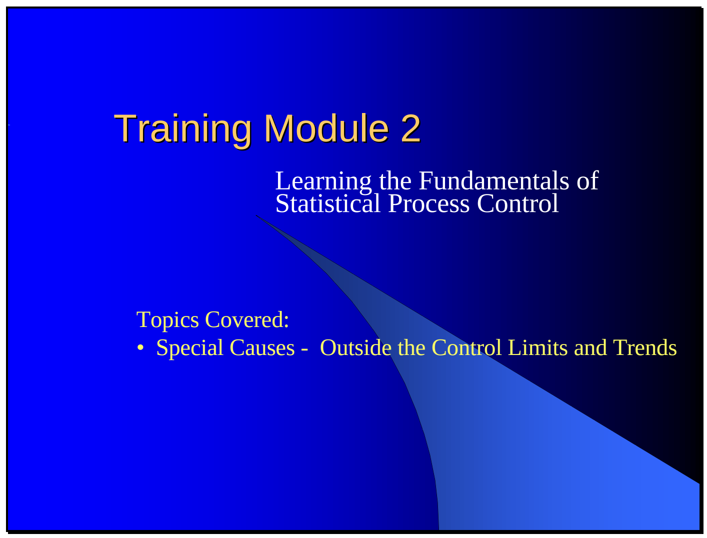
Slide 12



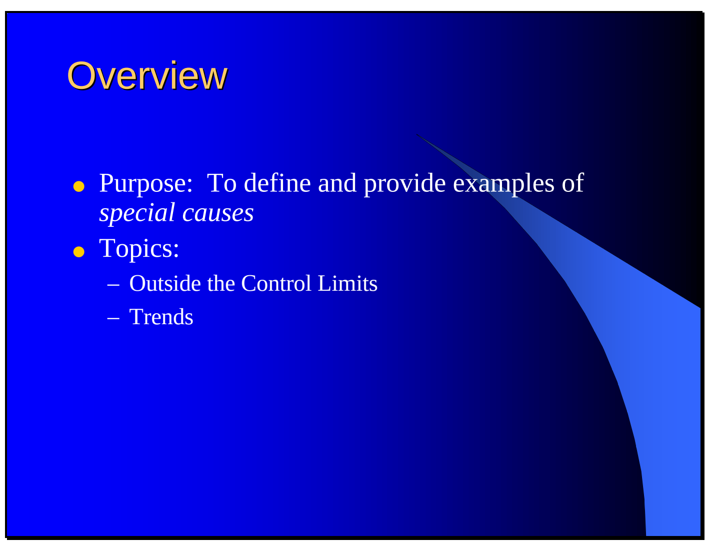
Slide 13



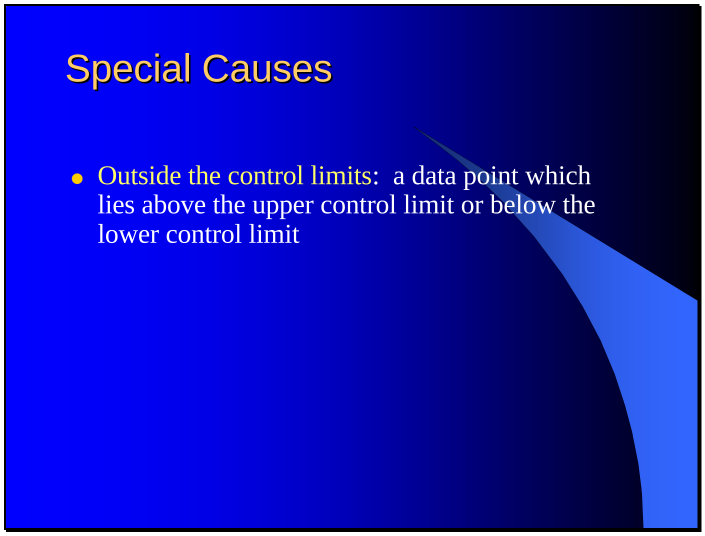
Slide 14



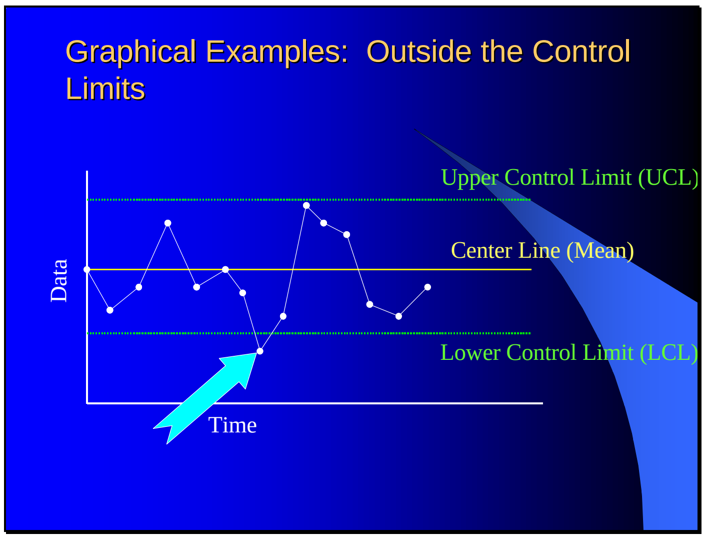
Slide 15



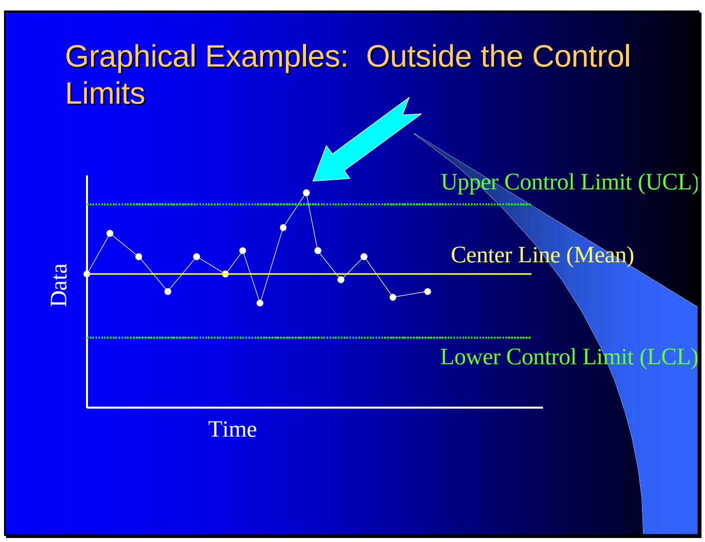
Slide 16



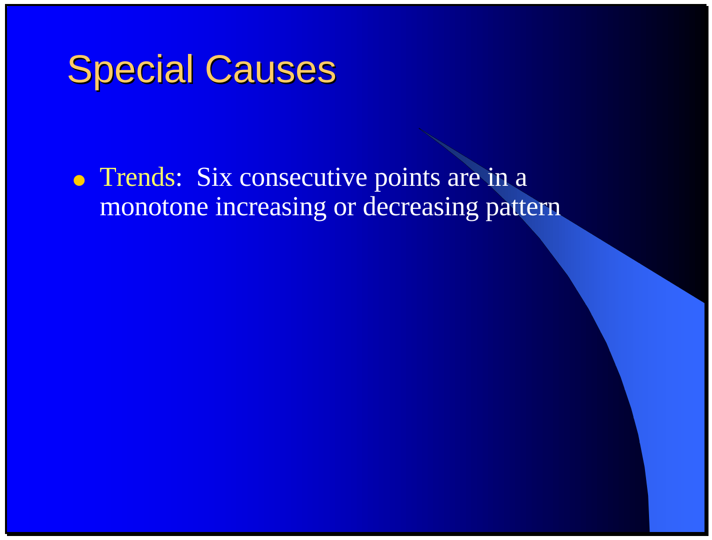
Slide 17



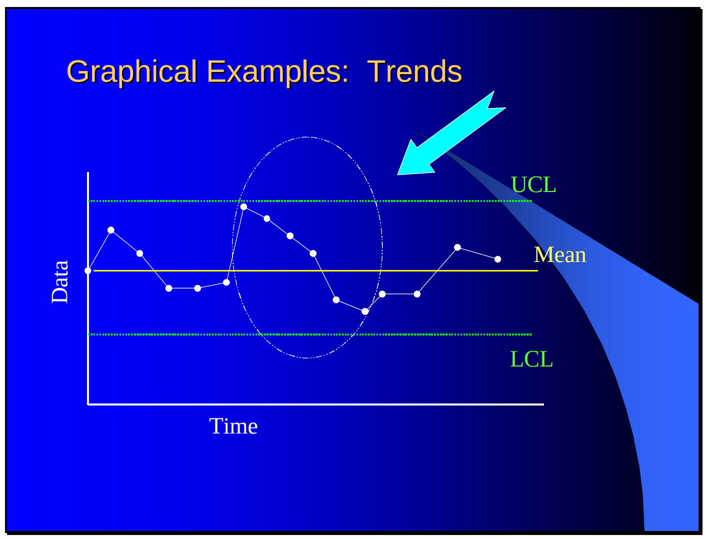
Slide 18



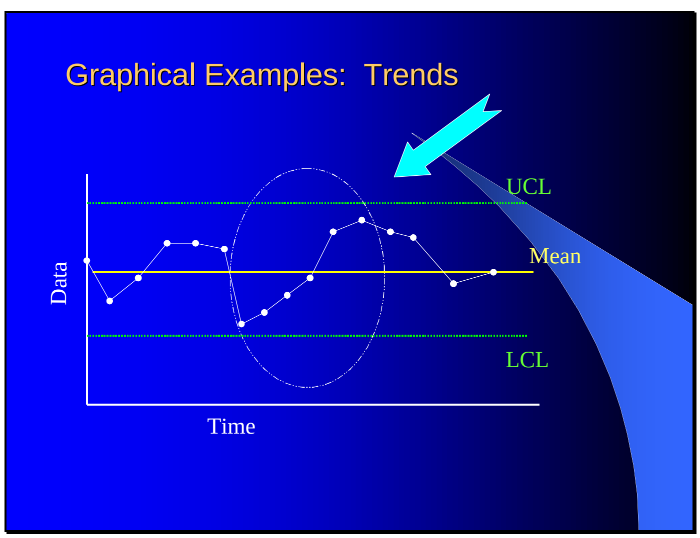
Slide 19



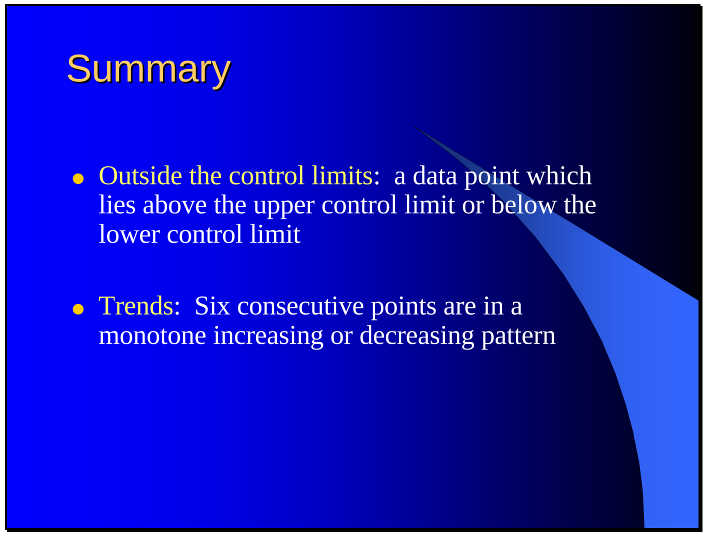
Slide 20



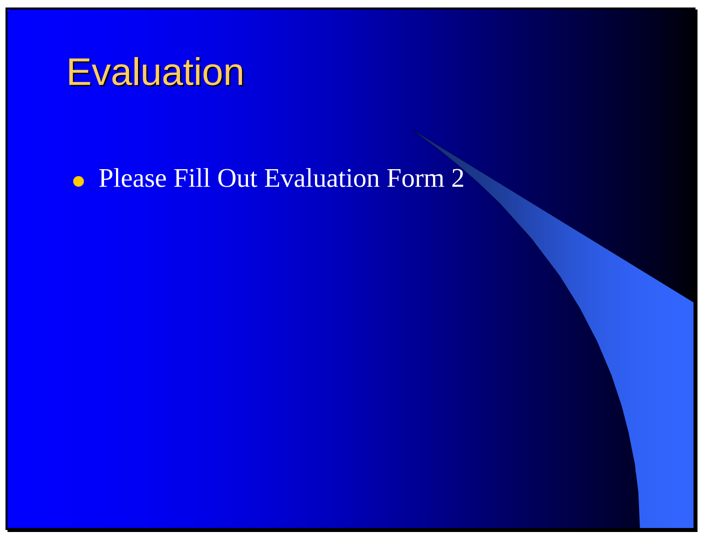
Slide 21



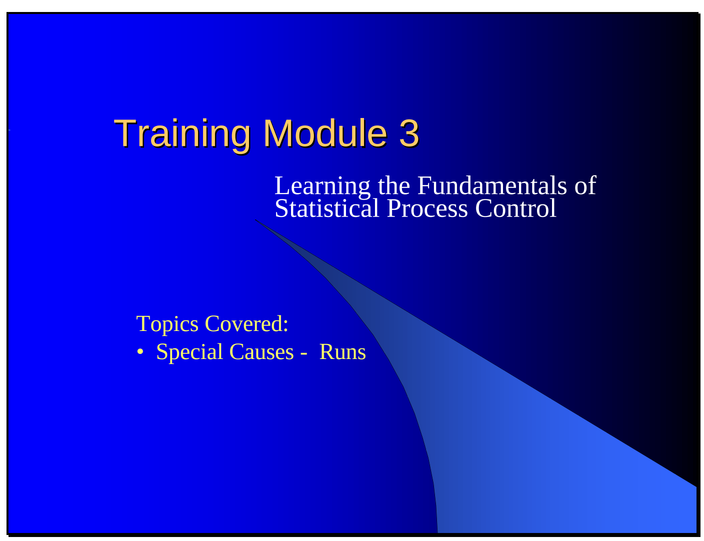
Slide 22



Slide 23



Slide 24



Slide 25

Overview

- Purpose: To define and provide examples of *special causes*
- Topics:
  - Runs

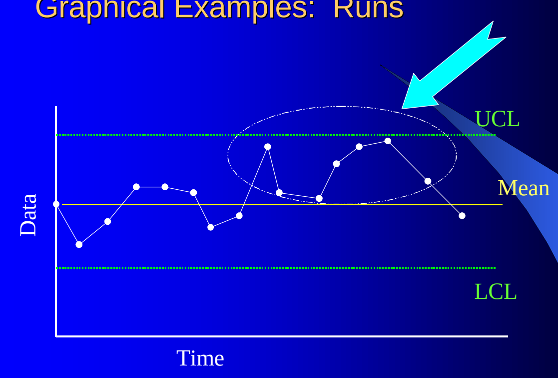
Slide 26

Special Causes

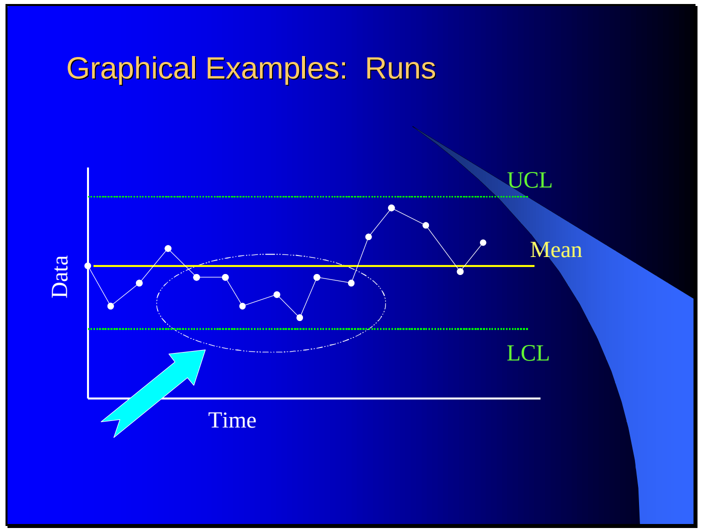
- Runs: Seven consecutive points above the mean or seven consecutive points below the mean

Slide 27

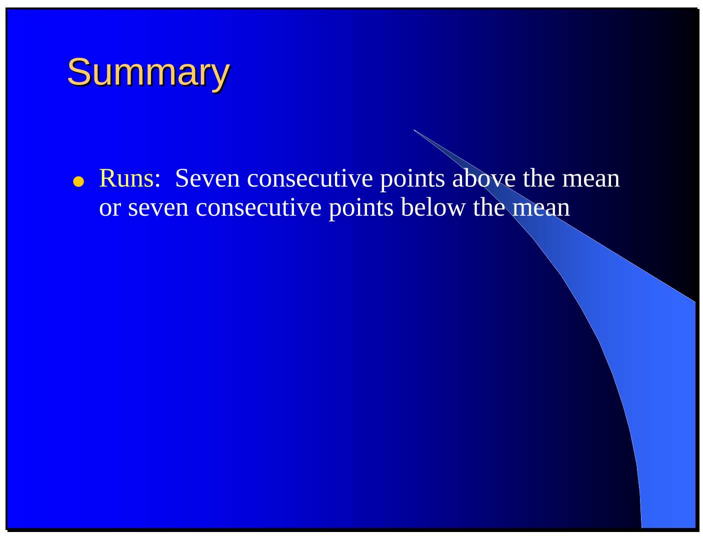
Graphical Examples: Runs



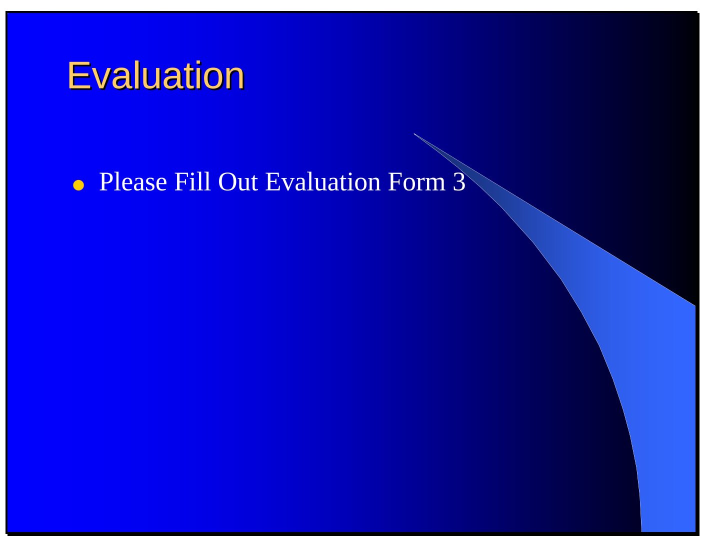
Slide 28



Slide 29



Slide 30



Slide 31

Training Module 4

Learning the Fundamentals of
Statistical Process Control

Topics Covered:
- Phase Related Correlation

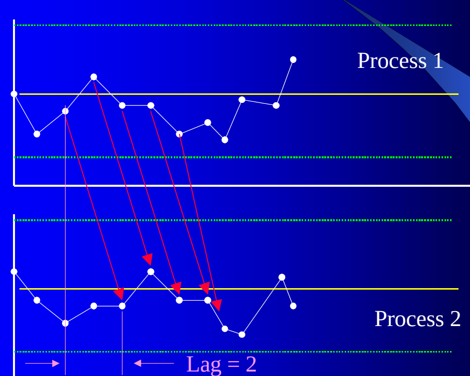
Slide 32

Special Causes

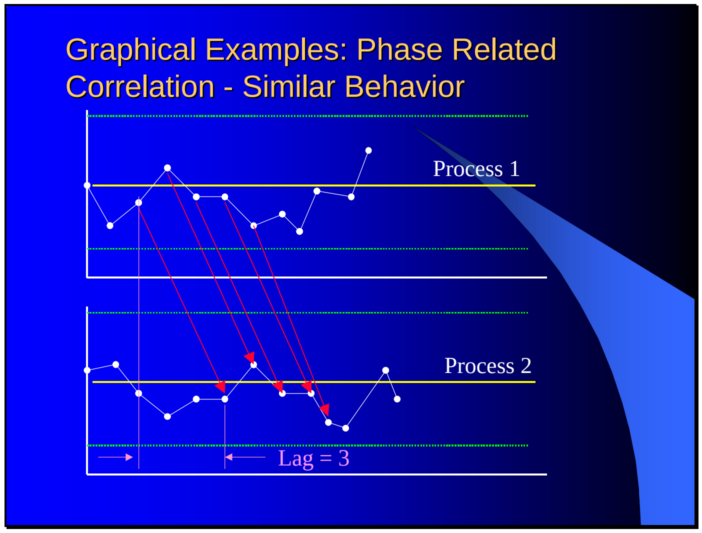
- Phase Related Correlation: Five consecutive points which exhibit similar behavior or inverse behavior (magnitude and direction) in two control charts based on some lag difference.
- Lag: The number of points between the start of a pattern on one control chart and the start of a similar pattern on a second control chart.

Slide 33

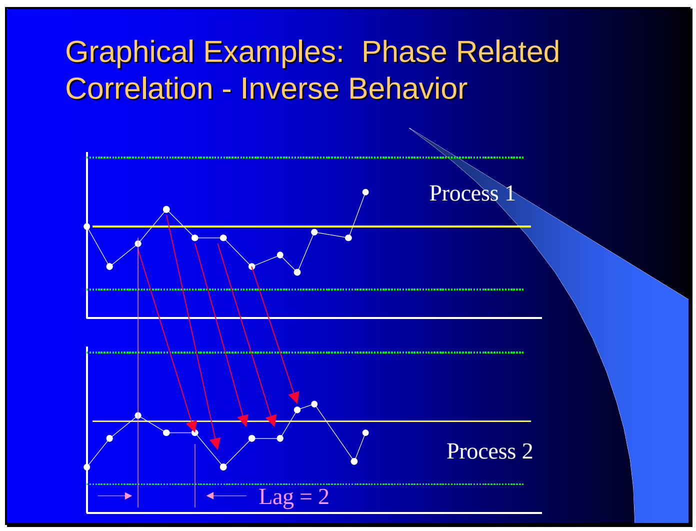
Graphical Examples: Phase Related Correlation - Similar Behavior



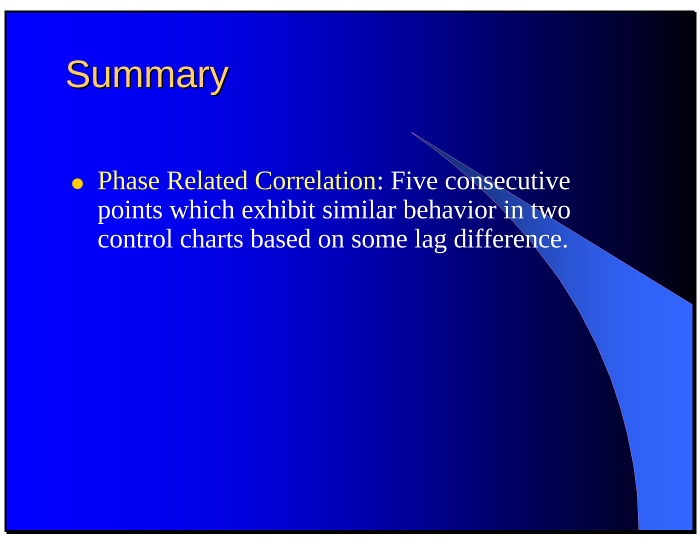
Slide 34



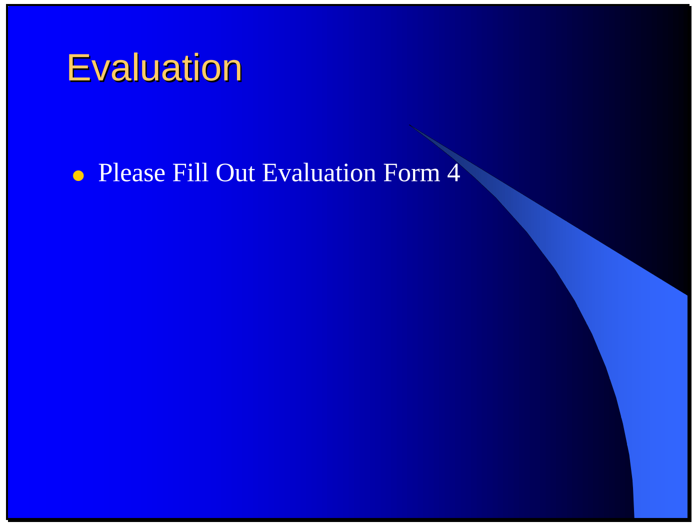
Slide 35



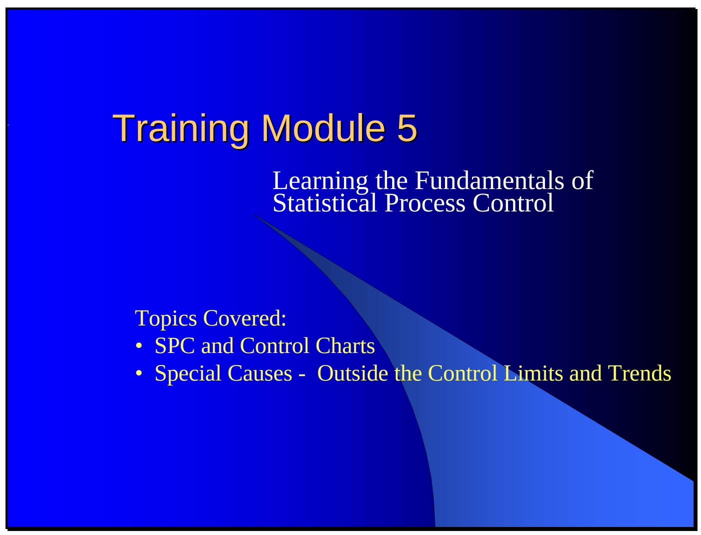
Slide 36



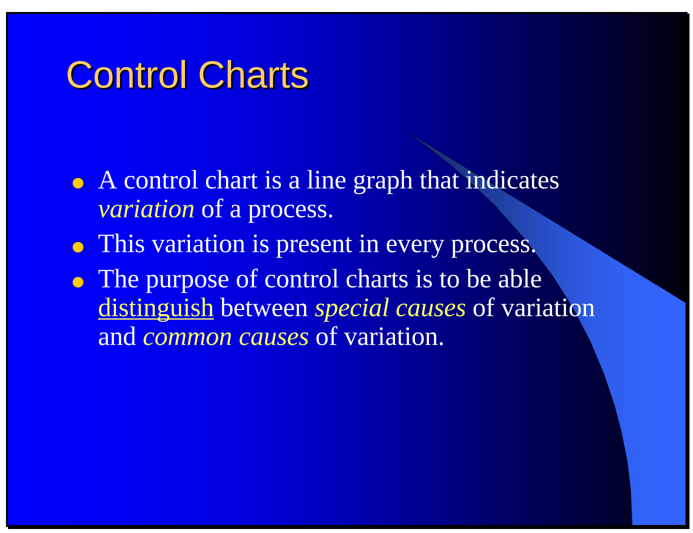
Slide 37



Slide 38



Slide 39



Slide 40

Definition of Causes

- *Common causes* are those factors which are inherent to the system and are present day to day
- *Special causes* deal with variation beyond what is naturally inherent in a system.

Slide 41

Process Variation

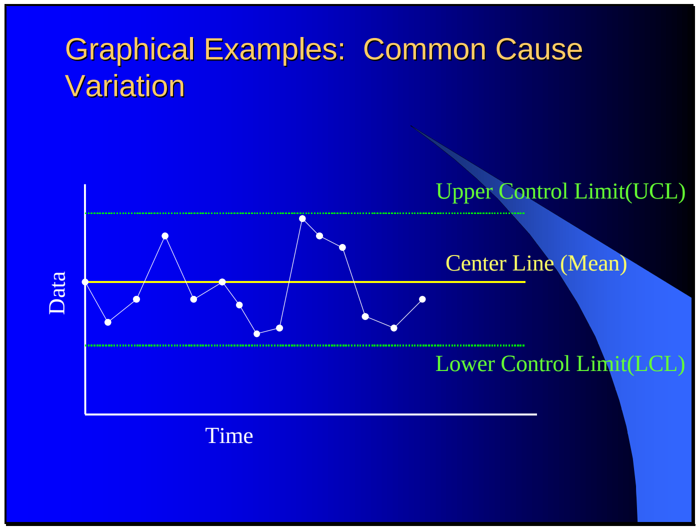
- Variation due to *common causes* is due to randomness and chance and should be left alone.
- Variation due to *special causes* should be identified and eliminated.

Slide 42

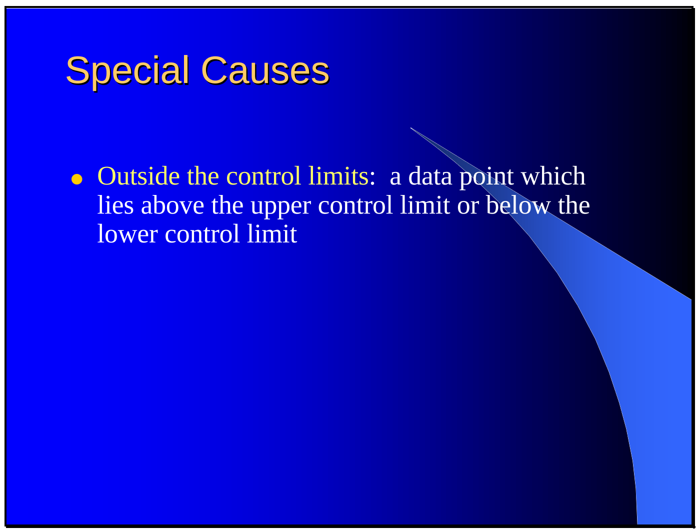
Statistical Control

- Processes are said to be in “statistical control” when the following conditions are met:
  - All data points fall within the control limits, and
  - The dispersion of data points appear to exhibit random behavior
- Statistical control is desirable since we can predict the future behavior of a process within the control limits
- Processes in statistical control should not be adjusted
- A process in statistical control still varies but only due to common causes

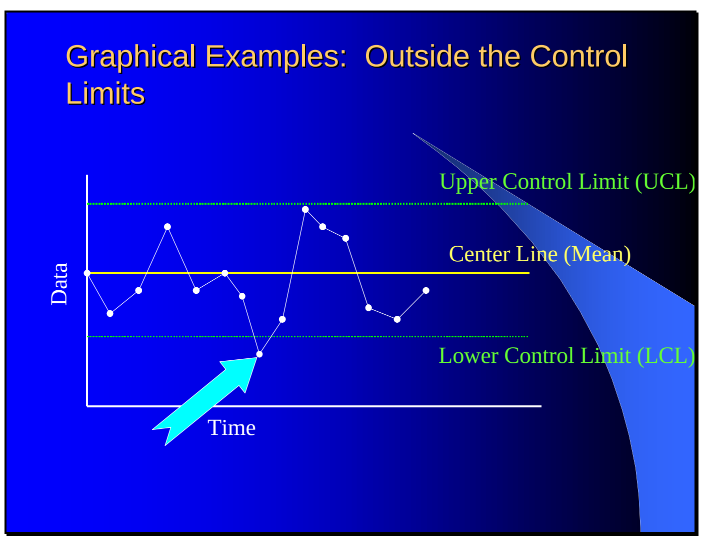
Slide 43



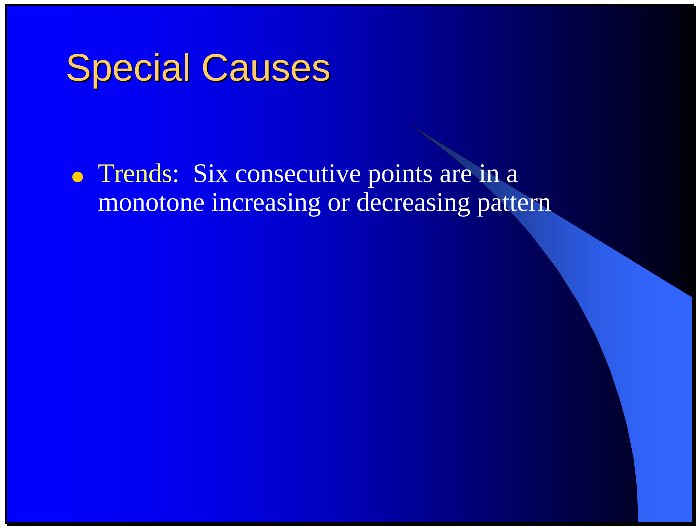
Slide 44



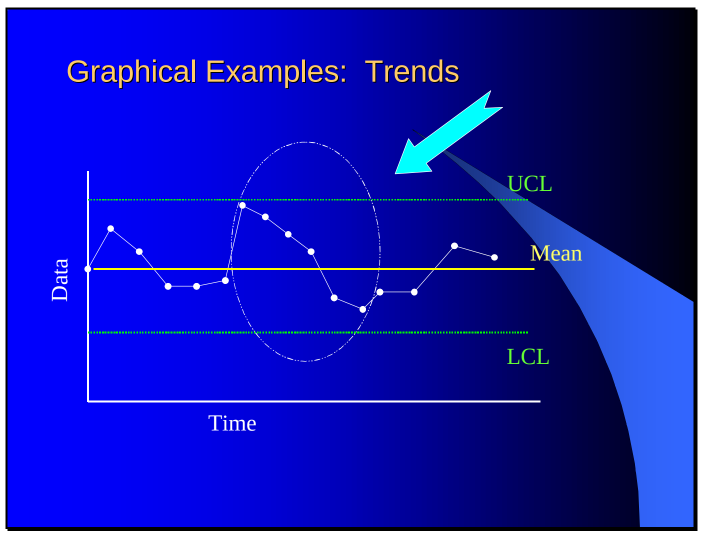
Slide 45



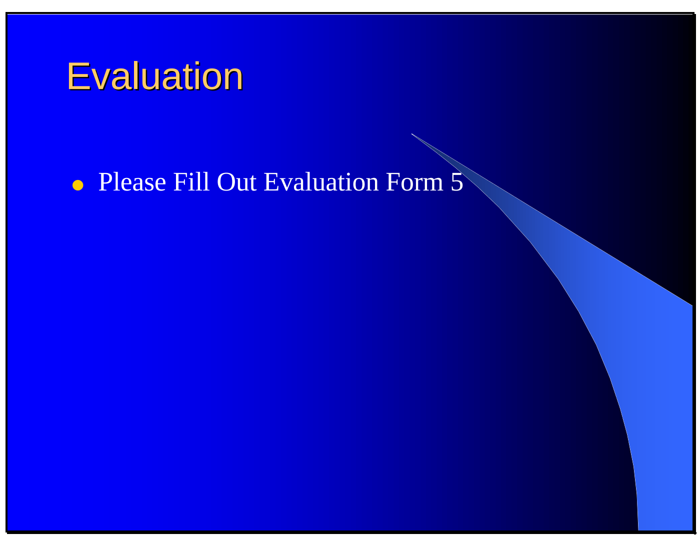
Slide 46



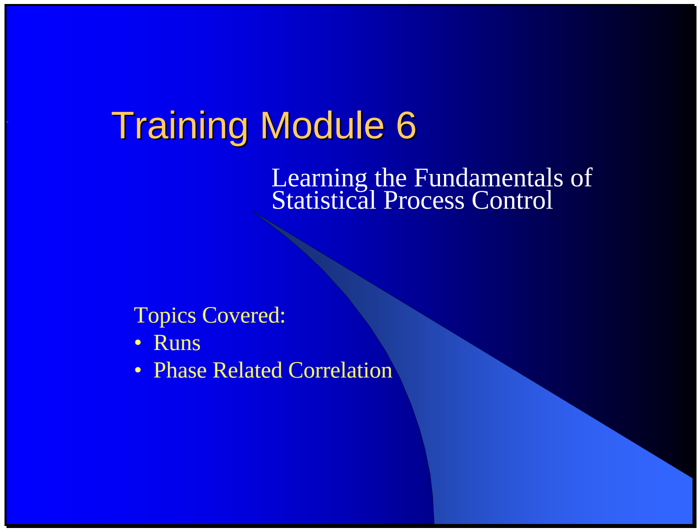
Slide 47



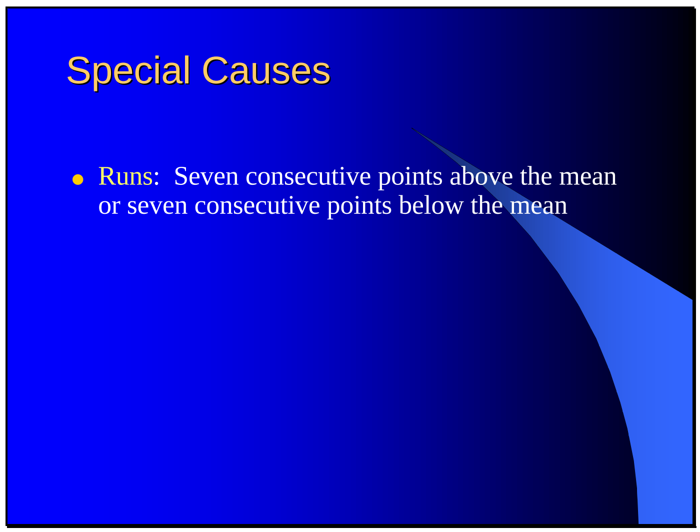
Slide 48



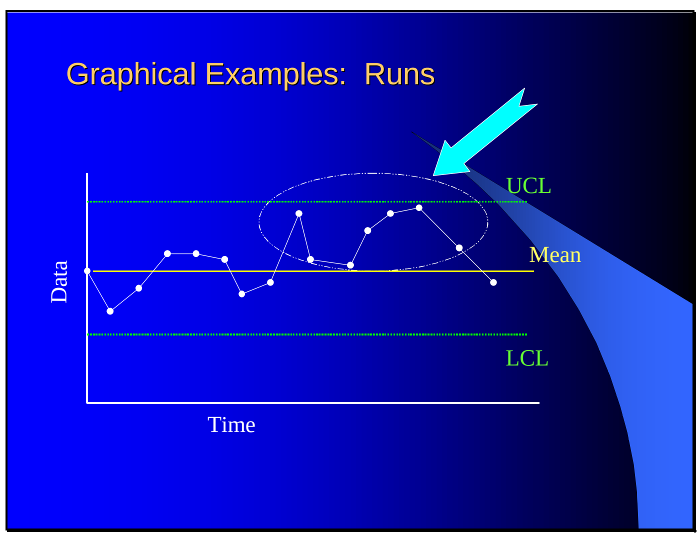
Slide 49



Slide 50



Slide 51



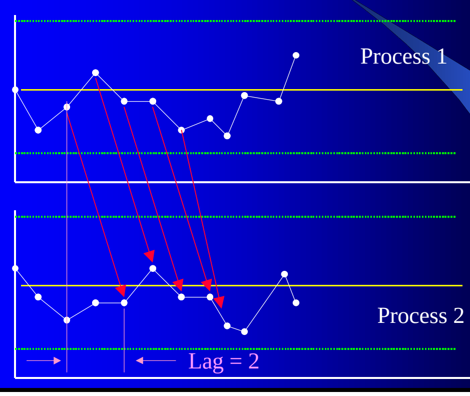
Slide 52

Special Causes

- Phase Related Correlation: Five consecutive points which exhibit similar behavior or inverse behavior (magnitude and direction) in two control charts based on some lag difference.
- Lag: The number of points between the start of a pattern on one control chart and the start of a similar pattern on a second control chart.

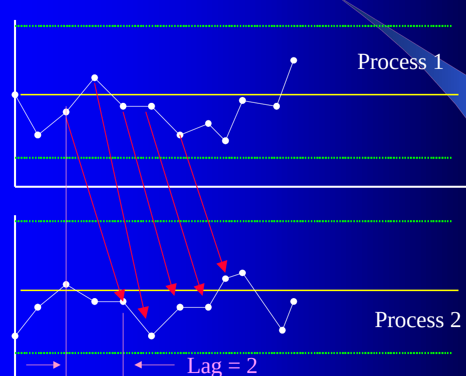
Slide 53

Graphical Examples: Phase Related Correlation - Similar Behavior

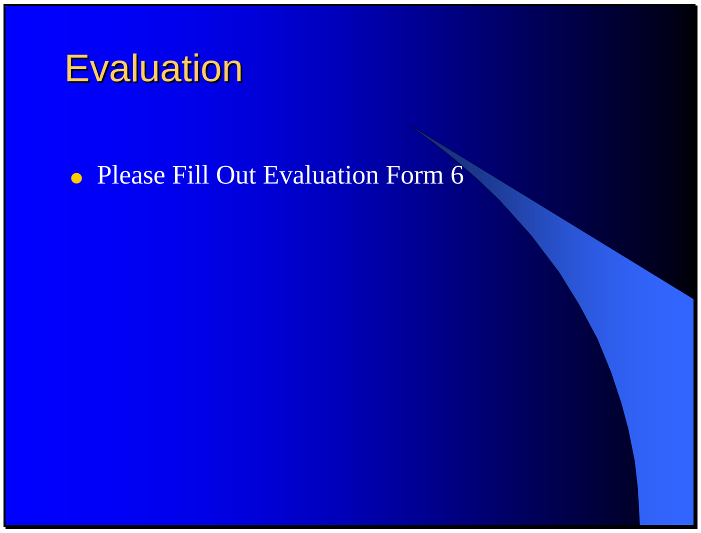


Slide 54

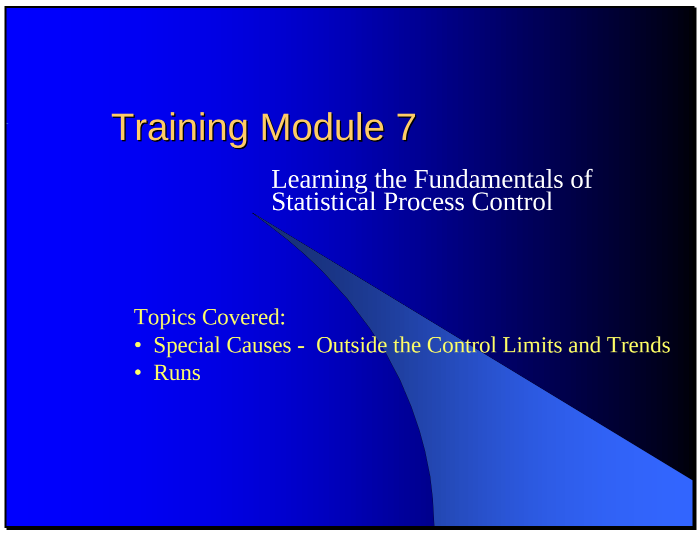
Graphical Examples: Phase Related Correlation - Inverse Behavior



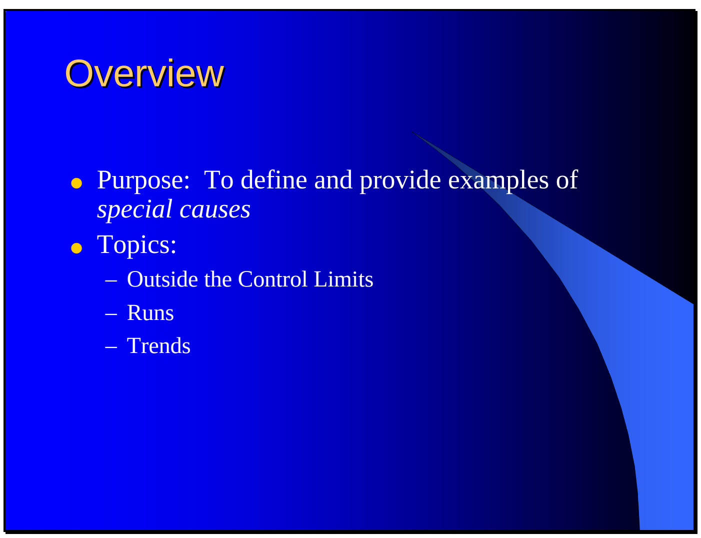
Slide 55



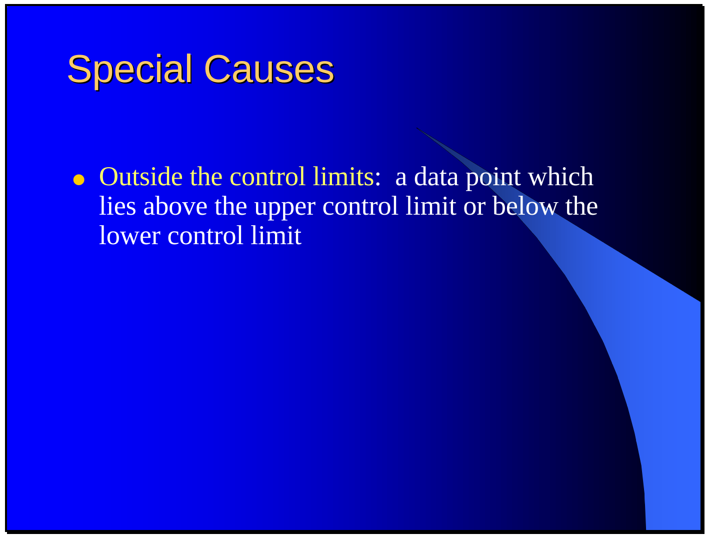
Slide 56



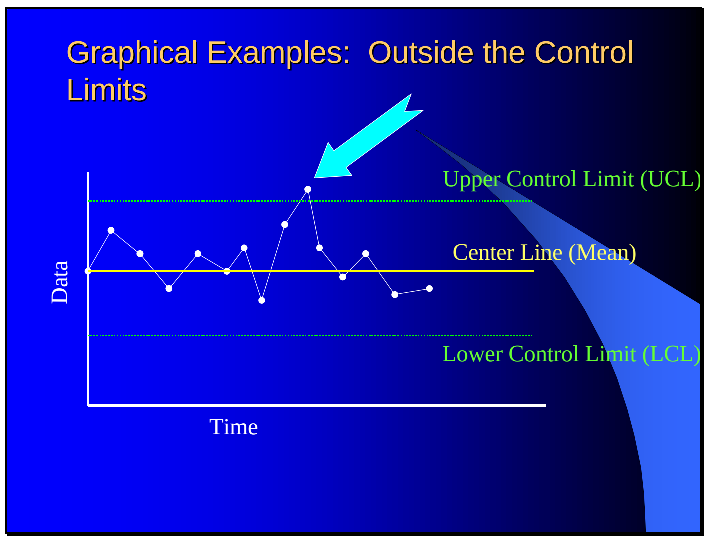
Slide 57



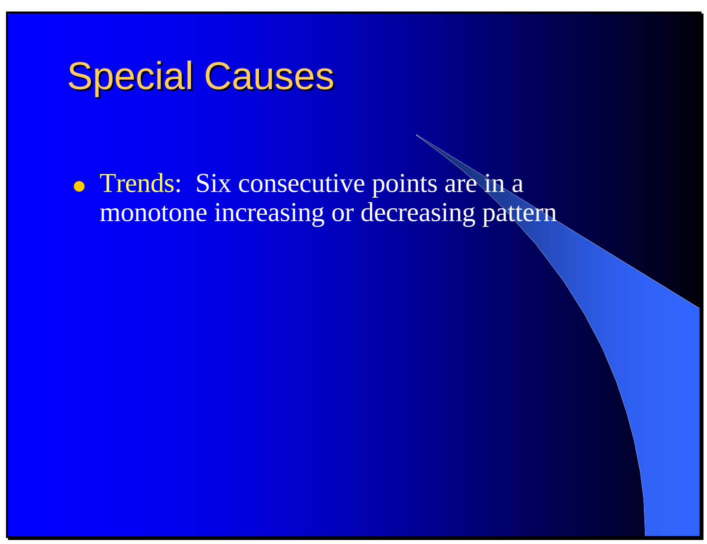
Slide 58



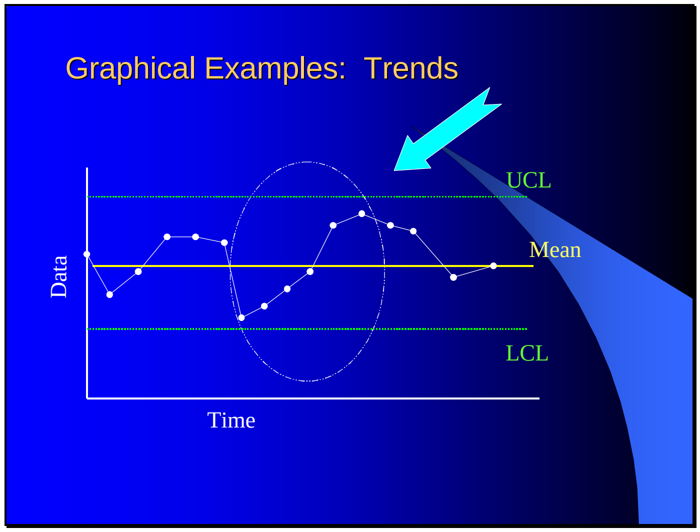
Slide 59



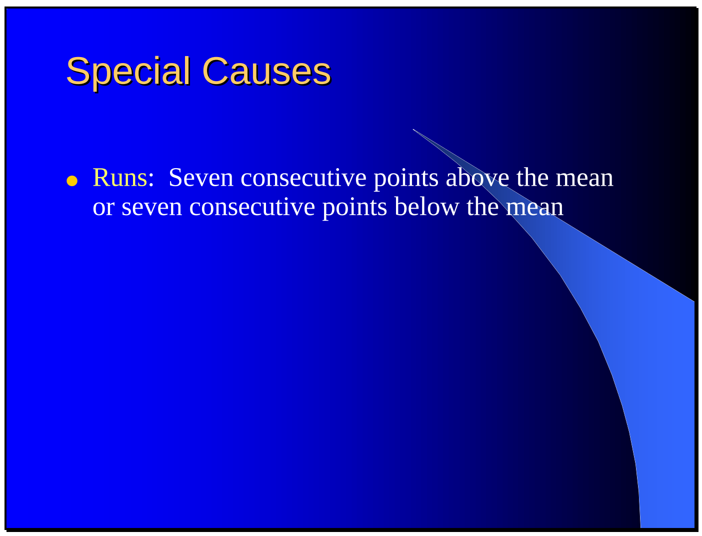
Slide 60



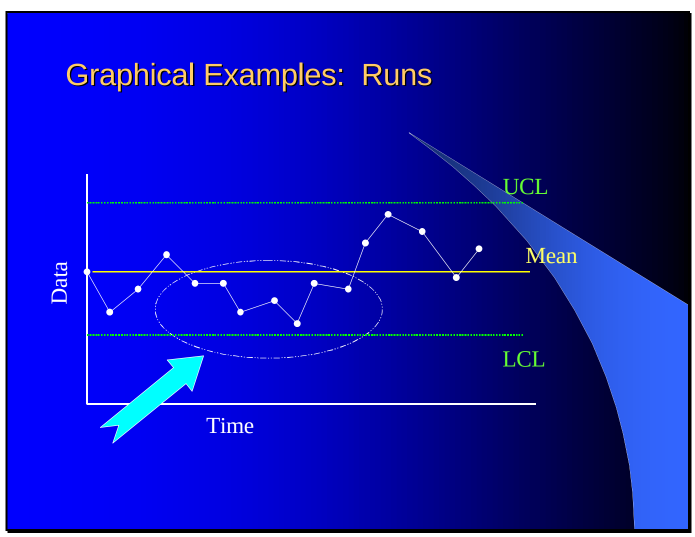
Slide 61



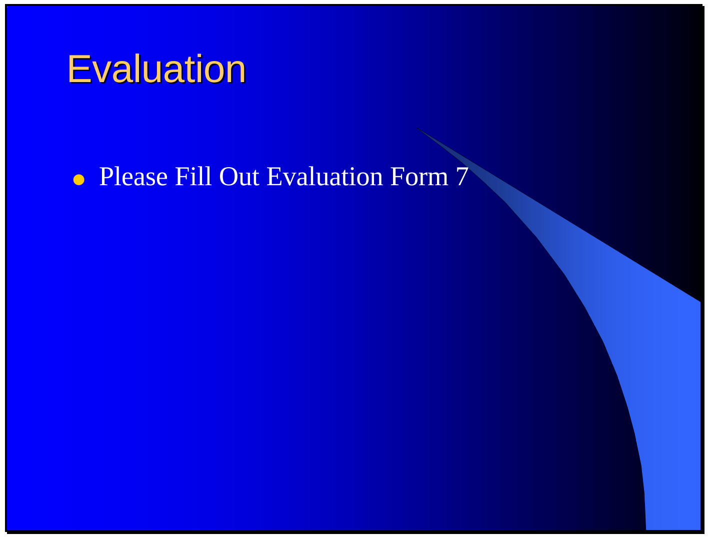
Slide 62



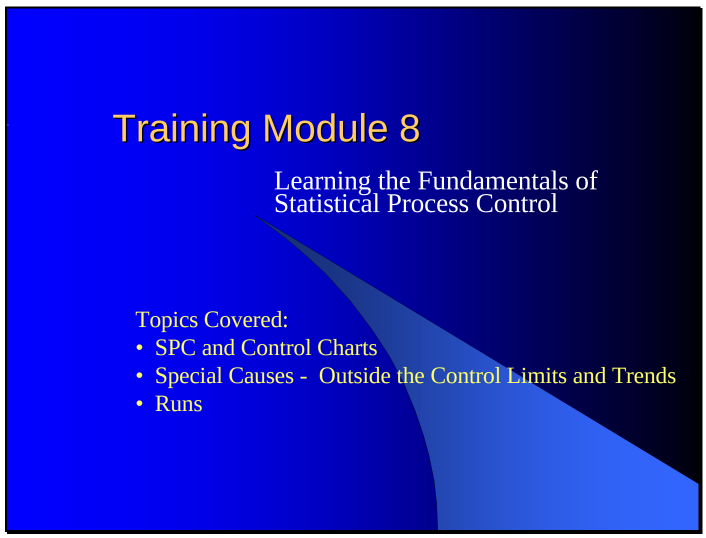
Slide 63



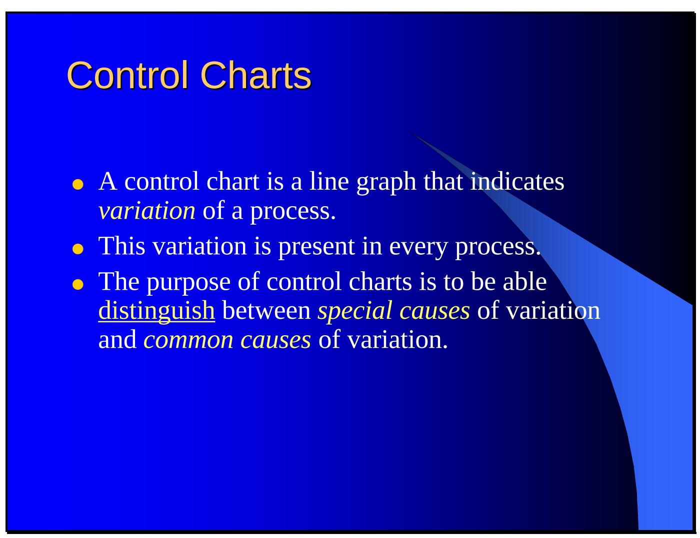
Slide 64



Slide 65



Slide 66



Slide 67

Definition of Causes

- *Common causes* are those factors which are inherent to the system and are present day to day
- *Special causes* deal with variation beyond what is naturally inherent in a system.

Slide 68

Process Variation

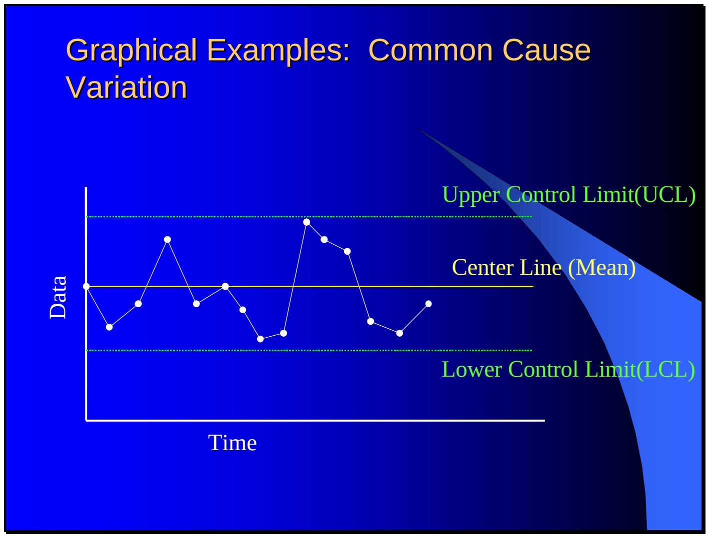
- Variation due to *common causes* is due to randomness and chance and should be left alone.
- Variation due to *special causes* should be identified and eliminated.

Slide 69

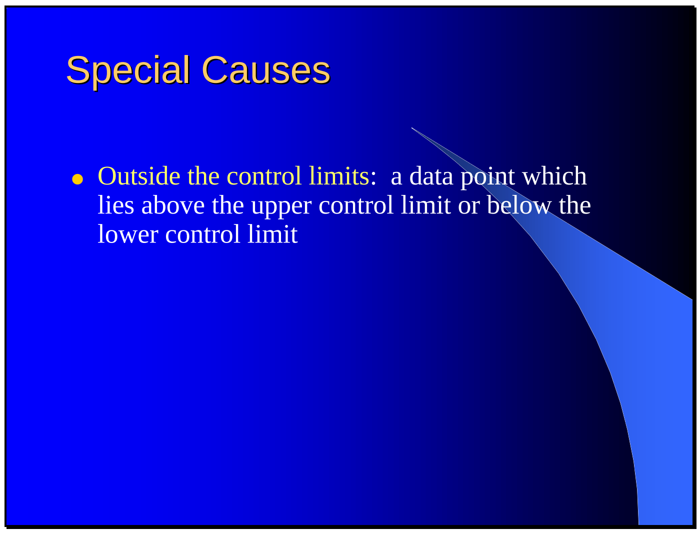
Statistical Control

- Processes are said to be in "statistical control" when the following conditions are met:
  - All data points fall within the control limits, and
  - The dispersion of data points appear to exhibit random behavior
- Statistical control is desirable since we can predict the future behavior of a process within the control limits
- Processes in statistical control should not be adjusted
- A process in statistical control still varies but only due to common causes

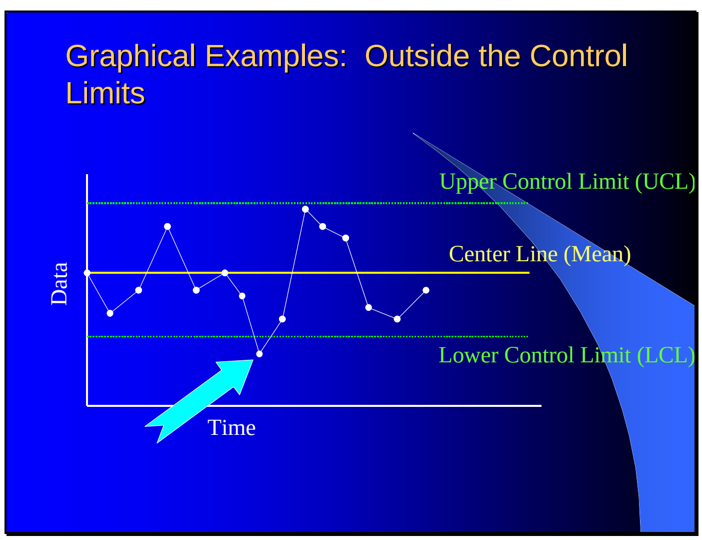
Slide 70



Slide 71



Slide 72



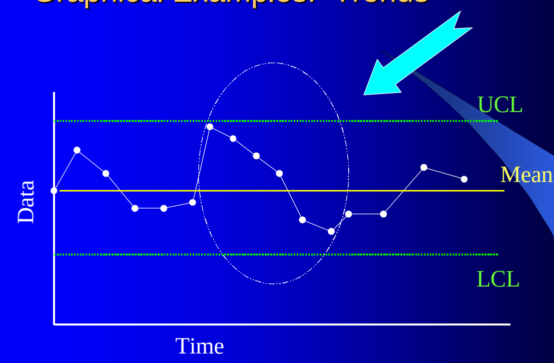
Slide 73

Special Causes

- Trends: Six consecutive points are in a monotone increasing or decreasing pattern

Slide 74

Graphical Examples: Trends

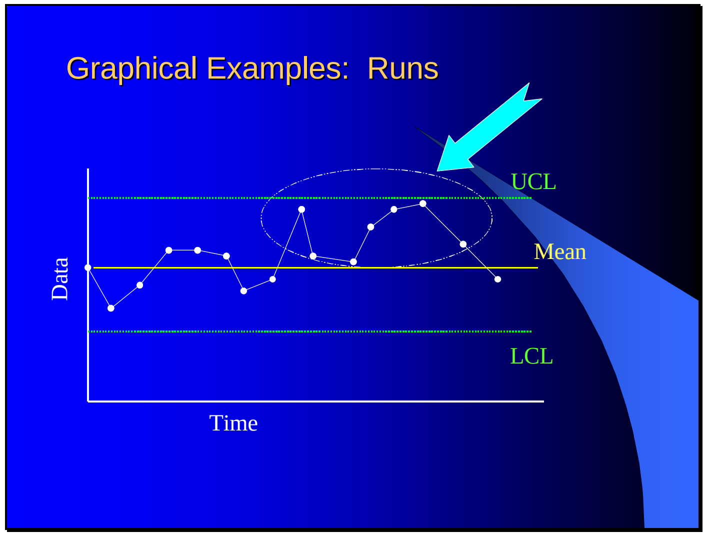


Slide 75

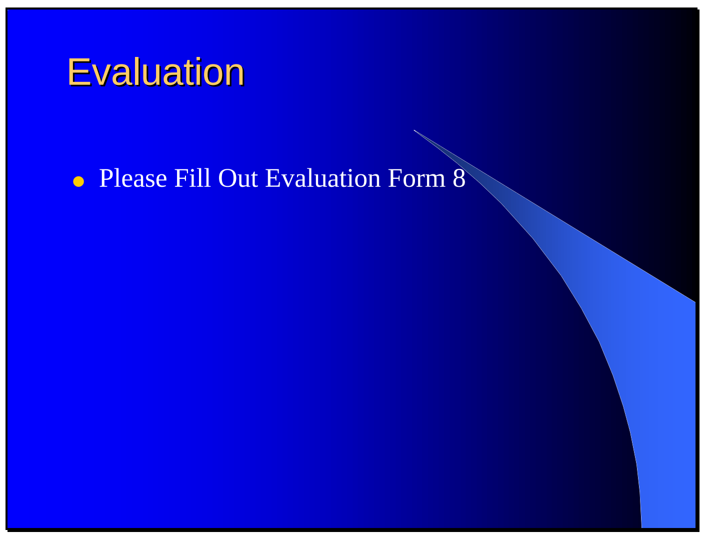
Special Causes

- Runs: Seven consecutive points above the mean or seven consecutive points below the mean

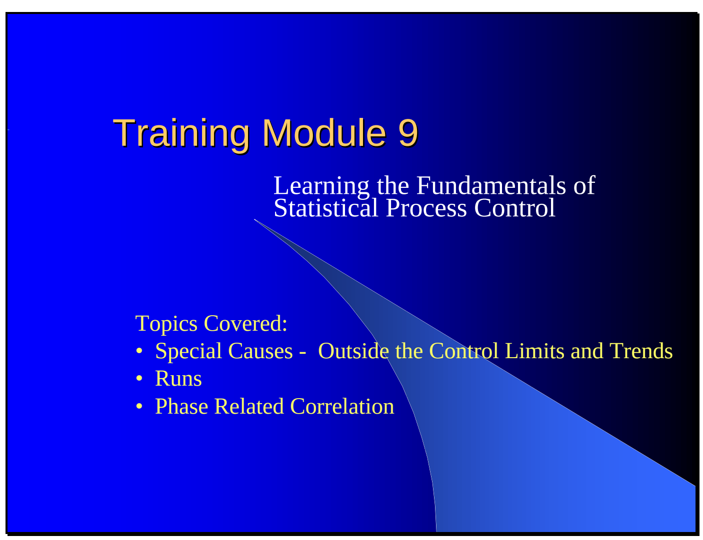
Slide 76



Slide 77



Slide 78



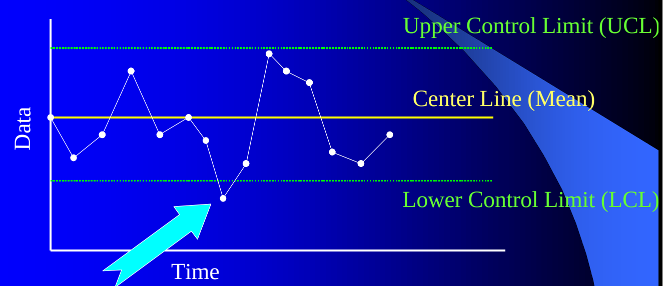
Slide 79

Special Causes

- Outside the control limits: a data point which lies above the upper control limit or below the lower control limit

Slide 80

Graphical Examples: Outside the Control Limits

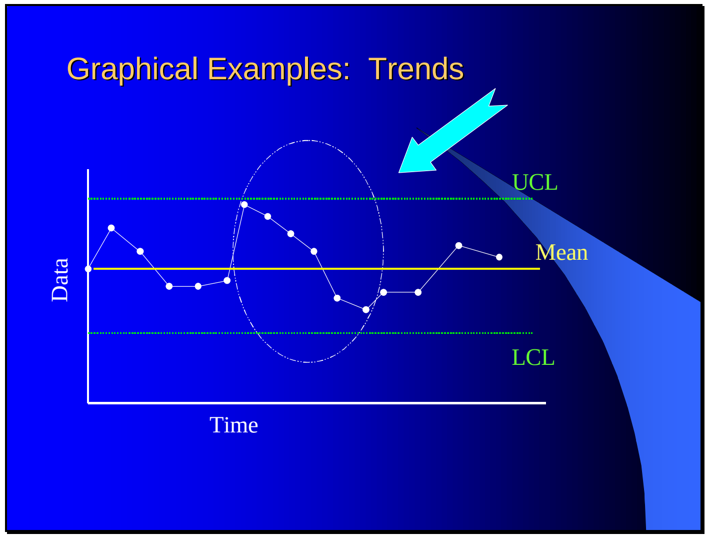


Slide 81

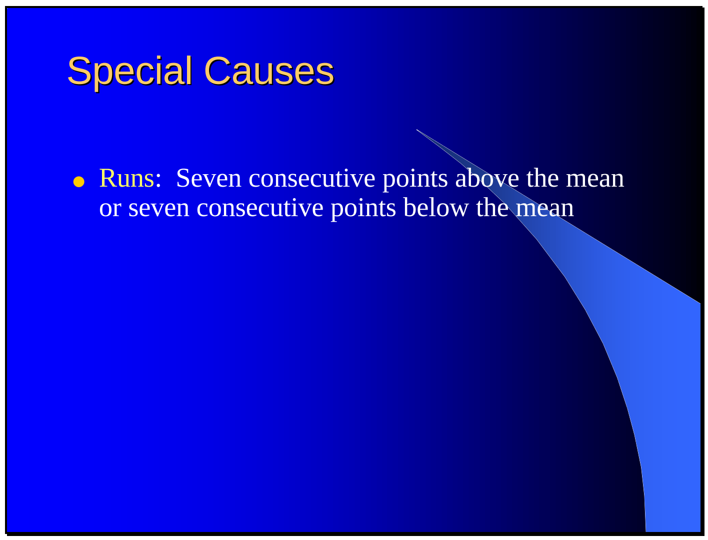
Special Causes

- Trends: Six consecutive points are in a monotone increasing or decreasing pattern

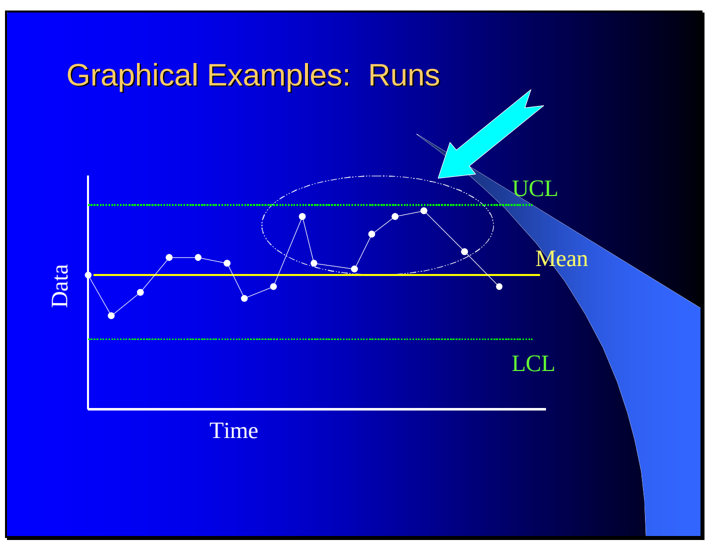
Slide 82



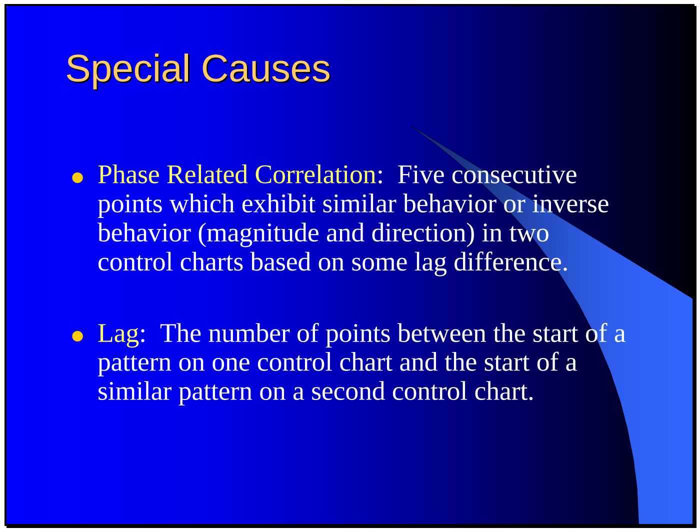
Slide 83



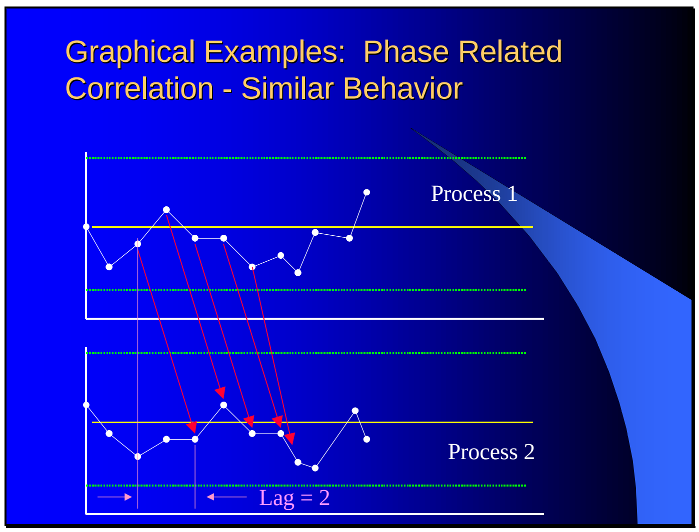
Slide 84



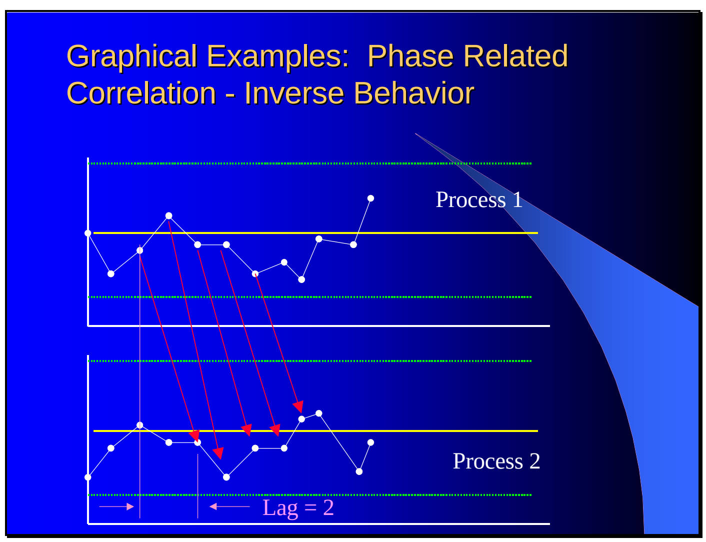
Slide 85



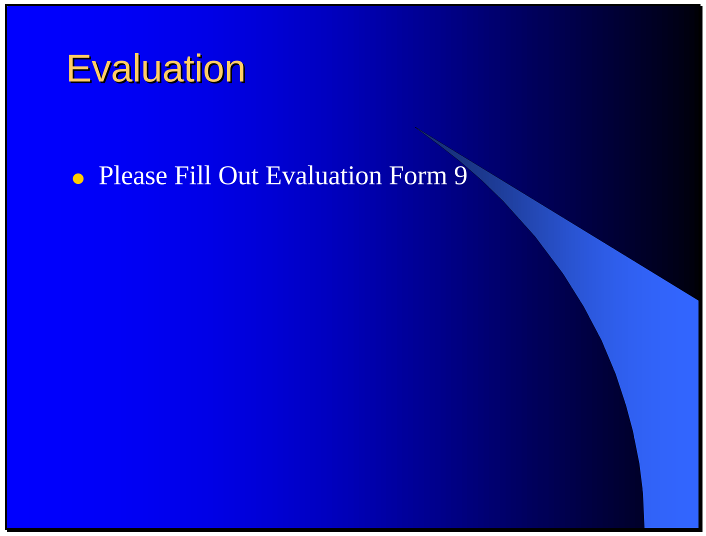
Slide 86



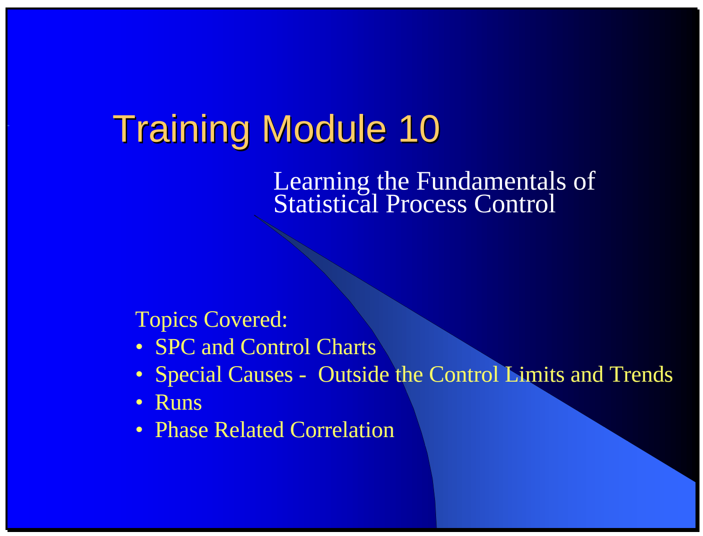
Slide 87



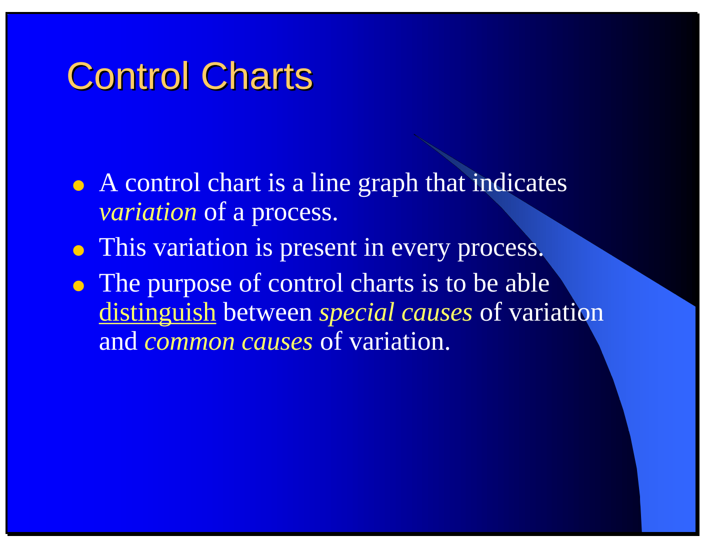
Slide 88



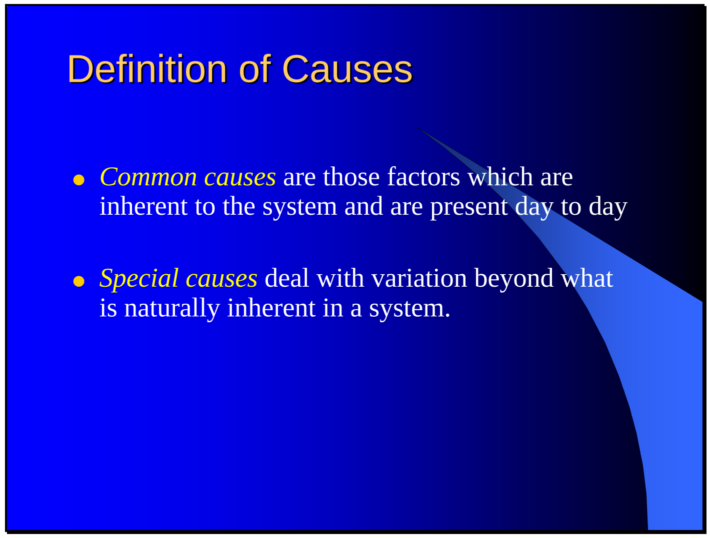
Slide 89



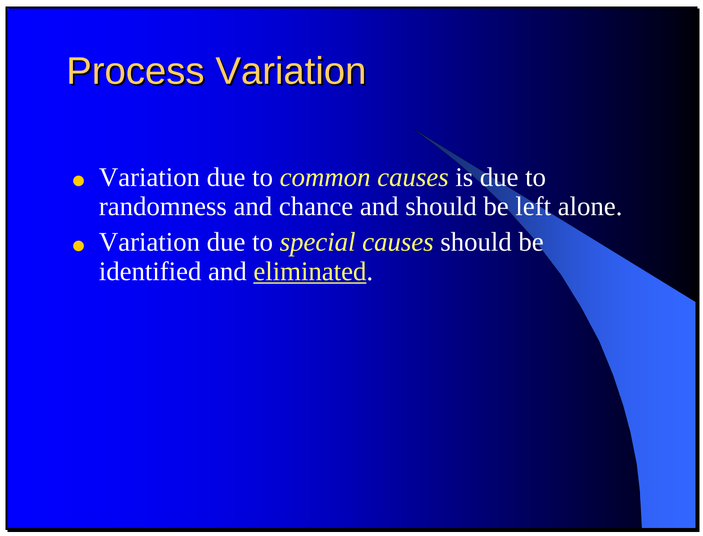
Slide 90



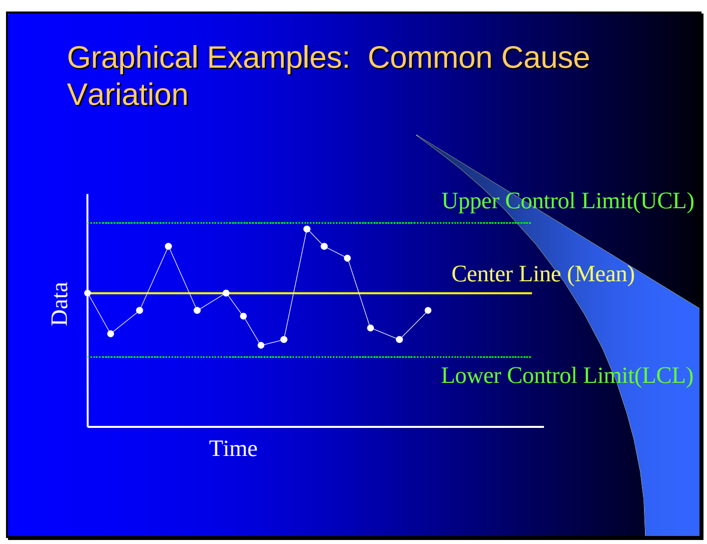
Slide 91



Slide 92



Slide 93



Slide 94

Statistical Control

- Processes are said to be in “statistical control” when the following conditions are met:
  - All data points fall within the control limits, and
  - The dispersion of data points appear to exhibit random behavior
- Statistical control is desirable since we can predict the future behavior of a process within the control limits
- Processes in statistical control should not be adjusted
- A process in statistical control still varies but only due to common causes

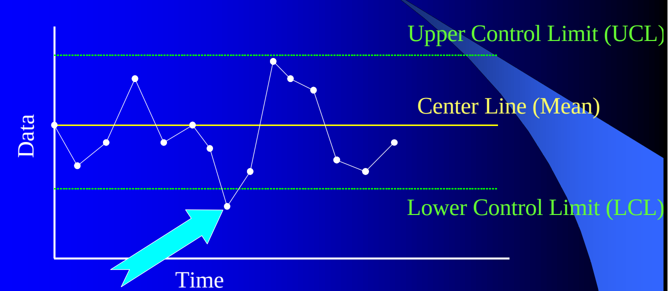
Slide 95

Special Causes

- Outside the control limits: a data point which lies above the upper control limit or below the lower control limit

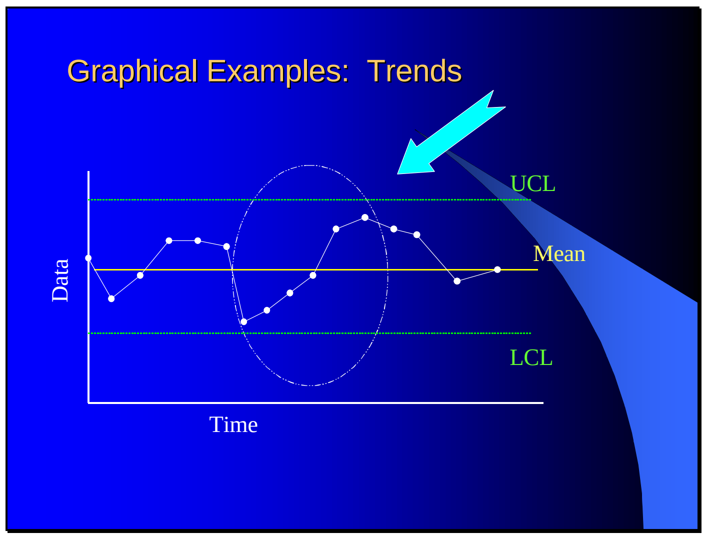
Slide 96

Graphical Examples: Outside the Control Limits

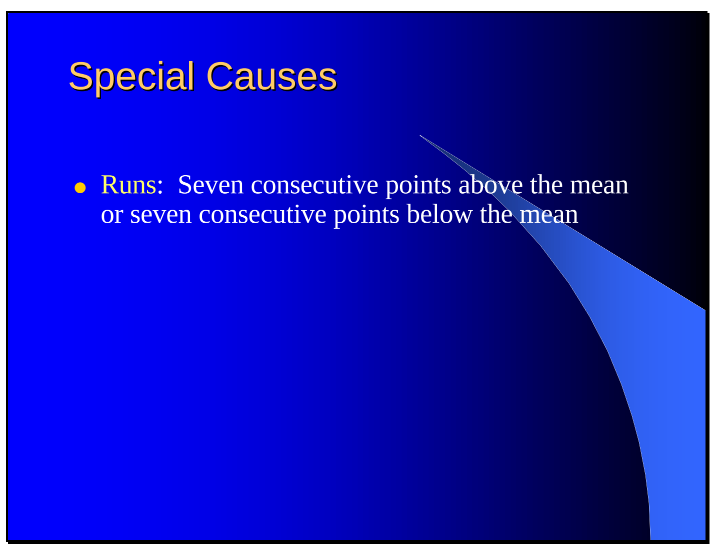


Slide 97

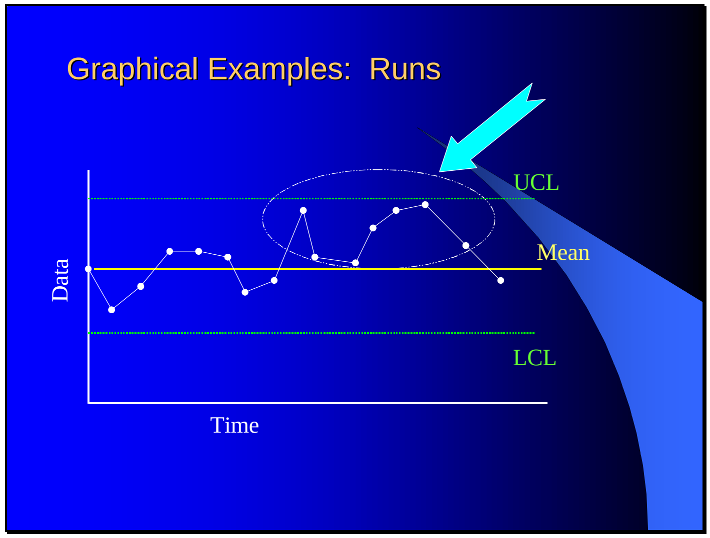
Slide 98



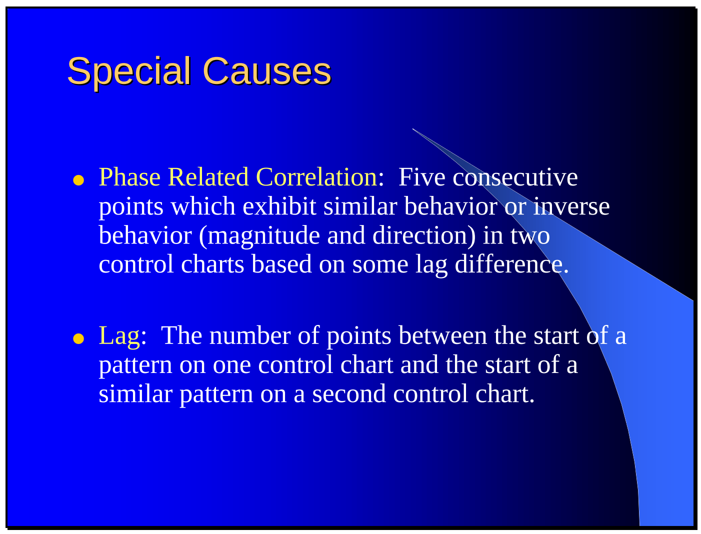
Slide 99



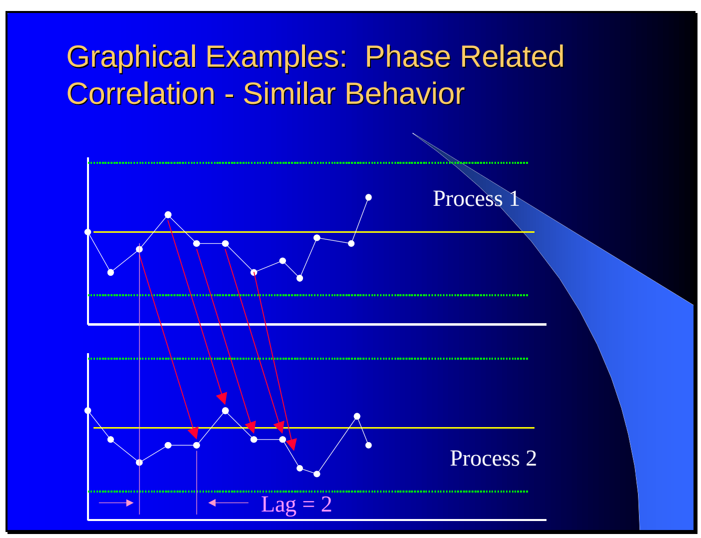
Slide 100



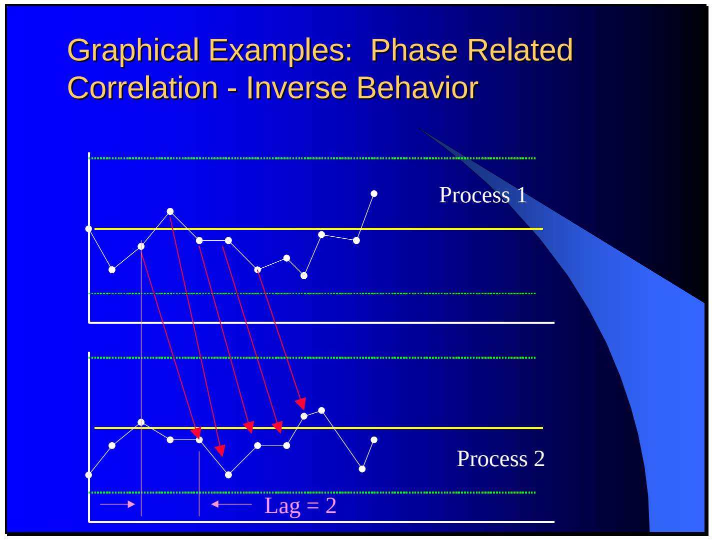
Slide 101



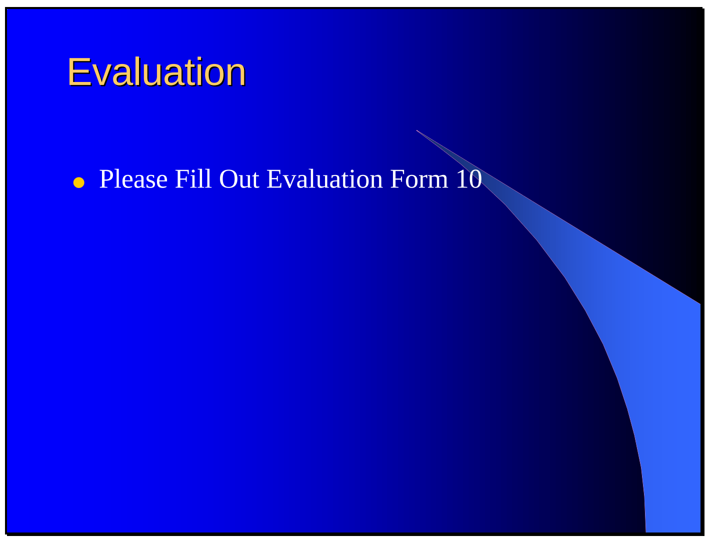
Slide 102



Slide 103



Slide 104



Appendix F2 Evaluation Forms

Evaluation Form 1

Multiple Choice

Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Variation is inherent in every process.
   A. True
   B. False

2. A control chart:
   A. Indicates variation in a process
   B. Distinguishes common causes and special causes
   C. Is a line graph
   D. All of the above

3. On the control chart, the mean of the process is represented by the control limits.
   A. True
   B. False

4. Statistical Control:
   A. Occurs when all data points fall within the control limits
   B. Indicates that the process varies due to common causes
   C. Occurs when the dispersion of data points appears to be random
   D. All of the above

5. Variation due to common causes should be eliminated, while variation due to special causes should be left alone.
   A. True
   B. False

Evaluation Form 2

Multiple Choice

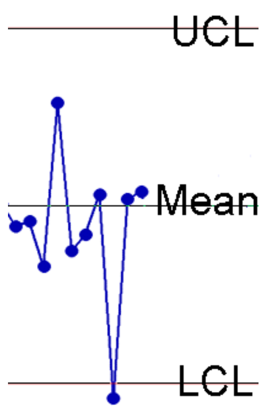
Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Outside the control limits and trends are the result of common causes.
   A. True
   B. False

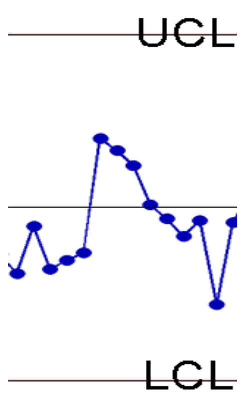
2. A trend occurs when ***five*** consecutive data points on a control chart are in a monotone increasing or decreasing pattern.
   A. True
   B. False

3. Outside the control limit occurs when a data point is near the mean.
   A. True
   B. False

4. The following control chart to the right represents:
   A. Outside the control limits
   B. Trend
   C. Random variation



5. The following control chart to the right represents:
   A. Outside the control limits
   B. Trend
   C. Random variation



Evaluation Form 3

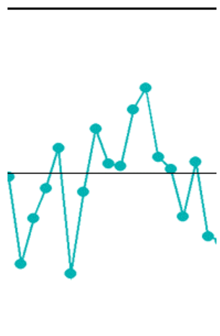
Multiple Choice

Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

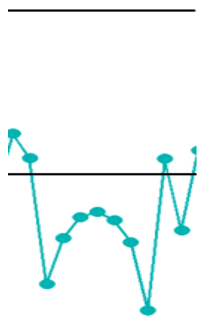
1. Runs are defined as six consecutive points above the mean or six consecutive points below the mean.
   - A. True
   - B. False

2. Runs are a type of special cause.
   - A. True
   - B. False

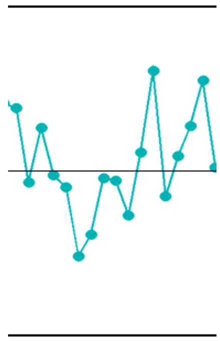
3. The following control chart to the right represents:
   - A. Run
   - B. Random Variation



4. The following control chart to the right represents:
   - A. Run
   - B. Random Variation



5. The following control chart to the right represents:
   - A. Run
   - B. Random Variation



Evaluation Form 4

<u>Multiple Choice</u>

Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Phase Related Correlation are defined as six consecutive points which exhibit similar behavior in two control charts based on some lag difference.
   A. True
   B. False

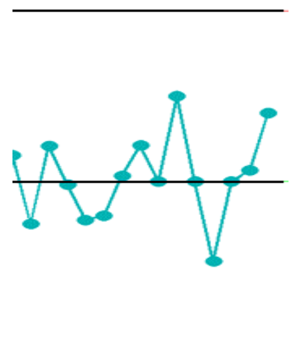
2. Lag is defined as the number of points between the start of a pattern in one control chart and the start of a similar pattern in a second control chart.
   A. True
   B. False

3. Phase Related Correlation is a type of special cause
   A. True
   B. False

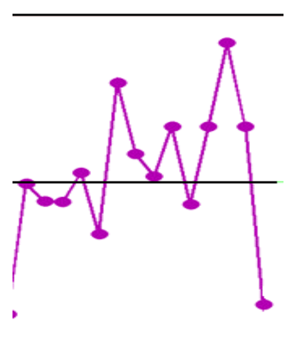
4. The following control charts to the right represent:
(Hint: Assume the two processes have a lag of two data points)

   A. Phase Related Correlation
   B. Random Variation

Process 1→



Process 2→



5. Phase Related Correlation can exhibit similar or inverse behavior:

   A. True
   B. False

Evaluation Form 5

<u>Multiple Choice</u>

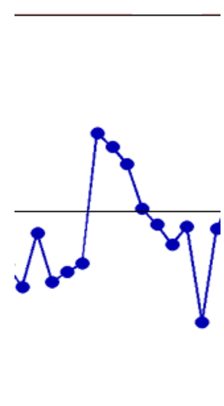
Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Variation is inherent in every process.
   - A. True
   - B. False

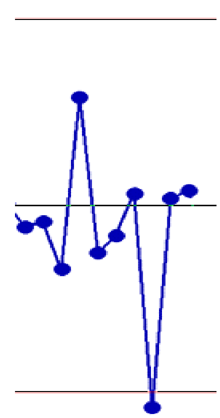
2. Statistical Control:
   - A. Occurs when all data points fall within the control limits
   - B. Indicates that the process varies due to common causes
   - C. Occurs when the dispersion of data points appears to be random
   - D. All of the above

3. When a data point is outside the control limits or when a trend occurs the causes of these events should be identified and eliminated.
   - A. True
   - B. False

4. The following control chart to the right represents:
   - A. Outside the control limits
   - B. Trend
   - C. Random variation



5. The following control chart to the right represents:
   - A. Outside the control limits
   - B. Trend
   - C. Random variation



Evaluation Form 6

Multiple Choice

Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Phase Related Correlation are defined as six consecutive points which exhibit similar behavior in two control charts based on some lag difference.
   A. True
   B. False

2. Lag is defined as the number of points between the start of a pattern in one control chart and the start of a similar pattern in a second control chart.
   A. True
   B. False

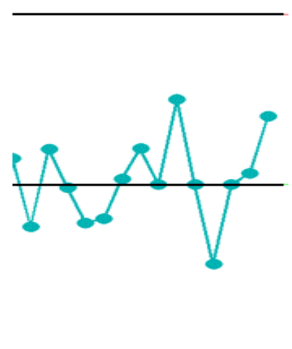
3. Phase Related Correlation and runs are types of common causes
   A. True
   B. False

4. Runs are defined as seven consecutive points above the mean or seven consecutive points below the mean.
   A. True
   B. False

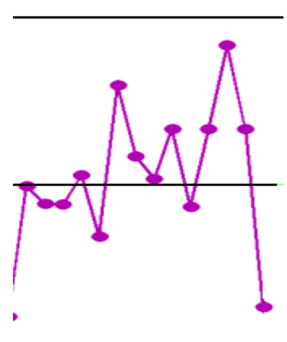
5. The following control charts to the right represent:
   (Hint: Assume the two processes have a lag of one data point)

   A. Phase Related Correlation
   B. Run in either graph
   C. Random Variation in both graphs

Process 1→



Process 2→

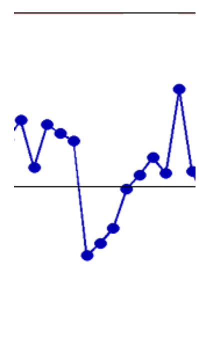


Evaluation Form 7

Multiple Choice

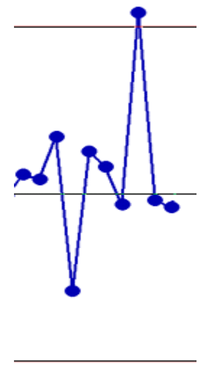
Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. An event where six consecutive data points on a control chart are in a monotone increasing or decreasing pattern is known as:
   A. Common cause
   B. Outside the Control Limits
   C. Run
   D. Trend

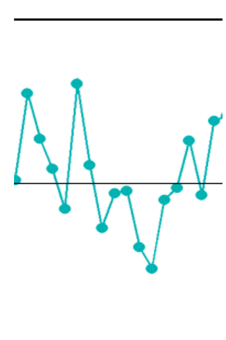


2. The following control chart to the right represents:
   A. Outside the control limits
   B. Run
   C. Trend

3. An event where a data point lies above the upper control limit or below the lower control limit is known as:
   A. Common cause
   B. Outside the Control Limits
   C. Run
   D. Trend



4. The following control chart to the right represents:
   A. Outside the control limits
   B. Run
   C. Trend



5. The following control chart to the right represents:
   A. Outside the control limits
   B. Run
   C. Trend

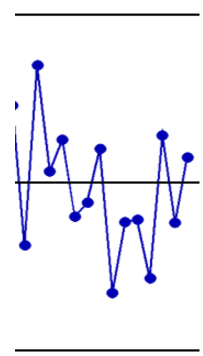
Evaluation Form 8

<u>Multiple Choice</u>

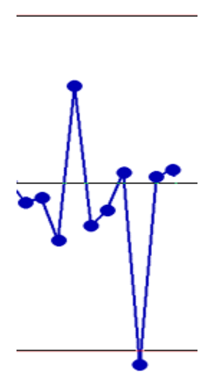
Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. Statistical Control:
    A. Occurs when a point is outside the control limits, or when a trend or run is present
    B. Occurs when all data points fall within the control limits and the points exhibit a random pattern
    C. Indicates that the process varies due to special causes
    D. All of the above

2. The following control chart to the right represents:
    A. Outside the control limits
    B. Run
    C. Trend
    D. Random variation

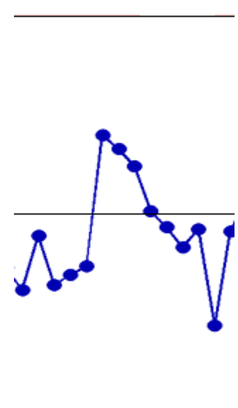


3. The following control chart to the right represents:
    A. Outside the control limits
    B. Run
    C. Trend
    D. Random variation



4. Variation due to common causes should be eliminated, while variation due to special causes should be left alone.
    A. True
    B. False

5. The following control chart to the right represents:
    A. Outside the control limits
    B. Run
    C. Trend
    D. Random variation

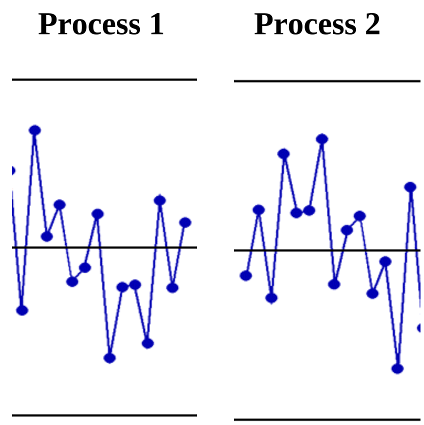


Evaluation Form 9

Multiple Choice

Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. The following control charts to the right represent: (Hint: Assume the two processes have a lag of two data points)



A. Outside the control limits (on either graph)
B. Run (on either graph)
C. Trend (on either graph)
D. Phase Related Correlation
E. Random Variation

2. Seven consecutive points above the mean or seven consecutive points below the mean is defined as:

A. Outside the control limits
B. Run
C. Trend
D. Phase related correlation

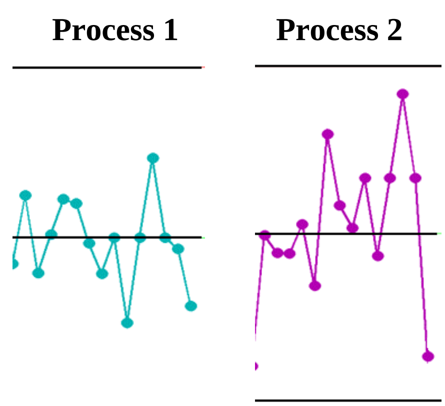
3. Six consecutive points which are in a monotone increasing or decreasing pattern is defined as:

A. Outside the control limits
B. Run
C. Trend
D. Phase related correlation

4. Five consecutive points which exhibit similar behavior in two control charts based on some lag difference is known as:

A. Outside the control limits
B. Run
C. Trend
D. Phase related correlation

5. The following control charts to the right represent: (Hint: Assume the two processes have a lag of one data point)

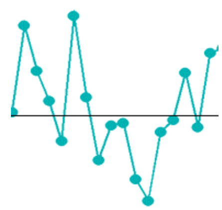


A. Outside the control limits (on either graph)
B. Run (on either graph)
C. Trend (on either graph)
D. Phase Related Correlation
E. Random Variation

Evaluation Form 10

Multiple Choice
Please answer the following questions regarding the training module you just completed. Circle the choice which best answers the question.

1. The following control chart to the right represents: ______________

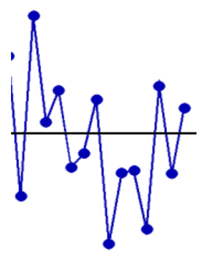


   A. Outside the control limits
   B. Random Variation
   C. Run
   D. Trend
   E. Phase Related Correlation

2. A process which exhibits only random variation is subject to: ______________

   A. Common Causes
   B. Special Causes
   C. None of the above
   D. All of the above

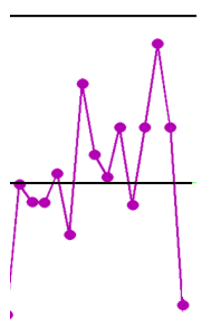
3. The following control chart to the right represents: ______________

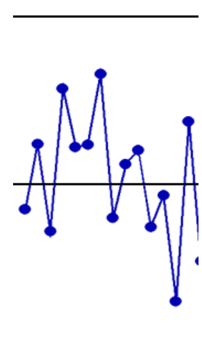


   A. Outside the control limits
   B. Phase related correlation
   C. Random Variation
   D. Run
   E. Trend

4. Five consecutive points which exhibit similar behavior in two control charts based on some lag difference is known as: ______________
   A. Outside the control limits
   B. Run
   C. Trend
   D. Phase related correlation

Process 1 **Process 2**





The following control charts to the right represent: (Hint: Assume the two processes have a lag of two data points)

   A. Outside the control limits (on either graph)
   B. Run (on either graph)
   C. Trend (on either graph)
   D. Phase Related Correlation
   E. Random Variation

Appendix G1 Data Sets used for Experiment

The following pages in this section represent the data sets used to create the control chart displays. There are four sets of data for each type of non-random signal. Each set of data is comprised of data for three processes used for the experiment.

Data for Outside the Control Limits

Point #	CL1A	CL1B	CL1C	CL2A	CL2B	CL2C
1	1.374	0.388	-0.859	1.122	-0.535	0.128
2	-1.130	-0.181	1.720	0.603	-0.242	-0.468
3	2.093	2.266	-0.737	-1.557	1.647	-0.503
4	0.201	0.647	0.326	0.881	0.554	-0.313
5	0.766	0.453	-0.851	0.078	0.205	-0.812
6	-0.612	0.061	1.138	0.030	-0.927	-0.542
7	-0.358	0.618	0.156	2.434	0.151	0.107
8	0.604	-0.183	-0.204	1.606	-2.092	1.453
9	-1.974	1.003	-0.020	-0.078	2.206	0.294
10	-0.710	1.382	-1.511	-1.639	-0.495	-0.747
11	-0.671	0.774	0.631	-0.637	-0.141	0.984
12	-1.722	-1.814	0.840	0.239	-1.048	-1.083
13	0.843	-1.173	-0.842	-0.915	-1.142	0.826
14	-0.722	-0.483	0.237	0.724	0.556	-0.900
15	0.446	0.040	-0.278	-0.306	0.365	-0.510
16	0.287	-0.528	0.339	1.916	-2.007	1.378
17	0.404	-0.088	0.110	0.546	-0.495	-0.779
18	-0.113	0.427	1.362	-0.175	0.854	0.142
19	1.223	0.244	0.864	1.206	-0.546	1.186
20	0.720	0.240	1.402	-0.450	0.728	-0.307
21	-0.844	-0.230	1.582	-0.340	-0.951	0.964
22	0.943	-0.156	-0.471	-0.252	-0.374	-0.552
23	1.536	-0.707	0.867	0.477	-1.488	-1.125
24	1.358	-0.362	-1.587	0.846	1.854	0.543
25	0.773	1.619	-1.327	-0.813	1.573	-0.713
26	-0.227	0.000	-0.044	0.548	2.186	-0.906
27	-0.957	-0.942	0.223	2.115	1.588	1.801
28	1.454	1.160	3.250	1.217	0.139	-0.539
29	-0.202	-0.104	0.698	0.289	-2.391	-1.970
30	0.782	-0.464	-0.054	0.913	3.250	0.386
31	-1.385	1.558	-0.929	-0.085	1.884	-0.154
32	0.859	-0.442	-0.483	1.858	-0.989	0.018
33	0.814	-0.395	1.306	-0.964	-0.816	0.862
34	0.064	1.781	0.431	-0.262	-0.821	0.147
35	-1.049	-1.242	-0.541	-0.091	-0.050	1.181

36	-0.321	0.631	-0.128	0.567	0.486	1.470
37	0.268	-0.935	0.323	-1.410	-1.126	0.618
38	0.494	1.433	-0.699	-0.412	0.794	-0.821
39	-2.092	-1.062	-1.262	-1.197	1.003	-0.549
40	-1.379	0.836	0.001	-0.161	-0.992	0.786
41	-0.543	1.116	-1.447	2.260	-0.775	-1.810
42	0.598	1.682	0.576	-0.065	0.070	-0.839
43	0.273	1.463	-0.848	1.394	-0.244	-0.091
44	-0.479	-0.302	-0.537	-0.968	0.833	0.188
45	-0.387	1.273	-1.345	0.159	0.978	-0.879
46	-0.565	0.349	0.494	2.541	1.313	-0.353
47	0.081	-0.779	0.314	0.762	-1.714	-0.221
48	0.186	0.693	0.757	-0.307	0.483	-0.749
49	-0.627	-0.741	0.109	0.596	-0.345	0.267
50	-0.773	0.980	-0.444	-0.168	-0.042	-0.661

Point #	CL3A	CL3B	CL3C	CL4A	CL4B	CL4C
1	1.714	0.090	-0.845	-0.074	0.758	0.388
2	-0.733	0.367	-0.680	-0.264	0.138	-1.985
3	0.621	-0.851	0.651	-1.000	0.223	0.268
4	-0.908	-1.135	1.010	0.758	-0.447	0.617
5	-0.203	-1.041	0.210	0.536	1.188	-0.278
6	0.285	-0.668	-0.946	0.814	1.109	-0.025
7	-0.771	-0.472	-0.219	0.355	0.175	-1.430
8	-0.706	0.357	0.108	-1.344	-0.907	0.496
9	-0.002	-0.128	1.427	1.922	0.898	1.037
10	0.760	1.722	-0.213	1.234	-1.828	-0.770
11	1.737	-1.129	1.938	1.734	-0.287	0.114
12	0.964	-0.194	0.833	-0.602	1.172	0.423
13	-0.687	-0.480	0.142	1.245	-0.012	0.785
14	1.480	0.809	0.240	0.390	-0.163	1.087
15	0.364	1.124	-0.458	0.300	0.644	0.726
16	0.248	0.164	0.227	0.046	-1.819	-1.911
17	-0.270	-0.133	0.121	-0.258	1.483	1.207
18	0.354	-0.232	0.640	0.026	1.457	0.852
19	-0.945	0.909	-0.248	0.598	-0.574	-0.959

20	0.499	0.908	0.079	-0.307	0.473	1.154
21	1.029	-0.238	-0.906	-0.618	-0.599	0.466
22	0.577	-1.668	-0.037	0.472	-0.182	-0.535
23	-0.823	-0.894	1.528	0.254	-0.536	-0.455
24	0.258	-0.484	0.386	1.191	-0.527	-0.621
25	-1.289	-0.271	-3.250	0.085	0.303	1.597
26	0.376	1.156	0.897	0.299	-1.022	1.230
27	-0.447	0.805	-0.324	0.176	-0.927	0.587
28	-0.452	0.115	-0.465	-1.593	-0.875	-0.889
29	0.249	-0.234	0.235	0.424	-0.090	-0.224
30	0.925	0.040	-0.844	-0.032	0.166	0.165
31	0.241	1.590	0.670	0.027	-2.092	-1.228
32	2.736	-0.562	-0.230	-0.347	0.214	0.659
33	-1.336	-0.567	0.202	-0.266	-0.880	-2.171
34	-1.106	0.043	0.134	-1.027	1.087	-0.865
35	0.425	-0.351	0.501	1.737	0.340	-1.782
36	0.233	0.473	0.272	-0.758	0.363	0.516
37	0.822	1.696	-0.869	-0.482	0.825	-1.162
38	1.587	0.055	-0.521	0.191	-0.676	-0.276
39	1.579	0.527	-0.850	-3.250	-0.100	0.576
40	-0.564	-2.002	-1.511	0.113	-1.088	-0.228
41	2.791	0.759	1.986	0.238	-1.318	1.175
42	0.544	1.088	-0.215	-0.785	-0.179	-1.196
43	-0.930	-0.250	1.274	-0.256	-0.698	0.356
44	1.323	0.812	0.090	-0.285	0.127	0.037
45	-0.211	-1.299	1.582	-1.324	-1.040	-0.483
46	0.115	-0.571	0.956	-0.173	0.410	-0.815
47	-0.118	1.082	0.362	0.832	-0.709	1.926
48	-0.764	-1.123	-1.242	2.471	-0.383	-0.919
49	0.138	-0.807	1.228	-0.028	1.729	1.099
50	-0.316	-0.906	0.188	-1.000	-1.212	-0.715

Data for Runs

Point #	RUN1A	RUN1B	RUN1C	RUN2A	RUN2B	RUN2C
1	0.086	1.077	0.329	-1.309	0.413	-0.316
2	0.113	0.691	-0.074	0.577	1.267	-0.513
3	0.269	-0.275	1.366	0.289	0.248	1.801

4	1.237	1.321	0.559	0.589	0.215	-0.373
5	-1.641	-0.278	-0.882	2.728	0.454	-1.036
6	0.612	-0.229	1.015	0.131	-0.923	0.276
7	-0.703	0.252	0.304	-1.954	-1.464	0.692
8	-1.458	-0.800	-0.875	-0.516	-0.306	-0.731
9	-0.896	0.616	-2.178	0.541	-1.062	-0.603
10	-1.272	0.324	-0.808	-0.644	0.791	-0.478
11	1.227	-1.328	0.034	-0.490	0.131	0.017
12	-0.151	-0.562	1.882	-0.444	2.066	-0.234
13	0.185	-0.225	-0.447	-0.304	-0.530	-0.005
14	-0.292	-1.509	0.318	0.681	0.029	2.056
15	-0.867	1.077	-0.306	-1.137	0.064	0.214
16	2.055	0.703	-0.195	0.120	0.182	-0.837
17	1.510	-0.363	1.525	0.829	-0.102	0.243
18	-1.507	0.194	0.457	0.201	-0.276	1.137
19	-0.236	-0.638	-0.929	-0.040	0.871	-0.542
20	-0.300	-0.555	-0.316	-0.591	0.827	-0.513
21	-0.805	0.584	-0.706	0.048	0.755	1.390
22	-0.204	2.149	2.304	-1.392	0.487	0.579
23	0.447	0.266	-2.666	-1.146	0.874	-2.676
24	1.705	0.502	1.054	0.277	1.190	-0.440
25	-0.559	0.924	-0.484	-0.455	0.695	0.150
26	-0.508	-2.363	-0.660	-2.576	-0.439	-0.367
27	0.394	-1.774	0.799	1.315	0.805	-1.161
28	1.559	0.861	1.666	-0.427	-1.262	-1.039
29	1.231	-0.001	0.333	-0.345	-0.614	-0.986
30	-0.604	-0.320	0.640	0.665	-0.146	1.317
31	-1.760	1.409	0.943	-0.197	1.952	-0.284
32	1.878	-0.815	2.090	0.486	-0.234	-0.075
33	1.669	-0.411	0.999	0.203	0.287	2.096
34	-0.258	0.859	-0.936	-0.579	0.472	-0.588
35	-0.321	0.599	0.513	0.886	-0.594	1.372
36	-0.797	-2.241	0.330	1.318	0.381	-0.439
37	-0.590	-0.333	0.729	0.919	-0.959	-0.238
38	-1.342	0.404	-0.279	-1.847	0.089	-1.416
39	0.142	-0.408	0.422	-0.770	-1.628	0.417
40	-0.765	-0.463	0.767	0.226	1.762	-0.222
41	-1.499	0.758	-1.886	-0.848	-0.821	0.023
42	0.146	0.162	0.499	0.308	-0.624	0.213

43	-0.245	0.596	-0.481	-0.985	-0.421	0.844
44	-1.881	-0.285	0.262	-0.163	-0.370	-0.553
45	-1.451	0.860	-0.039	-0.485	-0.739	0.098
46	-0.622	-0.070	1.751	0.309	-0.669	0.797
47	0.099	0.685	-0.415	0.416	0.628	-1.542
48	-0.077	-0.553	-0.893	-0.134	1.315	-0.366
49	1.045	0.098	-0.045	0.204	1.691	1.505
50	-1.226	0.797	-1.395	-1.955	-0.381	-1.498

Point #	RUN3A	RUN3B	RUN3C	RUN4A	RUN4B	RUN4C
1	-0.703	1.132	0.513	1.216	1.110	-1.300
2	1.792	0.793	-1.501	-0.705	-0.359	1.264
3	-1.304	0.700	-2.168	0.131	-0.123	-1.371
4	-0.451	-0.249	-1.557	-0.652	0.903	0.305
5	1.216	-0.251	0.420	0.595	-1.892	-0.445
6	1.160	0.202	-1.586	-0.839	-0.900	0.322
7	0.944	-0.132	-0.822	-0.592	0.826	0.015
8	-0.324	-0.856	-1.115	-0.182	-0.100	1.505
9	1.626	-0.244	0.086	0.363	0.075	-0.765
10	1.103	-0.453	1.300	0.750	-2.227	-0.214
11	0.117	-0.243	0.153	-2.197	0.807	0.498
12	-0.141	-0.278	0.859	0.155	-0.175	-1.941
13	-1.316	1.075	-1.313	-1.183	1.240	-0.629
14	0.660	-1.239	0.485	-0.494	1.145	-0.405
15	-0.699	-0.240	0.390	-0.463	-0.203	0.567
16	-0.257	1.384	0.992	-1.564	0.789	-1.124
17	-0.100	1.264	-1.465	-1.462	-0.072	-1.990
18	2.107	-1.554	-0.211	-0.109	-0.294	0.120
19	0.486	-0.805	-0.172	1.435	-1.547	-0.569
20	0.239	0.077	-1.241	-0.837	-1.161	1.325
21	-0.461	-1.692	-1.439	-0.991	-0.128	-0.626
22	0.562	-0.638	0.448	1.138	-0.179	0.479
23	-0.317	-0.937	-1.680	0.187	-0.806	-0.696
24	-1.088	-1.476	0.417	-0.911	0.336	2.035
25	-0.135	-0.154	-1.170	-0.143	1.825	0.399
26	-0.114	0.807	1.637	-0.979	-0.461	-1.930
27	-1.685	1.478	-0.161	0.089	0.276	0.188

28	0.098	0.462	-1.005	-0.359	0.822	-0.010
29	-1.148	1.128	1.830	2.652	1.649	-0.047
30	0.818	-0.227	0.414	-0.544	0.070	1.437
31	-1.687	0.176	-1.298	0.796	0.469	-0.469
32	0.383	-1.619	-1.911	0.660	-0.788	1.262
33	0.314	-0.986	0.511	0.176	0.074	0.220
34	-0.915	1.670	0.155	-0.097	0.279	0.930
35	0.751	-0.364	-1.280	-0.108	-0.953	-0.528
36	-0.546	0.741	-0.926	0.244	-0.220	0.405
37	0.327	0.296	-0.130	-1.206	-0.029	1.220
38	1.213	-2.012	0.062	0.169	-1.767	0.090
39	0.235	-1.171	-0.490	0.507	2.007	0.098
40	0.098	-0.788	0.245	-1.233	-0.515	0.860
41	-0.527	-0.693	-0.607	-0.863	-0.852	-0.148
42	-0.563	-0.837	0.171	0.299	-1.274	1.475
43	-2.109	-1.242	-0.396	0.957	1.848	0.158
44	0.836	-2.489	0.090	-0.227	-0.167	-2.252
45	0.629	0.279	0.324	1.477	-0.382	-1.888
46	-2.517	-1.029	2.025	0.491	-0.758	0.034
47	-1.256	0.433	-0.543	-1.270	0.766	0.373
48	0.428	1.341	0.104	-1.285	1.081	0.259
49	-0.212	0.231	-0.204	-0.457	0.314	-0.396
50	-0.135	-0.900	-1.302	-0.939	2.068	0.090

Data for Trends

Point #	TREND1A	TREND1B	TREND1C	TREND2A	TREND2B	TREND2C
1	0.091	1.834	0.365	0.046	0.488	-0.666
2	-0.584	-0.293	0.465	1.394	-0.942	-0.764
3	0.968	-0.735	1.588	-0.157	-0.807	-0.547
4	-1.388	-0.686	0.099	0.849	-1.067	0.115
5	0.118	-0.979	-0.023	-1.682	1.395	0.804
6	-0.033	-0.921	-0.886	-1.507	-0.366	0.879
7	1.304	-0.212	-0.771	-0.031	0.358	-0.723
8	-2.293	1.564	-0.109	0.561	2.429	-1.049
9	-2.742	-0.092	1.308	-0.349	1.202	0.147
10	-0.322	0.579	0.365	1.293	-0.371	0.231

11	-0.514	0.246	-0.557	-0.820	1.117	-1.190
12	-0.972	0.281	-0.189	1.522	-0.145	-0.915
13	-0.148	-0.199	0.635	2.593	-1.413	-1.358
14	0.041	-0.550	-0.435	0.204	-2.492	2.141
15	-0.163	0.058	0.898	-0.473	-0.397	-0.630
16	0.685	1.243	0.253	-0.467	0.136	1.719
17	0.654	-1.152	-2.009	0.514	1.126	-0.086
18	-1.187	-1.044	-0.128	-1.699	0.631	0.329
19	-1.322	0.466	0.260	-1.104	0.402	0.057
20	-0.485	0.191	0.481	-1.832	-0.306	2.080
21	0.368	0.046	-0.758	0.954	-0.453	-2.124
22	-2.357	0.510	-0.661	-0.116	-0.627	-2.542
23	-0.110	0.592	-1.457	-0.785	-0.297	-0.520
24	-0.863	-0.447	1.283	-0.269	0.631	-0.586
25	1.142	1.268	1.357	-1.892	2.402	2.144
26	1.150	0.742	-0.005	1.352	-0.306	-0.232
27	-0.768	2.566	0.084	0.580	0.453	0.200
28	1.116	0.619	-0.225	-0.879	0.728	-0.624
29	1.558	0.253	-0.207	0.548	0.684	-2.144
30	-0.283	0.437	0.344	-1.521	0.306	0.194
31	-1.089	-1.291	1.461	-1.496	-0.352	-1.626
32	-0.816	1.045	0.368	-0.434	-1.666	2.188
33	-1.063	0.301	-0.594	0.548	-0.039	0.541
34	-0.006	-0.678	1.108	-0.799	-2.516	0.944
35	-0.776	0.492	-1.197	0.434	1.931	0.808
36	0.912	0.281	1.048	-1.325	-0.097	-0.842
37	-0.079	0.021	1.947	0.458	-1.352	0.714
38	0.265	-0.653	-0.661	0.565	1.142	0.479
39	-0.208	-0.900	-1.475	-0.253	-0.519	0.141
40	1.505	-1.200	-0.223	0.452	0.478	-1.121
41	0.888	0.692	0.580	-0.642	0.457	0.680
42	1.212	1.281	1.915	-0.011	1.966	0.005
43	-1.439	0.021	-1.260	0.294	2.296	0.105
44	0.482	-0.653	0.439	0.024	0.021	0.134
45	1.703	-1.456	-1.189	0.147	-0.722	-0.630
46	-2.225	-0.265	0.288	0.479	-1.333	-0.065
47	-0.135	0.752	-0.179	0.966	0.652	0.704
48	0.222	-0.597	0.061	0.737	-0.442	0.153
49	-0.034	-0.587	-1.054	0.076	0.548	-0.367

50	0.837	-0.734	0.370	-0.018	-0.604	-0.281

Point #	TREND3A	TREND3B	TREND3C	TREND4A	TREND4B	TREND4C
1	0.063	1.204	-0.136	1.668	-0.648	0.370
2	0.377	0.002	0.584	0.693	0.334	0.821
3	-0.787	0.857	0.460	-0.791	-0.596	-0.763
4	-0.430	2.597	-0.945	0.667	1.496	-1.042
5	1.075	0.300	-1.093	0.813	0.857	-1.289
6	0.614	-0.014	1.103	-0.453	0.598	-0.066
7	-0.482	0.099	1.842	-1.423	-0.993	-0.960
8	-1.055	1.116	-0.367	-0.324	0.979	0.209
9	-1.227	-1.745	-0.831	-0.852	1.209	-0.424
10	-1.166	0.746	0.837	1.007	0.542	-0.588
11	0.362	-1.522	-0.015	-1.940	-1.065	-2.136
12	0.568	-1.475	-0.202	2.624	1.606	0.742
13	0.072	-1.135	0.121	0.068	2.621	0.810
14	0.649	0.640	2.081	-0.807	0.399	0.630
15	-0.396	0.329	0.537	-1.151	-1.643	0.421
16	-0.609	1.870	-1.753	-0.326	1.205	-1.393
17	-0.023	-0.605	2.439	-1.072	0.496	-2.495
18	0.308	1.216	1.043	-0.920	-0.203	-0.971
19	0.073	0.131	1.134	-0.790	-0.472	-0.567
20	-0.910	-0.430	-1.587	1.192	-1.191	-1.275
21	0.779	0.700	-0.876	0.981	1.048	-1.195
22	-0.889	-2.449	-0.392	0.721	-0.207	2.745
23	-0.396	-0.671	-0.400	0.047	-1.040	-0.379
24	0.417	-0.611	0.670	-0.200	-0.381	0.160
25	-0.353	-0.194	0.267	-0.500	0.843	-0.989
26	0.433	1.047	-0.179	-0.225	1.258	-0.879
27	-0.998	-0.291	-0.938	-1.682	0.899	-1.808
28	0.140	-0.190	-1.720	-0.261	-0.249	-2.581
29	0.133	-0.207	-0.131	0.269	-0.728	1.304
30	0.686	0.258	-1.574	-1.114	-0.632	0.343
31	-1.921	1.152	-0.852	0.114	0.460	-0.376
32	-0.583	-1.622	0.866	-0.814	-0.875	-0.901
33	0.139	0.007	-1.016	-0.447	1.356	-0.475
34	-0.107	-0.295	0.132	2.003	-0.664	-1.811

35	0.275	-1.373	0.397	0.547	1.827	0.196
36	1.382	1.664	1.286	-1.088	-0.504	-0.668
37	-0.778	-0.598	0.535	-0.560	-0.044	-0.041
38	-0.765	2.821	-0.213	-1.251	0.672	-0.216
39	-0.048	-0.571	0.569	1.517	0.067	-2.105
40	1.646	0.875	-0.327	1.406	0.452	1.205
41	0.884	0.641	-0.153	-1.219	1.116	0.505
42	0.543	1.507	-0.006	-1.225	-1.202	0.633
43	1.420	0.299	0.702	0.478	-0.173	1.806
44	-0.190	-0.347	0.931	-1.336	0.326	0.506
45	-1.411	-0.701	1.426	-0.823	-0.497	-0.592
46	1.053	-1.648	-0.454	-1.020	-0.565	-0.286
47	-0.065	1.799	0.255	1.430	1.049	0.562
48	-0.038	-0.384	-0.576	0.952	-0.968	0.570
49	0.145	-1.193	1.576	2.040	0.396	-0.029
50	-0.944	1.315	0.932	-0.180	0.301	-0.048

Data for Phase Related Correlation

Point #	PHASE1A	PHASE1B	PHASE1C	PHASE2A	PHASE2B	PHASE2C
1	-1.490	0.666	-0.275	-0.409	-0.386	0.630
2	1.295	-0.042	-0.328	-0.775	-0.881	1.494
3	-0.968	-0.112	0.077	-0.257	0.133	-0.147
4	-0.899	-0.793	-0.090	1.100	-0.199	0.173
5	0.621	-0.175	0.153	-1.600	-0.405	0.763
6	0.982	-0.853	1.594	-0.701	-0.445	-0.969
7	-2.470	-1.000	0.658	1.123	0.057	-0.376
8	0.452	0.523	0.786	-0.013	-1.810	-1.160
9	-0.183	-0.608	-0.373	-0.350	1.543	0.505
10	-0.626	-0.689	-0.842	-0.641	-0.176	0.218
11	0.269	0.102	-2.367	-1.121	-0.979	1.076
12	-0.259	0.368	-0.773	0.434	-2.098	1.090
13	1.234	0.223	0.118	0.215	0.966	-1.098
14	0.325	-1.183	-0.754	0.770	0.259	1.100
15	0.434	-0.373	-0.145	2.551	-0.298	0.858
16	0.229	-0.615	0.670	-0.983	-1.448	0.847
17	-0.286	0.394	-0.442	-0.230	-0.482	-0.169
18	0.000	-2.074	-0.599	-0.095	0.804	-0.957

19	1.500	-1.090	0.628	-0.018	-0.029	0.998
20	0.000	1.500	-0.434	0.174	0.873	1.479
21	1.500	1.900	2.281	0.086	0.796	0.320
22	0.000	1.000	0.000	-0.873	0.893	-0.130
23	-1.400	2.500	1.500	-0.073	0.009	0.532
24	0.000	1.000	0.000	-1.612	-0.951	-0.450
25	0.200	-0.400	-1.400	-0.062	-1.365	0.436
26	1.200	1.000	0.000	-0.402	-0.095	1.310
27	0.000	1.335	-0.667	1.342	0.138	-0.895
28	-1.400	1.281	-0.610	0.000	-0.500	-1.356
29	0.000	1.432	0.543	1.500	-0.900	-0.176
30	-1.130	-0.992	-0.598	0.000	0.000	-0.642
31	1.619	-1.202	0.549	-1.400	-1.400	-0.626
32	-0.824	0.336	0.460	0.000	0.000	0.635
33	-0.041	-0.320	-1.090	0.200	1.500	-1.669
34	-0.893	1.216	-0.541	1.200	0.000	1.318
35	-0.906	-0.777	0.451	-0.249	1.511	-0.662
36	1.493	0.739	-0.403	-0.427	-0.750	0.000
37	-0.031	-0.262	-0.707	-1.626	-0.781	-0.977
38	-2.404	-1.252	0.223	-1.083	2.207	-0.126
39	0.840	1.493	0.872	-0.027	1.433	-0.977
40	0.782	-0.432	0.907	0.346	-0.886	-0.409
41	-0.932	-0.397	-0.986	0.664	-0.067	1.876
42	-1.388	-0.043	-0.570	-0.006	1.058	1.420
43	-0.344	0.078	1.374	-2.486	0.962	-0.362
44	0.351	0.374	0.048	-1.415	-1.232	0.032
45	0.915	-0.544	0.291	0.212	-0.626	-0.217
46	1.741	0.484	-1.846	2.777	-0.147	1.760
47	1.635	1.305	1.053	1.476	1.611	-1.734
48	-0.816	-0.244	-0.227	1.153	0.919	0.884
49	-0.693	0.248	-0.526	-0.824	-0.043	0.079
50	0.202	-1.556	-0.796	2.657	0.868	0.756

Point #	PHASE3A	PHASE3B	PHASE3C	PHASE4A	PHASE4B	PHASE4C
1	0.724	-0.784	-2.243	-0.307	1.059	1.234
2	0.959	0.815	-0.342	0.962	-0.325	0.266
3	1.699	-0.506	-0.307	1.264	-0.228	0.123

4	-1.006	1.219	0.326	0.708	-0.294	-0.769
5	-0.976	-0.373	0.942	-0.810	0.246	1.040
6	-1.709	-0.505	-0.062	-0.619	-0.053	0.229
7	1.232	-0.389	0.950	-0.297	0.774	-0.071
8	0.041	-0.712	0.413	-0.979	-0.163	1.157
9	1.806	-0.148	-0.032	0.593	0.035	0.263
10	1.131	0.921	-0.355	-0.309	1.454	-0.998
11	-1.225	0.321	-1.777	-1.261	0.179	-0.535
12	1.352	0.639	-0.983	-0.992	-0.506	-0.792
13	-1.721	-0.865	-1.032	1.618	1.491	1.043
14	0.803	-0.815	0.121	0.508	-0.225	-0.581
15	-1.248	0.420	0.179	-0.960	-2.359	-0.507
16	-1.213	0.173	0.296	0.563	-0.071	-0.491
17	-2.751	0.023	0.740	-1.173	-2.144	1.641
18	0.615	-0.508	0.572	-0.095	0.461	0.029
19	-0.073	-0.742	-0.160	-1.402	-0.742	-2.386
20	0.581	2.265	-0.274	-0.968	0.625	-0.031
21	1.335	-1.434	0.132	1.775	-0.051	-0.351
22	1.281	1.169	1.426	-0.080	-0.673	-0.361
23	1.432	0.266	-0.840	-0.623	-0.602	0.160
24	-0.992	0.462	1.688	-0.064	0.096	-0.935
25	-1.202	0.000	-0.500	-0.760	0.631	1.781
26	0.336	1.500	-0.100	0.436	0.000	0.500
27	0.132	0.000	-1.000	0.236	1.500	0.100
28	-2.512	-1.400	0.500	-1.326	0.000	1.000
29	-0.359	0.000	-1.000	1.236	-1.400	-0.400
30	-0.650	0.200	-2.400	-1.457	0.000	1.000
31	0.750	1.200	-1.000	1.776	0.200	2.500
32	0.363	-0.633	1.664	-1.525	1.200	1.000
33	2.578	-0.740	0.781	2.227	0.365	-2.206
34	0.131	0.719	-0.960	0.411	-0.015	-0.797
35	-0.335	0.532	1.222	0.267	-0.159	-0.259
36	1.373	1.024	0.439	-0.726	0.434	1.490
37	0.342	0.565	-0.563	2.146	-0.886	1.891
38	-0.349	-1.505	1.056	0.731	0.968	-0.177
39	1.578	0.642	0.380	-0.456	-0.546	-0.391
40	0.666	-0.627	0.377	2.322	0.971	-0.414
41	0.571	0.560	-1.160	0.253	0.306	1.646
42	-0.423	1.257	-0.737	-0.141	-1.288	-0.147

43	1.467	-0.529	-0.940	-0.245	-0.271	1.166
44	-2.307	0.810	-0.476	-0.748	-0.630	-0.855
45	-0.595	1.090	-0.221	0.690	-0.801	-0.469
46	0.883	-0.481	0.419	0.414	-1.353	-2.429
47	2.019	0.156	0.031	-0.383	-0.588	-1.134
48	-2.562	-0.098	-1.730	1.750	2.120	-0.419
49	0.529	-1.861	-0.663	-2.231	2.748	-0.084
50	-0.292	0.765	0.989	0.418	-0.130	0.898

Appendix H1 Raw Data Sets

This section includes the following sets of raw data:

- Raw data from pre-experimental questionnaire
- Raw data for workload rating scores
- Raw data for objective decision making measures
- Raw data for subjective ratings for displays
- Raw data for Signal Detection Theory measures

To denote the type of display used for each subject, a coding scheme was created to facilitate the analysis of the data. The following represents the coding scheme used in the raw data: 1 = multiple two dimensional display, 2 = composite two dimensional display, 3 = composite three dimensional perspective display.

Raw data from pre-experimental questionnaire

Subject	Display	Academic	SPC	Visualization	Internet	CubeComp
1	1	4	2	1	2	32
1	2	4	2	1	2	32
1	3	4	2	1	2	32
2	1	4	2	2	3	3
2	2	4	2	2	3	3
2	3	4	2	2	3	3
3	1	4	2	1	3	18
3	2	4	2	1	3	18
3	3	4	2	1	3	18
4	1	1	2	1	4	26
4	2	1	2	1	4	26
4	3	1	2	1	4	26
5	1	1	1	2	3	12
5	2	1	1	2	3	12
5	3	1	1	2	3	12
6	1	1	1	3	3	9
6	2	1	1	3	3	9
6	3	1	1	3	3	9
7	1	1	2	3	4	31
7	2	1	2	3	4	31
7	3	1	2	3	4	31
8	1	2	2	2	4	33
8	2	2	2	2	4	33
8	3	2	2	2	4	33
9	1	1	1	3	4	30
9	2	1	1	3	4	30
9	3	1	1	3	4	30
10	1	1	2	3	4	41
10	2	1	2	3	4	41
10	3	1	2	3	4	41
11	1	2	1	2	3	35
11	2	2	1	2	3	35
11	3	2	1	2	3	35
12	1	2	2	3	4	22
12	2	2	2	3	4	22
12	3	2	2	3	4	22

Raw data for workload rating scores

Subject	Display	Mental	Physical	Temporal	ObsPerf	Frustrat	Effort	TotalWL
1	1	90	0	28	30	10	48	206
1	2	85	0	30	30	9	36	190
1	3	80	0	28	24	10	36	178
2	1	40	30	45	24	0	20	159
2	2	60	45	30	36	0	28	199
2	3	48	36	48	36	0	22	190
3	1	32	30	40	40	0	10	152
3	2	60	40	40	20	0	15	175
3	3	40	40	60	40	0	15	195
4	1	75	0	42	15	40	24	196
4	2	80	0	30	15	40	24	189
4	3	80	0	45	12	56	28	221
5	1	25	0	15	30	5	20	95
5	2	75	0	30	20	15	72	212
5	3	75	0	54	36	18	64	247
6	1	36	60	36	45	0	12	189
6	2	54	90	54	21	0	18	237
6	3	45	75	30	30	0	14	194
7	1	15	40	54	40	0	36	185
7	2	18	60	54	60	0	38	230
7	3	15	40	42	85	0	32	214
8	1	60	0	48	45	30	13	196
8	2	60	0	48	30	48	13	199
8	3	72	0	68	24	60	15	239
9	1	20	0	50	20	5	15	110
9	2	60	0	50	20	15	45	190
9	3	40	0	50	20	10	45	165
10	1	15	0	24	21	45	28	133
10	2	15	0	40	30	35	22	142
10	3	15	0	52	30	70	30	197
11	1	25	0	12	38	3	15	93
11	2	50	0	28	14	12	33	137
11	3	75	0	72	0	16	57	220
12	1	25	45	2	57	0	15	144
12	2	90	54	15	30	0	57	246
12	3	75	45	10	45	0	45	220

Raw data for objective decision making measures

Subject	Display	Search Time	Decision Time	Search+Decision Time	Stopping Time	Accuracy
1	1	5.83	4.50	9.50	21.05	0.750
1	2	6.60	3.80	10.40	26.42	0.625
1	3	3.00	4.40	7.40	29.05	0.625
2	1	8.50	4.25	12.75	23.35	0.500
2	2	9.00	3.75	12.75	24.95	0.500
2	3	9.75	3.00	12.75	24.08	0.500
3	1	6.20	4.60	10.80	20.97	0.625
3	2	9.50	6.00	15.50	26.97	0.500
3	3	16.50	4.00	20.50	20.80	0.250
4	1	7.86	4.43	12.29	28.35	0.875
4	2	7.80	4.80	12.60	25.33	0.625
4	3	5.17	4.33	9.50	25.83	0.750
5	1	5.20	3.80	9.00	26.85	0.625
5	2	4.40	4.00	8.40	20.92	0.625
5	3	8.00	3.00	11.00	14.77	0.125
6	1	5.60	4.20	9.80	22.08	0.625
6	2	3.00	4.00	7.00	22.30	0.250
6	3	6.00	4.50	10.50	14.88	0.250
7	1	8.20	4.80	13.00	21.70	0.625
7	2	10.60	3.60	14.20	27.00	0.625
7	3	12.14	3.29	15.43	25.85	0.875
8	1	6.50	5.67	12.17	27.42	0.750
8	2	4.00	4.33	8.33	22.03	0.375
8	3	7.80	3.80	11.60	22.35	0.625
9	1	10.60	3.60	14.20	17.88	0.625
9	2	10.00	4.00	14.00	19.73	0.375
9	3	6.00	5.00	11.00	17.07	0.125
10	1	7.00	4.00	11.00	20.52	0.750
10	2	5.83	4.33	10.17	26.10	0.750
10	3	6.20	3.80	10.00	25.55	0.625
11	1	13.80	6.60	20.40	28.70	0.625
11	2	7.33	6.33	13.67	26.47	0.375
11	3	10.00	5.50	15.50	28.83	0.250
12	1	5.33	4.00	9.33	24.52	0.750
12	2	4.00	4.75	8.75	30.12	0.500
12	3	8.00	6.33	14.33	13.33	0.375

Raw data for subjective ratings for displays

Subject	Display	SDecConf	SDecQuik	SDecAcc
1	1	3	3	2
1	2	2	1	2
1	3	2	2	2
2	1	3	4	3
2	2	3	4	3
2	3	3	4	4
3	1	4	4	3
3	2	1	1	2
3	3	2	2	2
4	1	5	4	4
4	2	4	4	4
4	3	3	3	3
5	1	4	4	4
5	2	4	4	3
5	3	2	2	2
6	1	4	4	4
6	2	1	2	1
6	3	3	4	2
7	1	4	5	4
7	2	4	4	3
7	3	4	4	4
8	1	4	5	4
8	2	3	4	2
8	3	2	2	1
9	1	4	4	4
9	2	4	4	2
9	3	2	2	2
10	1	4	3	3
10	2	2	3	2
10	3	1	2	2
11	1	5	4	5
11	2	2	2	2
11	3	1	1	1
12	1	4	5	4
12	2	2	2	2
12	3	2	2	2

Raw data for Signal Detection Theory measures

Subject	Display	Hit	FalseAlm	d'	β
1	1	0.667	0.051	2.06	3.51
1	2	0.500	0.251	0.67	1.25
1	3	1.000	0.217	3.86	0.01
2	1	0.500	0.203	3.09	1.43
2	2	0.364	0.186	0.52	1.38
2	3	0.444	0.196	0.69	1.41
3	1	0.500	0.130	3.09	1.88
3	2	0.400	0.211	0.55	1.34
3	3	0.250	0.225	0.06	1.05
4	1	0.438	0.071	1.33	2.95
4	2	0.500	0.172	0.95	1.58
4	3	0.600	0.105	1.54	2.19
5	1	0.556	0.204	0.99	1.41
5	2	0.714	0.055	2.11	2.87
5	3	0.500	0.093	3.09	2.46
6	1	0.714	0.032	2.43	5.03
6	2	1.000	0.108	4.32	0.02
6	3	0.667	0.055	2.00	3.04
7	1	0.455	0.172	0.83	1.57
7	2	0.333	0.203	0.40	1.29
7	3	0.318	0.067	1.01	2.67
8	1	0.500	0.155	0.99	1.64
8	2	0.750	0.178	1.59	1.21
8	3	0.455	0.157	0.87	1.63
9	1	0.385	0.085	1.06	2.37
9	2	0.375	0.091	1.04	2.35
9	3	0.500	0.180	0.92	1.52
10	1	0.429	0.136	0.90	1.76
10	2	0.660	0.092	1.75	2.27
10	3	0.500	0.168	0.95	1.58
11	1	0.313	0.145	0.54	1.52
11	2	0.429	0.272	0.44	1.19
11	3	0.400	0.302	0.27	1.11
12	1	0.667	0.124	1.62	1.81
12	2	0.800	0.212	1.65	0.97
12	3	0.500	0.078	1.41	2.68

Vita

SOMCHART THEPVONGS

Education:

December 1998 **Virginia Polytechnic Institute and State University**, Blacksburg, VA

Master of Science in Industrial & Systems Engineering
Concentration: ***Management Systems Engineering***

December 1995 **The Pennsylvania State University**, University Park, PA
Bachelor of Science in Industrial Engineering

Employment:

Aug 1997-Dec 1997 **Virginia Polytechnic Institute and State University**, Blacksburg, VA

Graduate Teaching Assistant

- Provided teaching assistance (i.e. lecturing and grading) for an introductory manufacturing class
- Developed internet site for class

Oct 1996-Aug 1997 **Virginia Polytechnic Institute and State University**, Blacksburg, VA

Graduate Research Assistant

- Conducted research and interviews at NASA, DoD, and industry to identify innovative program and project practices
- Provided administrative support for project management
- Developed and authored a case study and training materials for use by NASA
- Developed several internet sites dealing with program and project management, macroergonomics, performance measurement systems, and reengineering
- Presented information on behalf of the Macroergonomics and Group Decision Systems Laboratory(MGDSL) and the Socio-Technical Alliance for Industrial Research to the Air Force, the Swiss Embassy, and various industry members
- Network Administrator (Windows NT) for MGDSL computer facilities

May 1994-Aug 1994 **Bangkok Bank Public Co. Ltd.**, New York, NY

Intern

- Monitored work flow and employee productivity
- Worked as a trader, maximizing usage of overnight bank funds
- Evaluated bank documents with knowledge of UCP 500
- Trained as an accountant to maintain records of transactions and to write monthly reports

Jan 1992-Apr 1992 **S & L Enterprises**, Sykesville, MD

CAD Technician

- Developed drawings for model trains and farm equipment using AutoCAD
- Optimized AutoCAD menus to increase worker productivity

Activities/Honors:

- Pratt Fellowship Recipient (Fall 1997)
- Institute of Industrial Engineers
- Society for Engineering & Management Systems
- Golden Key Member(National Honor Society)
- Alpha Pi Mu (Industrial Engineering Honor Society)