Development of a Methodology for Assessing Military Team Processes

Brent DeWayne Fraser

A 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, Chairman
Dr. Kari Babski-Reeves
Dr. Tonya Smith-Jackson

4 September 2003

Blacksburg, Virginia

Keywords: Team Performance, Team Observation, Team Evaluation Methods, Command and Control Team

Development of a Methodology for Assessing Military Team Processes

Brent DeWayne Fraser

Abstract

This study is based upon the premise that overall team performance is the sum of the team's performance in several individual team processes. The purpose of this study was to develop a tool to measure performance in each of these individual team processes. This study begins the measurement development cycle by developing a tool that uses direct observation to collect data on team processes. The tool was then tested in a battle simulation being used as a C^2 training exercise. The study showed that (1) the team process performance can be measured using direct observation, and (2) non subject-matter experts can accurately and reliably rate team performance by using the data collection tool.

The tool developed here relied on an observer to rate the command team on 43 items concerning ten team processes. Ratings were collected on visual analog scales that were overlaid with a Likert-type template for analysis. The results showed that six of the ten constructs were measured reliably and validly. However, inter-rater reliability was shown to be an issue to be overcome in future studies. The implications of this weakness, the potential of this line of research, and potential tool design changes are discussed in this thesis.

Acknowledgements

The research presented in this thesis represents a massive undertaking involving many different individuals and organizations. Without these individuals, this research would not have been possible. First, I would like to acknowledge my research committee: Dr. Brian Kleiner, Dr. Tonya Smith-Jackson, and Dr. Kari Babski-Reeves. Their efforts in this endeavor ensured that I remained on the path to finding truth in the research, and they were invaluable in the formation of this research endeavor. Next, I would like to acknowledge Dr. Sam Middlebrooks and the Army Research Laboratory for providing the means to pursue this research. Additionally, Dr. Middlebrooks' guidance has proven to be valuable time and time again. It has truly been a pleasure working with Dr. Middlebrooks and the Army Research Laboratory. Last, but certainly not least, I would like to acknowledge Mr. Charles Clark and the staff at the III Corps Battle Simulation Center at Fort Hood. The tireless efforts made by these individuals are what made this research possible.

To each of you, my sincerest thanks.

List of Figures

List of Equations

List of Tables

Table of Contents

Chapter 1 - Introduction

The purpose of this study was to develop and demonstrate a method for employing direct observation techniques in evaluating team processes. The theoretical background for this study is based upon the Team Effectiveness Model developed by Tannenbaum, Beard, and Salas (1992b), and it uses a brigade-level military Command and Control (C^2) team for testing and validation purposes. However, the methodology developed in this study is applicable for many different types of teams.

Due to the observational nature of this study, and the benefits of naturalistic observation techniques, naturalistic observation techniques will be applied to the extent possible. Naturalistic observation is a type of direct observation that studies objects in their natural environments. These objects can be animals, people, teams, or any other object that may be of research interest. Naturalistic observation is truly defined by more than just the setting of the observations (Blacker and Mortimore, 1996). Blacker and Mortimore's define naturalistic observation as observing participants that are the target population in their natural environment, performing their real-life tasks, while maintaining a separateness from the participants that sufficiently allows for non-interference in participants' tasks. While the observation in this study does not truly classify as naturalistic, the guidelines for naturalistic observation are followed.

The C^2 team observed in this study participated in a weeklong battle simulation at Fort Hood, Texas. These battle simulations are primarily used as training exercises for command teams. Though these simulations are not naturalistic, in the sense that lives are not truly at risk, every effort is taken to ensure as realistic a setting as possible. Because of this removal from the natural setting, the active battlefield, this study classifies as a field study, rather than a naturalistic observation.

Military C^2 teams will be used in this study for several reasons. First, these teams are experts at making decisions in a high-pressure environment where the consequences of those decisions can be high. Second, battle simulations offer an opportunity to observe these teams in

a simulated naturalistic environment. Finally, C^2 teams are extremely interdependent, and therefore exhibit the observable signs of the team processes more than less interdependent teams.

The C^2 team represents the extreme case of high-pressure teamwork. However, the team's internal processes are the same as the processes found in decision-making teams in and out of the military. While the risks and consequences differ among military and civilian teams, the theoretical team processes being observed in this study are the same in any team. Therefore, the results of this study are very applicable outside the military C^2 environment.

Since this study will involve command and control teams, it is necessary to explore military C^2 and the methods in which these teams have been evaluated in the past. Military C^2 teams are groups of commanders and their support staffs that control and coordinate the placement, actions, and support of combat teams in the field of battle. This study is a first step to develop meaningful, quantitative C^2 measures of effectiveness in the team processes. This development path is important because team research has yet to develop quantitative measures of the team processes.

Sproles (2001) describes the difficulty in defining measures for C^2 in terms of the difficulty in separating C^2 from the rest of the combat system. It was once considered impossible to evaluate command and control separate from battle outcomes. However, command and control teams rarely actively fight in combat, so this measurement policy is, at best, an indirect measure of command team performance. There has been a shift to evaluate C^2 separate from mission outcome, but, these measurement systems are primarily qualitative in nature and extremely susceptible to bias (Sweet, Metersky, and Sovereign, 1985). While there is no clear method available for quantifying C^2 performance, there is the realization of the importance for developing such measures.

Command and Control is a very important component of an effective military force. Without the coordination provided by C^2, the U.S. Military would not be the foremost military power in the world. It is in the interest of the U.S. to maintain the status of having the most powerful military. Military technology plays a part in the extension of this power, but with

ineffective command and control, this technology is also ineffective. Therefore, steps must be taken to insure that C^2 is continually evolving and becoming more effective. However, before any improvements can be made to the system, it is necessary to gain knowledge and insights into how the command and control teams operate and interact, and this means the processes must be measured.

The United States Army has struggled for decades with how to evaluate command and control teams. The Army admits that there is no single accepted way to evaluate command and control teams (Halpin, 1992). The primary roadblock at this time is the absence of a method to collect data in the Tactical Operations Center (TOC) that is necessary to understand and evaluate the C^2 team.

Many of the current and proposed methods of command team evaluation require large amounts of post-exercise analysis time and enormous human resources to conduct. Since the C^2 team is largely an information processing and decision making entity, several methods focus on the collection and analysis of conversation. This data collection is costly in time and manpower. Even with the advances in analysis software and voice recognition software, the military language could present a significant problem in shortening this analysis time. In the end, the majority of these evaluations produce qualitative results anyway. There is a need to develop an observational method for evaluating TOC operations that does not require months of analysis time, that decreases the human resources needed to as few as one, and offers quantitative information on the team processes.

The research challenge presented for study is expressed in the following quote from Kirwan and Ainsworth (1992, p. 58): “Observation techniques cannot provide information about underlying thought processes, and so they will be of little use for highly cognitive tasks”. This study is the first step to developing an observational tool that can be successfully applied to the highly cognitive work environment of C^2, in order that the command and control process may be better understood and eventually evaluated.

The primary research question posed in this study is what existing naturalistic observational techniques can be applied, or modified for application, in the command and control environment that can be employed by a single observer and require an acceptable amount of analysis time, while maintaining observation accuracy and data relevance. Specifically, this study will attempt to answer the following question:

1) How can direct observation methods be used to evaluate military command and control teams?
    a) What data types can be collected in the TOC using direct observation?
    b) What are the tradeoffs among the different observation methods?
    c) In what ways can the potential data be collected and analyzed?

Based on previous research, the following hypothesis has been formulated:

1) Direct observation techniques can be employed to explore the command and control team processes in a quantitative manner with non-expert observers
    a) The data collected will be reliable across observers
    b) The data results will be equivalent to the command team's self-assessment

Chapter 2 - Review of Literature

2.1 Group and Team Performance

As the term "Command and Control (C^2) team" implies, C^2 relies on the collective efforts of the individual commanders within the team. An understanding of team processes and functioning is necessary to explore the processes and functions of the C^2 team, just as it is for any team.

2.1.1 Team Definition

The definition of a "team" may be as varied as some of the individuals you may find in a typical team. A slightly different definition can be found in almost every literature entry concerning team models. The preferable definition of a team may be some composite of many of these individual definitions. A team requires a group of individuals in order to exist, but a group of individuals does not necessarily constitute a team. For a team to exist, this group of individuals must possess a relationship that fosters an interdependent effort to achieve a common goal. A "team" is set apart from a "group" by this relationship (Swezey and Llaneras, 1997).

Having knowledge of the definition of a team is essential in this research endeavor. Since the defining characteristic of a team is the relationship among the team members, TOC (or team) performance cannot be evaluated by observing the individuals. Observations must, therefore, focus on the team members' interactions, rather than the actions of a single player.

The following diagram is the Team Effectiveness Model as proposed by Tannenbaum, Beard, and Salas (1992b). This model was designed to explore how teams function and to evaluate team effectiveness (Tannenbaum et al., 1992b). In addition to providing the components that contribute to team effectiveness, example variables have also been included for each component.

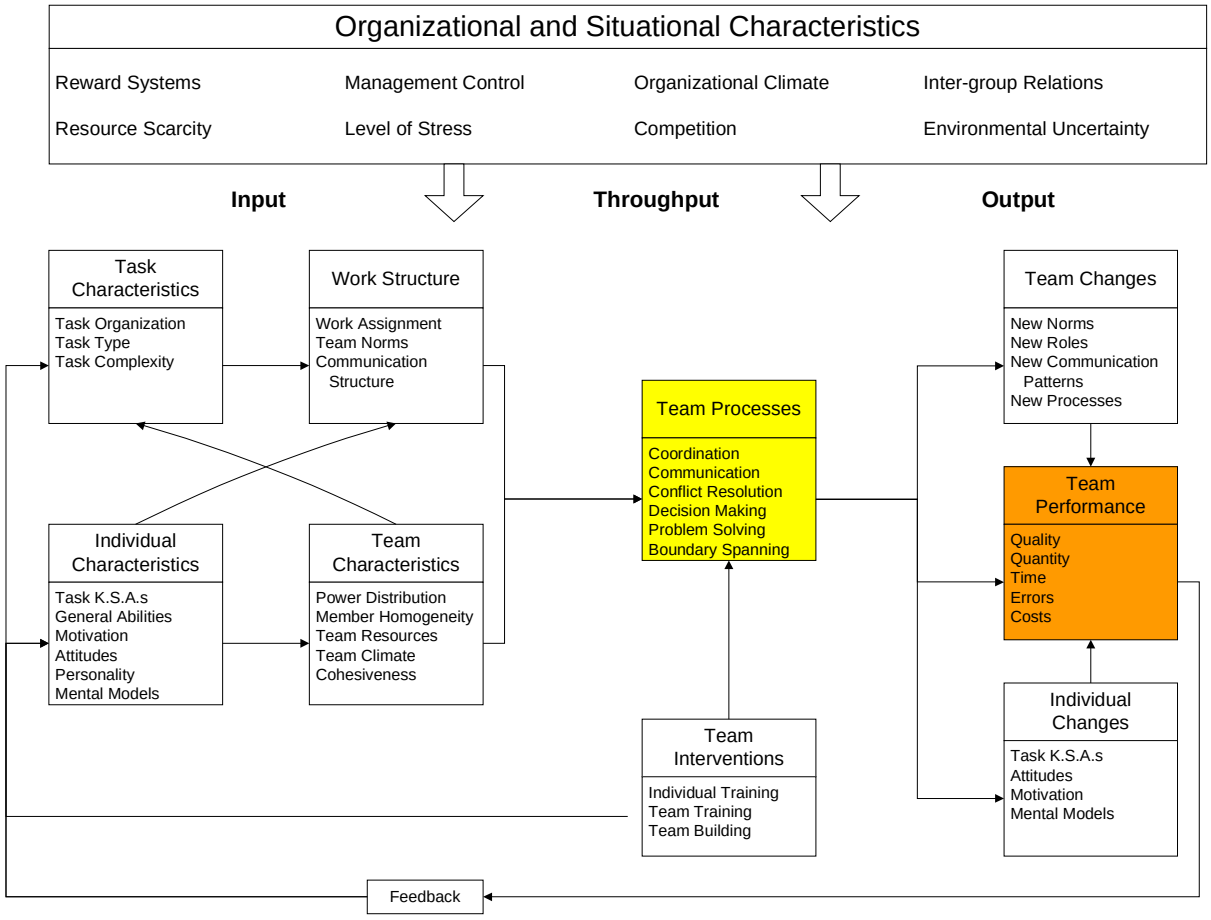


Figure 1 - Team Effectiveness Model (Tannenbaum, Beard, and Salas, 1992a)

The thrust of this study is the development of a framework for observing and measuring the team processes illustrated in this model, in the hopes that future research can use team process observations to predict team performance.

2.1.2 Stages of Team Development

Teams can be categorized by their team processes. Scholtes, Joiner, and Streibel (2002) describe four stages of team development: Forming, Storming, Norming, and Performing. Each of these stages has unique characteristics that are displayed in the team processes. This section illustrates the four phases and the characteristics described by Scholtes, et al. (2002)

The Forming stage of team development is the first stage of team development. It is during this time that the problem is defined and team discussion is focused on abstract development of the action plan for solving the problem. Team members have a tendency to venture off-topic quite often during this stage. Because of this, team workload might be high, but not well focused.

The next stage of team development is the Storming phase. The Storming phase is named such because this is an argumentative stage in team development. Arguments arise because individual differences and expertise are revealed to the group. Team members begin to become defensive of ideas and methods. An informal authority structure also appears in this phase, which can cause further tension in the team. Goals formed during this phase are often unrealistic.

Assuming the team survives the Storming phase, it enters the third phase: Norming. In this phase the team starts to show its first signs of cohesion. Discussion moves from argumentation to discussion of the team's problems and forming solutions. Criticism is conveyed constructively as levels of defensiveness are decreased. Team rules are formed, either implicitly or explicitly. With less energy spent on defending the individuals' egos, more energy can be spent solving the problems the team was formed to solve.

Teams that reach the final stage, Performing, are the more effective teams. Every team should have this stage as its goal when the team is formed. Team members begin to self-identify with the team, and they understand the other team members' strengths and weaknesses. This somewhat intimate relationship with the other team members allows conflict to be quickly resolved or avoided in entirety. This further reduction in conflict allows even more energy to be dedicated to solving the problems at hand. The knowledge of the team members' expertise can also make certain assignments automatic, thus reducing time in the delegation process. The Performing stage represents the apex of a team's problem solving efficiency.

The application to command and control teams is that if it can be ascertained which stage of development a certain team is in, the performance level of that team can be inferred in relation to the performance potential of that team.

2.1.3 Team Process Attributes

The measures of team process fall into several groups, called "team process attributes" in this study. The team attributes contained in this section are those that must be measured to effectively evaluate the functioning of the team processes. These attributes are illustrated in Swezey and Llaneras (1997). While these attributes are explored separately, many of them are interrelated and cannot be measured separately.

These attributes are separated into "team processes" and "cognitive processes" as shown in the following table.

Table 1 - Team Attributes

Team Processes	Cognitive Processes
Coordination	Joint Knowledge
Communication	Joint Task Ability
Conflict Resolution	Motivation
Decision Making	Personalities
Problem Solving	Shared Thought Processes
Boundary Spanning	

The team processes are the actual steps taken by the team during teamwork. The cognitive processes are input factors into how the team processes are conducted.

2.1.3.1 Coordination

Coordination refers to the team's ability to work collectively in an efficient manner. It has been suggested that high-stress teams, which military command and control teams would qualify as, become high performance teams by adapting their decision-making and coordination strategies over time, as the situation dictates, in order to maintain a manageable stress level and similar performance levels. The change in coordination strategy is from an explicit strategy in

low stress environments to a more implicit strategy in high stress environments. In order for implicit strategies to work, the team members must develop a congruent shared mental model (Serfaty, Volpe, and Entin, 1993). Thus, effective teams actively manage the environmental stress' effects on the team.

In order to make measurements of coordination, it is necessary to understand how coordination manifests itself in the command team. Xiao, Mackenzie, and Patey (1998) list different forms of coordination. This list follows.

- Following Protocol – Team members follow a pre-determined set of procedures that define individual roles
- Following the Leader – Doing what is told them by the leader or focusing on the leader's actions
- Anticipation – Team member's knowledge of the events and the probable reaction of the leader guide members to perform certain tasks
- Activity Monitoring – Activities guided by knowledge of task sequences. This is another form of Anticipation
- Information Flow – Team members spontaneously give opinions or information that was not specifically asked for. This is yet another form of Anticipation

2.1.3.2 Communication

Bowers and Jentsch (2002) suggest that relative skill in communicating is the trait that differentiates effective and non-effective teams. Communication refers to the information sharing functions of the team. This information sharing can be inter- or intra-team. Communication is a necessary element for effective teams, as it forms the basis for other team attributes, such as conflict resolution and coordination.

One study demonstrated the effect of communication on team performance while measuring the effects of speech intelligibility (Whitaker, Fox, and Peters, 1993). This study demonstrated that as intelligibility, which is a performance measure of speech communication,

was reduced, team performance was reduced as well. This demonstrates the value of clear communication within the team. Some other points of good communication are staying on topic, refraining from hostile comments, refraining from excessive humor, and whole-group participation (Baird and Weinberg, 1981).

2.1.3.3 Conflict Resolution

Conflict, obviously, is contrary to effective teamwork. Conflict can negatively affect many other team process attributes as well, such as communication and coordination (Morgan and Bowers, 1995). Conflict can be seen in verbal and non-verbal ways: members "tuning out" other members, members holding back information or feelings and opinions, hostile conversation, expressions of incompetence, pulling team conversation off-topic, excessive criticism, monopolizing the conversation, and so on (Baird and Weinberg, 1981). Baird and Weinberg (1981) assert that conflict resolution can occur in four ways:

1. Tolerance – Ignoring the transgressor until the episode passes and the individual returns to a normal, non-combative state. This method has a low probability of success in most cases.
2. Avoidance – Avoiding the topic that sparked the conflict. This method is only effective as long as the topic is not discussed.
3. Restriction – This method represents a range of actions from alienating the transgressor to implementing a formal communication procedure, such as Parliamentary Procedure, to devising a punishment for such aggressions.
4. Treatment – Actions taken to resolve the conflict. This method encompasses team building, objective discussions concerning the conflict, involvement of the aggressor in the conflict resolution, and role-playing.

Obviously, the treatment method of conflict resolution is the preferred method, as it produces lasting results. However, it is the most difficult and time-consuming method to implement.

One gray area of concern when observing a group is the fine line between verbalizations of emotions, including frustrations, and true conflict. Theoretically, verbalizations of emotions can sometimes aid in conflict resolution within a team by identifying the probable sources of conflict prior to the point where the problem has escalated and introduces hostility to the team (Baird and Weinberg, 1981; Powell, 1969).

2.1.3.4 Decision Making

A good description of the military command and control team and its function in decision making is given by the following quote:

> Tactical teams within the military exist (1) to help a leader assess a given scenario involving imminent danger or threat, (2) to provide information to the leader in a form that he or she can use in making a decision, and (3) to implement the action implied by the decision that the leader comes to. In this sense, the tactical team is a decision-making team." (McIntyre and Salas, 1995, p. 9).

Decision-making is the process in which commitments to a certain plan of action are made. This process is composed of four distinct phases (Guzzo, 1986):

1. Intelligence – Gathering the necessary information to make the decision and recognizing the need for a decision
2. Design – Developing alternative courses of action
3. Choice – Choosing a particular alternative as the overall best
4. Review – A reflective phase where the decision maker studies the results of the action and determines the correctness of the decision

2.1.3.5 Problem Solving

Problem solving can be considered part of the decision making process. It is the application of team knowledge and skills to an existing problem in order to develop a solution. Problem solving and decision-making are part of the same process as each involves the other.

Francis and Young (1979, p. 217) developed a questionnaire to elicit a team's problem solving capacity from the team's members. The points that are outwardly observable follow.

- Clear objectives shared by everyone in the team
- Clear success criteria were established
- Action taken was effective
- Thorough review of outcomes was performed with team members
- Time was well used
- All involved team members participated positively

As this questionnaire represented a means to identify areas for improvement, there is no performance-based analysis involved with these categories. However, it gives categories that can be used to collect problem solving performance data.

2.1.3.6 Boundary Spanning

Boundary spanning refers to the manner in which a team interacts with those outside the team (Tannenbaum et al., 1992b). This type of interaction occurs regularly as a normal function of the command and control team. The team must communicate with the units under its command, to receive intelligence and transmit orders, as well as communicate with higher command. This type of communication is generally conducted via a remote communication device, such as two-way radio or tactical email.

In one study that was conducted, boundary spanning, or boundary management, was determined by the degree of misunderstanding between teams (Gladstein, 1984). In the battlefield, this type of misunderstanding can occur in two ways: between higher command and

the TOC and between the TOC and the lower commands. In the typical, simulated battle exercise, the higher commands are not simulated, so the only boundaries that can be observed are between the TOC and the lower commands. Evidence of breakdowns in boundary spanning would be shown by the actual movement of forces on the battlefield being contrary to the orders intended in the TOC.

2.1.3.7 Joint Knowledge

Joint knowledge is knowledge that is conceptualized as a shared mental model within the team. This knowledge is required for effective team functioning and consists of shared, consistent knowledge of the following items (Cannon-Bowers, Tannenbaum, Salas, and Volpe, 1995):

- Teamwork skills
- Team goals, objectives, missions, and strategies
- Team interaction protocols with outside entities
- Team members' roles and responsibilities
- Team members' characteristics (abilities, preferences, tendencies, strengths, and weaknesses)
- Team coordination cues

2.1.3.8 Joint Task Ability

Joint task ability, or collective skill, is a phenomenon that has been identified in high performing teams (Shaiken, 1996). Shaiken (1996) defines skill as the ability to diagnose and correct problems. Skill is composed of three functions: theoretical knowledge, experience, and technique. Theoretical knowledge and experience are the necessary components of skill, and technique is the observable component of skill.

Shared task ability has been shown to be less than the sum of individual abilities in teams (Tziner and Eden, 1985). Tziner and Eden (1985) demonstrated that team performance was disproportionately skewed toward the level of performance of the least able team member.

2.1.3.9 Motivation

Motivation is the force that energizes activity toward a particular goal. It is exhibited as enthusiasm for one's work. Few would argue that motivation is essential to performance. Most of the motivation literature is specific to individual motivation, and there is an underlying assumption that team motivation is the sum of individual motivation. There are a few suggestions in the literature to ensure motivation. These suggestions can give clues to the aspects of teamwork that are necessary to observe team motivation.

McManus (2001) asserts that good, relevant feedback is necessary to promote employee motivation. Additionally, according to Pollock (2001), the following items increase or maintain motivation in employees:

- Ensure balanced workloads
- Allow employees enough time to complete the work
- Supply the correct equipment for the task
- Distributing well-defined assignments

2.1.3.10 Personalities

The effect of personalities on team performance has been of research interest for many years. Unfortunately, there has been little success in determining the true effects of personality on performance, partially because there is great uncertainty involved in defining personality traits, and that uncertainty has manifested itself in the research (Barrick, Stewart, Neubert, and Mount, 1998; Neuman, Wagner, and Christiansen, 1999).

The following personalities have been identified as being counter-productive to the team process (Arnold, 1999):

- The Clown – Displays excessive humor
- The Cynic – Disagrees with every team decision
- The Dominator – Dominates all interaction forums
- The Interrupter – Disrupts or does not participate in the speaker's discussion
- The Loner – Does not communicate with the other team members
- The Movie Star – Attempts to focus all attention on him/herself
- The Prisoner – The member forced to be on the team
- The Last Word – Must have the last word in every conversation
- The Latecomer – Is habitually late to all team functions

Further research has developed five personality dimensions that relate to team performance. The line of research is summarized by McCrae (1989). These dimensions are:

- Extraversion
    - Positive: Assertive, social, talkative
    - Negative: Reserved, cautious
- Agreeableness
    - Positive: Cooperative, tolerant
    - Negative: Irritable, suspicious, inflexible, uncooperative
- Conscientiousness
    - Positive: Careful, responsible, organized
    - Negative: Irresponsible, disorganized, lack of self-discipline
- Emotional stability
    - Positive: Calm, secure, enthusiastic
    - Negative: Anxious, depressed, emotional, insecure
- Openness to experience
    - Positive: Open-minded, curious, imaginative
    - Negative: simple, narrow-minded

A comprehensive list of observable personality traits is available in Hofstee, de Raad, and Goldberg (1992).

Conscientiousness, which appears to be an additive property in teamwork, tends to be positively correlated with performance, but only if the variance in conscientiousness is minimized, as this variance is negatively correlated with performance. The other personality factors exhibited traits much as joint task ability. It was found that adding a single team member that exhibited traits on the extreme negative end of the conscientiousness continuum brought team performance down dramatically (Barrick et al., 1998).

2.1.3.11 Shared Thought Processes

Cannon-Bowers, Salas and Converse hypothesize in their exposition on team shared mental models that there are a number of items that must constitute the shared mental model. They list the following knowledge items that each team member must have in common (Cannon-Bowers, Salas, and Converse, 1993):

- Familiarity with the functions of the other team members
- Expectations of the way in which events will proceed
- Team strategies
- Expectations of team member performance

On the other hand, there are forms of shared thought processes that are not conducive to effective team functioning (Cannon-Bowers et al., 1993). One form of this phenomenon is manifest in Groupthink. Groupthink is a phenomenon where the team members agree on a particular decision because the team does not have the time, resources, or desire to pursue the matter further. It is assumed that groupthink does not allow the optimal solution to surface, and thus, a sub optimal solution ensues from this process.

2.1.4 TOC Evaluation

Since the Tactical Operations Center (TOC), in which military command and control teams operate, will be the setting for this study, an exploration of the current methods of TOC evaluation is necessary. Command and Control (C^2) teams are typically removed from battle areas and plan and orchestrate the art of war. It is the duty of C^2 to process incoming intelligence information and mission directives to direct the units in its command to effectively meet the mission objectives. As is apparent by this description, C^2 is a highly cognitive process, and thus very difficult to measure.

The United States military has attempted for many years to develop an empirical method for evaluating TOC performance. This evaluation is necessary to determine and eliminate operational problems within the TOC. With continual improvement, the United States Army can maintain its status as the world's foremost army. Unfortunately, a set methodology has yet to be set for evaluating a TOC in a way that improvements can be effected.

One study that tackled the problem of TOC evaluation was performed by Olmstead, Christensen, and Lackey (1973). This study opens with the premise that TOC evaluation should be based on two factors: Organizational Effectiveness and Organizational Competence. Organizational Effectiveness is defined simply as the level at which the organization achieves its goals. Military organizations, for the most part, have very clearly defined goals, called "tactical objectives". Many of these objectives would be easily evaluated for achievement (Olmstead et al., 1973).

The second criterion, Organizational Competence, defined fuzzily as the organization's ability to cope with an every-changing environment, is considerably more difficult to objectively evaluate (Olmstead et al., 1973). Competence contrasts effectiveness in that competence embodies the process of achieving the goals that determine effectiveness. Competence would include such fuzzy variables as leadership strength, how well the group communicates, and how well the individuals in the team work with each other. One would be hard-pressed to find objective measures for these attributes. Olmstead based much of the competence measure on

communication. The method he used required processing, line-by-line, the transcript of all communications throughout the exercise. This process is extremely time-consuming and makes this method of measuring competence infeasible.

Layton and Feld (1996) developed one potential measurement system. Command performance was assessed using the following criteria:

1. Quality of Understanding – Staff perceptions of the situation (rated as correct, not incorrect, and incorrect)
2. Options – The number of potential future scenarios considered in the planning stage
3. Planners – the number of staff members participating in planning
4. Queries Required – requests for additional information needed for planning
5. Plan Time Less Than Understanding Time – (Yes/No)
6. Option Rejection, Commander – Was the recommended course of action rejected by the commander (Yes/No)
7. Option Rejection, Other – Was the recommended course of action rejected by someone other than the commander (Yes/No)
8. Lead Time Adequacy – Did the commanders give subordinates enough time to carry out planning

These metrics represent the heart of Command and Control, but many of them are not easily collected. Further research must be completed to determine ways to collect these data; else alternate, more easily collectable data needs to be defined.

2.2 Issues in Team Research

Aside from the complexities of evaluating military command and control teams, there are several issues that pertain to team research in general that must be explored. These issues are: distributed cognition, naturalistic decision making, action research, and situated action models.

2.2.1 Distributed Cognition

Distributed cognition is defined as:

> …a new branch of cognitive science devoted to the study of: the representation of knowledge both inside the heads of individuals and in the world …; the propagation of knowledge between different individuals and artifacts …; and the transformations which external structures undergo when operated on by individuals and artifacts … (Flor and Hutchins, 1993, p. 272)

Nardi (1996) asserts that the combination of human-instrument cognitive systems can only be understood when the relationships among the system components, and how those relationships contribute to the system as a whole, are understood. The study of distributed cognition moves away from the study of individual cognitive patterns and attempts to view system-level cognition. Hutchins illustrates that a system's cognition is composed of the interactive cognition among the team members, as well as the interactions with non-human system components. The illustration demonstrates how non-human components exhibit cognitive traits, such as memory (Hutchins, 1995).

2.2.2 Naturalistic Decision Making

Naturalistic decision making, in the team context, refers to teams making decisions in real time as situations arise in the natural environment. Typically, as is the case in C^2 teams, these decisions must be made under considerable time constraints, which can also equate to insufficient information on which to base the decision. Decisions made in naturalistic settings are also prone to personal biases. Since not all information can be attained and processed in the given amount of time, some decision makers will focus on information that supports their beliefs (McDaniel, 1993). One example that is often cited in the literature is the USS Vincennes incident, where an Iranian passenger airplane was mistaken as a hostile threat and downed. Possible contributors to these decisions include a high expectation among the crew that an attack

will be made by the Iranians. A good, brief account of this incident is available in McDaniel (1993).

As is evident in the account of the USS Vincennes incident, the time pressure and personal biases in the team can lead to disastrous mistakes (McDaniel, 1993). Such decisions often rely on shared mental models (Cannon-Bowers et al., 1993). The quality of the decisions made will depend on the congruence of the shared mental models across the team members.

2.2.3 Action Research

Action research uses naturalistic observation methods, so it deserves recognition here as well. The topic is similar to the "Plan, Do, Study, Act" methodology for change in organizations. Action research theorizes that the way to effect positive change is to study the process by direct observation or participation, define what needs to be changed based on that observation, implement that change, and revisit the organization to ensure the desired changes attained the desired effects. It is deeply rooted in the theory that the only way to learn about nature is to immerse the research in nature, instead of viewing it from afar (McNiff and Whitehead, 2000). The Action Research perspective is taken in this study. In order to improve C^2 function, C^2 must be measured, and the only way to make accurate measurements is to be present in the C^2 environment.

Conventional laboratory empirical research is being scorned in many research venues because those being studied rarely see any benefits coming from the study once it is completed (Uzzell, 1995). Military venues are not immune to this fact. Military research and development has seen a decline in funding over the past decade because the military no longer sees the connection between R&D and the war-fighting end result (Best, Gentner, Cunningham, Tiller, and Schopper, 2001; Gentner, Best, and Cunningham, 1999). The action research approach, if adopted correctly, ensures that the research participants reap some benefit from the study. Uzzell (1995) also implies that there is too great a distance between the researcher and the researched, to the point that the researcher often does not allow himself/herself to learn from the participants involved. Action research is also designed to alleviate this concern.

One example is found in the realm of human-computer system design. Greenbaum and Kyng (1991) assert that the traditional human factors approach to design was insufficient to solve the majority of the problems encountered in computer system design. The emphasis should be on the unique situation of each problem, rather than a set of static rules governing design. In order to develop a conception of the situation, it is necessary for the designer or researcher to be exposed to that situation. Only then can the system be designed to support the unique environment encountered by the end users.

2.2.4 Situated Action Models

Situated Action theory asserts that activity is a product of the interaction between the individual and the environment (Nardi, 1996). It forces the researcher to remove intensive focus from the individual and dedicate resources to tracking the ever-evolving nature of the environment in which the individual is operating. As it applies to the current study, situated action modeling warns that team processes cannot be researched by focusing on the individual team members and ignoring the environment. Indeed, as suggested by this theory, the more dynamic the environment, the more influence it exerts on the team processes, and combat environments are some of the most dynamic environments teams can face. Therefore, it is important to take the environment into account when evaluating naturalistic team performance.

Another study being performed simultaneously with this study, by another researcher, is focusing on collecting environmental data. This study is concerned with collecting data related to the physical environment and the communications transmitted between the main TOC and other areas. It is hoped the two methods can be fused in the future to satisfy the concern presented by situated action models.

2.3 Ethnography/Naturalistic Observation

Observation has long been a foundation of the natural sciences. Throughout history, anyone interested in learning about the world and the people living in it began study by

observing. It could even be stated that all of human knowledge stemmed from observation (Adler and Adler, 1994). Observation is so fundamental that most people fail to recognize it when it occurs. In fact, every person performs a number of observations everyday. People observe traffic patterns on the daily commute to determine the best time to depart for work in the morning. For recreation, many people visit the local mall and "people-watch" (Moore, 2002).

Naturalistic observation is simply that: observing things in their natural environments. However, it is defined by a little more than just the setting of the observations (Blacker and Mortimore, 1996). Adapting Blacker and Mortimore's definition of a naturalistic observation to the needs of this study, naturalistic observation is defined as observing participants that are the target population in their natural environment, performing their real-life tasks, while maintaining a separateness from the participants that sufficiently allows for non-interference in participants' tasks.

Naturalistic observation appears in simple forms that are familiar to most people. Naturalistic observation can involve the direct visual observation of the participants, or remote observation, or even photography (Kirwan and Ainsworth, 1999). The technique is extremely broad and versatile. The power, importance, and prevalence of naturalistic observation should not be underestimated.

References to differing naturalistic observation techniques abound in sociological and psychological literature. Almost all of these references use observation techniques as part of a qualitative research strategy. Many scientific circles dismiss the validity of qualitative research, because in qualitative research, the observer is the data collector, analyzer, and interpreter (Borman, LeCompte, and Goetz, 1986). Qualitative naturalistic observations have a definite place in research. Even those who prefer quantitative research can see the benefits of the qualitative side of naturalistic observation in determining what to look for and where to look for it (Moore, 2002). In fact, Kirwan and Ainsworth (1992) echo the feelings of many other researchers when they assert that it is absolutely necessary to know what you are looking for before beginning the observation process. However, beyond the beginnings of the research, this qualitative nature introduces many problems in the analysis of the results of the observation.

How much of the research has been influenced by the observer's thoughts, feelings, beliefs, values, expectations of the results, et cetera? This problem may be a reason why so many scientists and engineers reject qualitative research and any methods associated with qualitative research (Lipinski and Nelson, 1974).

Many scientists have discarded naturalistic observation techniques because of their beginnings in qualitative research. Laboratory experimentation seems to be the preferred method facilitating research with many scientists and engineers. This fact is not without good reason as it is the method that allows for the easiest data collection and analysis. The experimenter has control of nearly every aspect of the experiment. Unfortunately, this control affects any results received by eliminating potential variables, and thus the experiment has a low external validity. Additionally, if the research is focusing on a real-world problem, the data collected from the laboratory experiment represents a sterile world very unlike the real world (Hendrick and Kleiner, 2001).

By contrast, naturalistic observation is, in ideal form, as close as possible to real-time, real-life (Hendrick and Kleiner, 2001). Human-computer interaction researchers have realized how artificial laboratory research was and have endeavored to incorporate naturalistic principles, where possible (Blomberg, 1995). Naturalistic observation should be considered an essential component of any Human Factors research, particularly Macroergonomic research, where group dynamics can introduce variable interactions that would be lost in the laboratory. If real-world problems need to be solved, the real world is the best place to collect the data needed to solve the problem.

Laboratory experiments have their place in research. Naturalistic observations typically have high external validity, but the results of the observation may not be very generalizable and typically cannot be replicated easily (Graziano and Raulin, 1989). This issue exists because the observation takes place in a very specific environment, over which the observer has no control. Laboratory experimental results are typically more generalizable because of the environmental controls available in the laboratory setting.

Because of the trade-offs between laboratory experiments and naturalistic observation, a field experiment will be conducted in this study. A field experiment is a fusion of naturalistic and laboratory environments. The experimenter has more control over the environment than possible in a purely naturalistic environment. However, the natural environment is maintained to the extent possible. The usage of the field experiment techniques should serve to make the results of this study more generalizable and easier to replicate in both the naturalistic and laboratory environments.

Using the field experiment method compromises the non-interference principles necessary in naturalistic observation. Observing in close proximity increases the probability of observer effects. The proximity also increases physical interference, such as the observer physically standing in the way of a participant while the participant is attempting to complete the assigned task (House, 1978). Every square inch, every piece of furniture or equipment in a TOC has a purpose. An observer in the TOC must be exceptionally aware of what is occurring in the TOC and anticipate the movements of the players in order to avoid hampering the function that is being observed.

Non-interference is very important in observation because of the potential for the observer to alter the actions of the observed. Typically, a person will modify his behavior when being observed (House, 1978). This behavior modification can be further exaggerated if the observation is taking place in the laboratory (Christensen and Hazzard, 1983). Shaughnessy and Zechmeister (1985) go so far as to suggest that participants being observed will attempt to guess how the observer wants them to act and act accordingly. Since research endeavors are typically interested in the participant's natural response to the situation being presented, and not the participant's response to the observer's desires, observer influence must be carefully controlled.

The literature gives many guidelines for decreasing the interference effect of an observer, which have potential use in this project. House (1978) notes that participant reactivity to the observer tends to diminish as the observer becomes a familiar staple in the participant's environment. This would suggest that an observer in the TOC would need to be phased in. In essence, there is reason for concern that data collected while the observer is still new to the

players in the TOC would not be truly representative of that TOC. As time progresses with the observer in place in the TOC, the observer should become somewhat transparent (or organic) to the TOC. Data collected after this point should be representative of that TOC's normal operations. Of course, this is made very difficult with short-term studies, as there is not much time to adjust the participants to the presence of the observer (Kirwan and Ainsworth, 1999).

A better solution for ensuring non-interference is training the observer in non-interference procedures. These procedures would instruct the observer in methods of collecting the necessary data while maintaining a certain distance, both physically and socially, from the activity. Due to the confined space of the TOC, the emphasis will be training the observers to be mindful of personnel movement, so that the observer is not physically in the path of activity, and refraining from socializing with the individuals in the team. Overall, the observer must make efforts to minimize his/her presence in the experiment.

With longer-term studies, this phasing in can be taken to the extreme with Participant Observation. In Participant Observation, the observer joins the group and participates as an active member during the observation session. This method requires that the observer's actual agenda be kept from the group being observed. One classic example of participant observation is a study that was performed in the 1950's, where a group of researchers tested their theories on the effects of failed prophesy on people (Festinger, Riecken, and Schachter, 1956). Festinger, et al. infiltrated a religious cult that was prophesizing a flood that was to devastate humanity. The researchers entered the cult as active participants in order to study the people before and after the prophecy had failed. The cult's followers were never privy to the fact they were being studied. Certainly, there are ethical issues concerning this deception, regardless of the research benefits obtained by this method. However, if this method is to be employed, the observer using this method should be aware of the laws governing the groups being infiltrated.

The observer's effect on the participant is not the only source for interference. Observers are biased toward predicted outcomes, and those biases can make their way into analyses very easily. Prime areas for these expectancy effects are in the coding of the observation session and the analysis of the code sequences (Shaughnessy and Zechmeister, 1985). Shaughnessy and

Zechmeister give two suggestions for minimizing observer bias: automatically coding/analyzing observation sessions and using an observer that is not affiliated with the research project and limiting the amount of information given to the observer. In this experiment, participants not affiliated with the project will be used as observers. The research hypotheses will not be disclosed to the participants in this study. The only interaction between the research team and the participants will be explanation and clarification of the team assessment tool items.

Automatically coding/analyzing observation sessions could prove extremely difficult, even with modern technology. While it may be possible to utilize this method in the future, it is currently infeasible for dynamic human environments where cognitive processes are the events of interest, because there are currently no methods available for clearly mapping the cognitive processes to easily observable artifacts. The second option, however, is viable, though it also introduces additional problems. It might be possible to use a naïve observer in the TOC. While the security clearance of the individual might pose a problem, the most difficult challenge would be the person's inexperience with TOC observations. This challenge could possible be overcome if the observation tool given to the observer was very specific in the actions/interactions that are to be observed.

Naturalistic observation methods have been shown to be useful in recording military team organization and behaviors (Roth, Hritz, and McGill, 1984). The following sections explore several naturalistic observation methods that could be employed in team research.

2.4 Selected Observation Methods

2.4.1 Activity Sampling

Activity Sampling involves making generalizations of the physical actions of those being observed. These generalizations are typically in the realm of core Industrial Engineering, and focus, as the title implies, on describing the activities being performed. Cognitive activities, which have no truly discernable physical manifestation, are not well evaluated using this method of observation (Kirwan and Ainsworth, 1999).

Kirwan and Ainsworth give the following guidelines in using this method of observation:

1. Observe only distinct, easy-to-describe activities
2. Observe no more than 20 distinct activities in each observation session, though this limit can be extended if electronic recording devices are utilized
3. Determine what activities to observe before session
4. To obtain accurate samples, utilize the Nyquist Criterion (the Nyquist Criterion states that in order to receive an accurate representation from the sampling, samples must be taken at twice the rate of occurrence for the activity being sampled)
5. Focus mainly on occurrence frequency data

This method of observation is most useful to determine the amount of time that is spent on a particular task. It also has the strength that a single participant does not require the constant attention of the observer, though the amount of divided attention allowed is constrained by the Nyquist Criterion.

This method is typically used with a timing device so that generalizations can be made concerning the percentage of time spent in each activity. If an equal-interval or random sampling method is used systematically across all participants in the observation session, these percentages can be compared, and relative workload can be inferred.

There are many examples of the use of activity sampling or work sampling in the literature to choose from. One simple example was used to evaluate the distribution of workload in nursing staff across three shifts (Upenieks, 1998). Utilizing a random sampling procedure, Upenieks (1998) determined there existed a difference in the composition of work among the three shifts.

Activity sampling has the benefit of allowing the observer to collect quantitative data on multiple activities simultaneously. However, this method has the disadvantage it can only collect data on discrete, easy to identify tasks, which do not include cognitive tasks.

2.4.2 Conversation Analysis

As a subset of content analysis, conversation analysis is, as the term suggests, an analysis of conversation. To qualify as true conversation analysis, however, the conversation must occur in natural settings. While this method can be employed in an observational nature, it is typically employed using permanent records of the conversation (transcripts, recordings, etc.).

Conversation analysis is used when the researcher is interested in analyzing the communication patterns of individuals or groups. This analysis searches the conversation for indicator words and phrases to which the researcher has tied particular meanings. The method is descriptive and qualitative, and is used to gain insight into the ways in which humans interact and share information. While the method centers around the verbal aspect of the interaction, there have been additions made to the method to include items such as laughter, voice inflections, and non-verbal gestures while conversing (Heritage, 1989). This method of observation has been used to analyze dominance in communication (Roger, 1989), evaluate social skills in the handicapped (Thomas, 1989), and the effects of non-verbal cues on conversation (Rutter, 1989), among other things. In fact, this method is applicable in any research where the content of communication is of interest. This method gathers information on communication, which may be the heart of teamwork. However, the process of collecting the conversation, transcribing the conversation, coding the transcript, and analyzing the codes is extremely resource intensive.

2.4.3 Interaction Analysis

Interaction analysis focuses on the communication among the people of an organizational system. This section will concentrate on the Flanders Interaction Analysis System and its variations.

The Flanders Interaction Analysis System (FIAS) was initially developed for the classroom setting to try to determine what behaviors and attitudes are beneficial or detrimental to pupil learning (Flanders, 1965). It was used to record the communication interactions between

the teacher and the student body. For example, there were codes for lecturing, questioning, responses, et cetera. The end result of an observation session is a matrix of interactions. The analysis of this matrix depends on the individual situation, the interactions recorded, and the importance given to the interactions by the researcher (Darst, Mancini, and Zakrajsek, 1983). As such, this method is very versatile.

The FIAS removes a large amount of observer influence by dictating that the observer only records what mode of interaction that is occurring in a given time frame. The subjectivity is moved from the observation session to the analysis stage of research. Since the data collected represents only time-discrete interaction units, the analysis must be based on pre-defined definitions of the interactions and their meanings in the research.

The FIAS has been modified many times to suit individual situations. The Cheffers' Adaptation of FIAS (CAFIAS) added a qualifier to each discrete observation time to record whether the interaction was verbal or non-verbal (Cheffers, 1972). The second generation of the CAFIAS, CAFIAS II, added categories to record emotions expressed. A specialized adaptation of CAFIAS, called the Dyadic Adaptation of the CAFIAS, shifted the focus from individual-group interactions to purely dyadic, or individual-individual, interactions (Darst et al., 1983). However, since TOC operations may and do consist of more than dyadic interactions, the CAFIAS will be investigated.

The CAFIAS system development added more to FIAS than just a qualifier allowing non-verbal interactions; it developed new methodologies and techniques to employ FIAS. The first of which was the possibility for using video recorded observation sessions in order to minimize data loss (Cheffers, 1972). The only drawbacks mentioned in the study, such as cost and equipment failures, have been overcome by current technology.

Second, the construction of CAFIAS allowed for multiple observers during the setting with little data conflict (Cheffers, 1972). Cheffers trained his observers in the proposed system for 15 hours, a relatively small amount of time for group observation at this level. This amount of training would seem to be inadequate to produce the homogeneity demonstrated in the

observers' interpretations of the interactions. This capability demonstrates the relative objectivity the method can bring to team observation.

By itself, interaction analysis does not relate performance information. This method only identifies the interactions taking place. As such, the researcher has the freedom to stipulate what is preferred and what is not. The primary resource needed to collect this type of data is a human observer. The number of observers needed to perform this method depends on the number of different interactions that are collected and the length of the observation session. In most cases in the literature, one observer was sufficient.

2.4.4 Link Analysis

Link Analysis is used to illustrate relationships between subsystems. In the context of the TOC, these relationships could be between teams, individuals, or an individual and equipment within the TOC. This linkage information can be weighted in terms of expected importance and link frequency to evaluate the effectiveness and importance of the interface mechanisms (Wilson and Corlett, 1995).

A major strength of this method is in what is considered a link. Links are highly adaptable to the situation and can range from the traditional human motion within a system to communication between humans to the attention focus of a single human in the system (Kirwan and Ainsworth, 1999). Link analysis of inter-human communication borders on Interaction Analysis, so this method is further discussed in Interaction Analysis.

Kirwan and Ainsworth (1992) assert that the primary advantages of this method of observation are that the measures are completely objective, if what constitutes a link is well defined prior to the observation session, and that a skilled observer is not required to collect the link data. The primary disadvantages of this method are that it collects superficial data and that it does not include a temporal aspect with the data.

Due to the nature of the TOC and the restrictions placed upon an observer in the TOC, the superficiality of the data is not considered a disadvantage. The observer is not allowed to interact with the participants in a battle simulation; therefore, any observation would have to be superficial in nature.

One of the more interesting applications of link analysis is in police intelligence. Link analysis was introduced in law enforcement in the 1970's as a method of assembling and analyzing intelligence regarding organized criminal activity (Harper and Harris, 1975). Through understanding the relationship between individuals and organizations, it became possible to predict activity and formalize the relationships that eventually lead to identification of complete criminal networks. The successes shown in this article are a testament to the power of understanding group interrelationships.

2.4.5 Narrative Record

The most common usage of data collection in the psychological and social sciences is the narrative record. This method is the most complete observational data collection that can be compiled in a naturalistic setting, and the only resource it requires is, typically, a single observer. Many ethnographical studies use this method because of its completeness and its lack of reliance on interacting with the observed. Unfortunately, the narrative record represents a purely qualitative record of events and typically involves some level of observer interpretation of the events being studied.

Field notes are a form of narrative record. Sometimes these notes are called "written narratives". Due to their written nature, field notes are typically high-level notes. However, it is possible to expand the comprehensiveness of the notes by using voice recording devices to record the field notes (Shaughnessy and Zechmeister, 1985). The observer must be conscious of the observational situation, if voice notes are to be used. If the participants being studied can hear the observer making the notes, the observer runs the risk of influencing the situation he/she is studying.

Illustrations are another form of narrative record. Illustrations are very detailed notes of certain occurrences within the observational session. These occurrences must be clearly defined before the start of the observation sequence and must be of relatively short duration (Shaughnessy and Zechmeister, 1985).

Shaughnessy and Zechmeister give the following guidelines for using narrative records in research:

1. The level of detail for each observation session must be predetermined.
2. Keep each observation session short to increase the accuracy of the notes.
3. Determine who the key players and what the key events are and focus on them.
4. Use a tape recorder to take notes, if at all possible.
5. If observer interpretation is allowed into the notes, multiple observers are necessary, and only those interpretations that all observers agree upon can be included in the final research.
6. If notes cannot be made at the time of observation, they must be recorded as soon as possible, for accuracy's sake.

These guidelines are included because of the infinite amount of information that can be collected from the study. First, the level of detail must be pre-determined. This entails determining whether macro-level events on the large scale or micro-level events involving a relatively small subset of the population will be recorded.

Second, the volume of information requires short data collection intervals because of human memory capacity limits. Short intervals allow the researcher to remember the details of the events in order that they may be fully documented. The use of recording equipment, such as a tape recorder, can serve to lengthen the observation interval by providing a means of documenting the events faster, perhaps even in real-time.

Moore (2002) elaborates on the sixth guideline by stating the pros and cons of transcribing notes from memory. The principle advantage of relying on memory for notes is that

this method further obscures the observer's presence, and thus reduces the risk that the observer is influencing the outcome of the study. However, there are some serious drawbacks to relying on human memory to record a session. First, human memory is far from infallible. Additionally, human memory becomes less reliable with greater amounts of information and greater expanses of time. Thus, relying on memory effectively shortens the amount of time an observer can spend in each observation session (Moore, 2002).

Finally, in order to preserve the objectivity of the study, observer interpretation is typically not allowed into the narrative. If interpretation is allowed, there must be multiple observers present and making interpretations. The interpretation is considered valid if all observers agree on the interpretation of the events in question. Validity concerns arise when observers disagree on interpretations. In essence, the extra observers serve as a validation check on any interpretation made in the course of the observation.

While narrative records serve well to give a description of the events occurring in the study, they do not allow for quantitative analysis very easily. Perhaps this is the reason physical science and engineering researchers do not often employ this method in their research. While this method, in entirety, might not be useful in the evaluation of the TOC, there are several parts of this method that can serve this project.

First, the guidelines that Shaughnessy and Zechmeister give can be applied to other forms of observational research. These guidelines point to a major problem with investigating TOC performance. For observation notes to be accurate, the observation sessions must be kept relatively short. This poses a problem in this research in that TOC exercises can possibly run for 24 hours a day for several weeks. This weakness in observational methods must be overcome, in order for these methods to be useful in TOC research. Second, this method of observation could be of particular use in preliminary studies and to acquaint those unfamiliar with TOC operations with the subject being studied. These descriptive details can be used to develop the key observational events occurring within the TOC operations.

Narrative records utilized relatively infrequently in current research, due in part to the overwhelming subjective nature of the method. However, this method received some attention in the 1990's in a study of mother-infant relationships (Biringen, 1990). Biringen believes that too much information is lost in certain research settings when purely objective, quantitative data is collected. Biringen asserts that the context that is provided with narrative records adds depth to the research that can aid in the interpretation of the events studied.

While the subjective nature of this method may disqualify it from use in the final research product, the method may be used effectively to determine the focus of the research. Narrative records can also facilitate easy collaboration among researchers that would before require all researchers be present in the preliminary observations.

2.4.6 Protocol Analysis

Protocol Analysis, often termed as a "think aloud" method, is often used to elicit methods of operation and problem-solving strategies and methodologies from individuals (Ericsson and Simon, 1984; Kleiner and Drury, 1998). While this method is similar in methodology to conversation analysis, there are some differences. First, while conversation analysis analyzes interpersonal communication, protocol analysis typically only involves an individual verbalizing his/her thought processes and actions. Second, conversation analysis researchers typically make permanent recordings of conversations, and those recordings are later coded and analyzed (Bull and Roger, 1989). Finally, it is not absolutely necessary in protocol analysis to observe the process or require an on-going narrative during the session; it can be used as a reflective technique, where the participants in the study are asked to recall their procedures and processes (Takemura, 1993). In protocol analysis, the researcher often directly codes from the observation session and never has a chance to reexamine the observation session (Ericsson and Simon, 1984). This is of concern, because most implementations of this method call for a single observer per study group.

The protocol analysis method is often used to research decision-making processes and strategies (Kleiner and Drury, 1998). One study that was reviewed researched the decision-

making process involved with individuals deciding to apply for jobs or not to apply, based on information on the job posting (Barber and Roehling, 1993). While much information can be elicited using this method, the content of the verbal report is not the only source of information concerning the process. Significant information can also be elicited from non-verbal artifacts (Bainbridge and Sanderson, 1995). These non-verbal cues are not typically incorporated in protocol analysis.

2.4.7 SYMLOG

The Systematic Multiple Level Observation of Groups (SYMLOG) technique is one of the more powerful and widely used methods for group observation. It has been applied in almost every conceivable team or group situation. Primarily, the method focuses on identifying the behaviors shown by the group members, and is thus a descriptive tool (Polley and Stone, 1988). SYMLOG classifies members of teams and groups along three dimensions. The dimensions are forward and backward, up and down, and positive and negative. Forward and backward refer to the member's behaviors relating to that member's freedom of emotional expression; up and down refer to dominance and submission; positive and negative refer to friendly and unfriendly behaviors (Bales, Cohen, and Williamson, 1979).

One study, performed with Naval teams, sought to find the differences between average and superior commands, using SYMLOG (Bachman, 1988). Bachman administered a questionnaire to the team members in a number of commands that had been established by military standards as being either average or superior. This questionnaire asked each team member to rate the commanding officer's (CO) behaviors. The questions were descriptions of behaviors, such as is the CO positive and outgoing, unfriendly, and egocentric. The team members also rated other team leaders on the same questionnaire.

Bachman's study found significant differences in the behaviors displayed by the CO in the two groups. Average teams showed a significant amount of behavior variety among the team leaders, whereas the superior teams' leaders displayed similar behaviors. Superior teams were led by commanders with very strong behavior sets, and the leaders in the chain of command

beneath the commander displayed the same behaviors as the commander, though they were not as dominant (Bachman, 1988).

2.5 Selected Non-observational Methods

This section contains methods for eliciting information concerning events that occurred in the natural environment. These methods are not truly naturalistic in nature, but they can be used to learn about the processes and events from naturalistic settings.

2.5.1 Critical Incident Technique

The Critical Incident Technique is a post-hoc form of protocol analysis; hence it cannot be used real-time in the TOC. It focuses on salient, rarely occurring events (Kirwan and Ainsworth, 1999). This method could provide insight into the decision-making procedures, other cognitive processes, as well as physical activities performed during an event (Prunkl and Boyles, 1968). As such, this method could be useful in determining indicators leading to the incident that could be observed during operation.

This method would be limited by the memory and objectivity of the participant. It, also, requires the investigator to have sufficient knowledge of the event to determine that it is of interest. On the positive side, this method requires minimal resources. Typically, only one observer and the time to document the session is needed.

2.5.2 Interview

Interviews come in many different forms, but they all accomplish the same goal of eliciting information from participants in an interactive forum. Interviews can be structured or freeform. They collect data on individual or group attitudes, opinions, and knowledge (Fontana and Frey, 2000). Primarily, interviews would be used pre- or post-hoc, and could not be used as a real-time collection instrument. Therefore, their employment would still classify as

unobtrusive to the situation being queried, because there is no observer interference in the subject of the study.

There are several benefits to employing interviews in research. First, it allows interactive querying and capture of non-verbal cues. Participants also have a tendency to be more complete in their answers in an interview, as opposed to a questionnaire. However, this method does rely on the individual or group memory of the events being questioned, and anonymity is not preserved, which can affect results in certain situations.

One example of interviews being used in team research was in a study to determine the relationship between team learning, psychological safety, and team performance (Edmondson, 1999). Edmondson conducted interviews with at least one member of each team he used in his study. These interviews were used to determine if each team's functions would facilitate the research. In this case, the interview was used as a selection tool. However, interviews can be used to accomplish other objectives, as well as facilitate other observation methods.

2.5.3 Questionnaire

Questionnaires are considered observation techniques in much the same fashion Critical Incident Technique is. The questionnaire requires reliance on the participant's memory of an event or exercise. They are useful in collecting opinions, attitudes, and knowledge from people (Kirwan and Ainsworth, 1999). However, the greatest weakness of the questionnaire is its general inability to infer causal relationships (Shaughnessy and Zechmeister, 1985).

Questionnaires are popular in research because they are extremely easy to administer and later analyze, but there are also several pitfalls to using this method. First, truthful answers are not always guaranteed. Second, questions may not have the same meaning between participants. Finally, responses to the questionnaire are sometimes hurried, limited, or non-existent, especially when the administrator is not present when the questionnaire completed (Kirwan and Ainsworth, 1999). However, there are some techniques that can be employed to increase the usefulness of the data returned on a questionnaire.

Kirwan and Ainsworth suggest the following guidelines for creating the items on the questionnaire. Use simple, positive sentences that do not employ technical jargon or acronyms. Take precautions to eliminate bias in the items; a trial run of the questionnaire is beneficial in identifying and eliminating biased items. Additionally, make a point to ensure the confidentiality and anonymity of the item responses to help ensure honest answers (Kirwan and Ainsworth, 1999).

For those who need an in-depth guide for building almost any type of questionnaire, one of the best resources for questionnaire construction, administration, and analysis is the Questionnaire Construction Manual produced by the US Army Research Institute for the Behavioral and Social Sciences (Babbitt and Nystrom, 1989).

Questionnaires have been used in attempts to measure performance. In the 1970's, the U.S. Army developed a questionnaire to help measure the effectiveness of the war-fighting units. This questionnaire elicited performance, job satisfaction, and morale data from individual soldiers (Downey, Duffy, Shiflett, and Helme, 1979). The questionnaire was designed to be administered after a Field Training Exercise (FTX). This data was self-reported and completely subjective. The results of this pilot study were to be used in a much larger study to help determine a holistic method to measure performance. However, no further mention of this project exists in the non-classified literature.

While the Downey et al. study seemed to hit a dead end, the questionnaire method of performance measurement surfaced again in the 1980's. The next questionnaire was designed to determine subjective categories of data that correlated with performance (Olson and Borman, 1987). Olson and Borman's questionnaire collected data on job satisfaction, resources / tools / equipment available, support available, utilization of skills, perceived job importance, unit cooperation, and unit cohesiveness. If correlations were found between the categories and C^2 performance, such a questionnaire could be used as a leading indicator of performance.

2.5.4 Role Analysis

Role Analysis is a process by which the relationships within an organization are mapped and analyzed. The method is used to identify variances in the organization's expected and perceived structures (Dayal, 1969; Hendrick and Kleiner, 2001). These variances can be critical to maintaining or improving job satisfaction, and thus performance to a degree. Each member of the team has a conception of what his/her role is in that team, and that role is accepted or rejected to varying degrees. A higher degree of acceptance of the conceived role leads to higher member performance (Morris, 1971).

There appears to be an underlying assumption in the literature that there is an optimal organizational design for each situation. Role analysis can be used to evaluate the current organizational design as well as identify variances between the intended design and the perceived roles of the team members. Identifying and controlling these variances increases the effective control over the organizational structure to ensure that the intended structure is truly that which is implemented.

2.6 Data Collection Taxonomy

The literature review provided the basis for developing an observation taxonomy that could be utilized in the command and control research. This taxonomy represents a summary of the observation literature reviewed, and presents an answer to the first two research questions. This taxonomy contains observational methods that could be employed in command and control research, but is by no means a completely exhaustive list. As was evident in the literature review, observational techniques are continuously being applied in new ways, and this taxonomy only represents the observational methods as they are or have been used in observational team research. This taxonomy also illustrates the types of data and the tradeoffs associated with each method. Figures 3, 4 and 5 display this taxonomy on the following three pages.

Observational Methods					
Data Collection Technique	Intrusiveness	Costs/Benefits	Relevant Studies	Measures	Source
Activity Sampling	Low to moderate intrusion, depending on method of collection	Fast, easy, cheap method of data collection		Amount of time spent on activities	
		Minimizes number of data collectors by allowing multiple data collection areas for each observer	Upenieks, 1998	Workload (indirectly)	Kirwan and Ainsworth, 1999
		Objective/Quantitative data		Relative workload/utilization	
Conversation Analysis	Low to moderate intrusion, depending on method of collection	Requires considerable time and resources to perform (collecting conversation, transcribing conversation, coding transcription, analyzing codes)	Olmstead, et al., 1973 Roger, 1989 Thomas, 1989 Rutter, 1989	Content of communication (verbal/written)	
		Objective/Quantitative data		Type of communication occuring and builds communication models	Heritage, 1989
Interaction Analysis	Low to moderate intrusion, depending on method of collection	Requires little/no instrumentation and, typically, very few observers			
		Requires full attention of observers participating			
		Robust: definition of interaction can be defined	Flanders, 1965	Interaction types and frequencies	
		Removes most subjectivity from data collection	Cheffers, 1972	Builds group/team communication models	Flanders, 1965
		The number of observers required depends on the number of interactions studied			
		Objective/Quantitative data			

Figure 2 - Observation Taxonomy (Part 1)

Observational Methods (cont.)					
Data Collection Technique	Intrusiveness	Costs/Benefits	Relevant Studies	Measures	Source
Link Analysis	Low intrusion	Requires minimal number of unskilled observers Definition of "link" is flexible Objective/Quantitative data	Harper and Harris, 1975	Identifies relationships within a system Infers behaviors and important relationships Illustrates interaction patterns	Wilson and Corlett, 1995
Narrative Record	Varies, depending on the measures taken to disguise the observer	Requires minimal resources and often a single observer Subjective/Qualitative data	Biringen, 1990	Describes the situation, including the context, encountered in detail	Shaughnessy and Zechmeister, 1985
Protocol Analysis	High intrusion	Requires minimal resources and often a single observer Allows interactive querying Objective/Qualitative data	Barber and Roehling, 1993	Processes (cognitive and physical) and sequences	Bull and Roger, 1989 Takemura, 1993 Ericsson and Simon, 1984
SYMLOG	Not intrusive to acts being questioned	Requires substantial training and moderate instrumentation Allows flexibility in data collection, but not in output categories Subjective/Quantitative data	Bachman, 1988	Behavior traits of group/team members Shows the behavior relationships between group/team members	Bales and Cohen, 1979 Polley and Stone, 1988

Figure 3 - Observation Taxonomy (Part 2)

Non-observational Techniques					
Data Collection Technique	Intrusiveness	Costs/Benefits	Relevant Studies	Measures	Source
Critical Incident Technique	Not intrusive to acts being questioned	Requires minimal resources Relies on subject's memory of events Subjective/Qualitative data	Prunkl and Boyles, 1968	Gathers narrative information on specific past events	Kirwan and Ainsworth, 1999
Interview	Not intrusive to acts being questioned	Requires little instrumentation, but substantial time Allows observation of non-verbal cues and interactive querying Subjects tend to be more involved in answering the questions compared with the questionnaire, but they are limited by their memory Subjective/Qualitative data	Edmonson, 1999	Problem identification Probes individual and group opinions, attitudes, and knowledge	Hendrick and Kleiner, 2002 Cooper, 1974 Fontana and Frey, 2000
Questionnaire	Not intrusive to acts being questioned	Requires minimal resources to perform Participation is often not high and relies on subject's memory Does not allow interactive querying Method is subjective, but can be quantitative or qualitative	Downey, Duffy, Shiflett, and Helme, 1979 Olson and Borman, 1987	Identifies problems Measures the effects of changes being made Probes individual opinions, attitudes, and knowledge Can be used to predict performance	Hendrick and Kleiner, 2002 Downey, Duffy, Shiflett, and Helme, 1979
Role Analysis	Not intrusive	Requires moderate resourses, mostly time Involves the participant in the problem-solving process Subjective/Qualitative data	Morris, 1971	Variances in the expected organizational role and the experienced organizational role	Hendrick and Kleiner, 2002 Daval, 1969

Figure 4 - Observation Taxonomy (Part 3)

2.7 Instrumentation

One way to reduce participant reactivity to the observer is to have the observer not take notes in the presence of the participants (Hickson, 1974). This would require the observer to rely on his memory to retain his field notes until he was afforded time to record them. Hickson (1974) believes this is possible as long as the recording takes place as soon after the observation session as possible. In fact, Kirwan and Ainsworth (1992) assert that most observation now employs electronic recording devices.

Taking mental field notes is not an option for this project, due to the complex nature of TOC operations and the long duration of the observations, but there are other ways to decrease the appearance of note taking. For example, in a noisy environment, such as a TOC, a discrete tape recorder and a personal microphone could allow the observer to take audio notes. The observer's voice would be lost in the background noise, but would be recorded. If sufficient background noise is present, speaking softly should be far less intrusive than stopping the observation to look down at a pad and frantically scribbling notes. This method also solves the problem of selective attention in the observer. An observer in a naturalistic setting must continually look for the critical events he seeks and record them (Moore, 2002). Most recording methods currently used require the observer to look away from the transpiring events to document the event. With audio notes this distraction is minimized, though not completely eliminated.

Another example of a less intrusive recording device is the use of covert electronic surveillance. This could be audio, video, or audio-visual. These methods completely eliminate the selective attention problem by removing the observer from direct observation. Due to the level of noise present in a typical TOC, ambient audio recording is not feasible (Kirwan and Ainsworth, 1999). However, video surveillance is feasible. Video surveillance offers many advantages. A video record of an observation allows the observer to review events as many times as necessary (Fisher, Hunter, and MacRosson, 2001). This is particularly helpful in the beginning of an experiment, when the observer might not know exactly what to look for or has underestimated the importance of a dismissed event, and when complex events, defined here as a

series of events worthy of recording occurring in a very brief time, occur (Molseed, 1994). Molseed (1994) further points out that cameras used for observation are often quickly and completely forgotten about during the course of the observation. Much as the participants grow accustomed to the observer as time passes, they grow accustomed to the camera. Since the camera makes no noise or motion to draw attention to it, once it is forgotten, it may not be remembered before the exercise is concluded. Because it does not move, the camera has the disadvantage of having a fixed field of view and a fixed focal point.

If reactivity is not a prime concern in the observation sequence, Kirwan and Ainsworth (1992) suggest that the best method of collecting data is to use a combination of electronically recorded data and written notes. This would be particularly helpful if the observer is keenly aware of the operations involved in the group being observed. While the recording gathers all of the action in its field of view, it cannot capture some, more subjective, data. The observer present, if he/she is familiar with the situation, can make written notes of an obscure or unexpected event sequence that would later help explain the events taking place on the permanent recording.

While there are many advantages to using electronic recordings of observation sessions, there are some disadvantages as well. While the data collection will be potentially more complete with a recording of the session, the analysis time increases substantially (Kirwan and Ainsworth, 1999). Additionally, for high quality recordings, there is a need for high quality equipment. This equipment represents an added expense to the research project.

Sanson-Fisher, Poole, Small, and Fleming (1979) present the following, applicable guidelines for recording equipment:

1. It should be robust and reliable in the observation environment, as well as being portable and as unobtrusive as possible.
2. It should rely on battery power and utilize a permanent storage medium.

3. It should be able to record at least four hours of data without requiring a medium or battery change.

Due to the length of many battle simulations, power and storage capacity are of great concern. However, these concerns may be minimized by thoroughly evaluating the opportunities available with the equipment being utilized. For instance, the use of digital recorders can increase the storage capacity and decrease the power usage of the equipment. Some equipment can also use power packs that enable large power storage. Many of these simulations occur within a close proximity to a powered building, so there is easy access to power just outside the TOC area. This allows for certain recharging and storage solutions that would be unavailable in the field without a mobile power source.

Modern electronic equipment has the potential to greatly aid in performing naturalistic observations, as it makes observing participants in the natural settings not dependent on the observer actually being in those settings. These instruments are, essentially, the observer's eyes and ears in the observation. There are several to choose from, and each must be fitted to the environment being observed. Once the method and instrumentation is chosen for the study, it is important to maintain that method and instrumentation throughout the study, because no two methods or instruments will obtain the same level of detail. This could lead to serious discrepancies in the data being analyzed. It is important to realize that, because of these discrepancies, the same process observed with differing methods and instrumentation cannot be compared (Peregrine, Drews, North, and Slupe, 1993). This issue will not present a large problem in this study, because each study will commence and terminate in the specific TOC being observed.

Chapter 3 - Methodology

3.1 Purpose

The purpose of this study is to develop and demonstrate a method for employing direct observation techniques to evaluating team processes. This study uses a military Command and Control (C^2) team for testing and validation purposes. However, the methodology developed in this study is applicable to any high-pressure, tactical decision-making team in the military and civilian realms.

Past observational research has demonstrated that direct observation methods provide an accurate data collection method that can easily be employed in the field. Because the environment is a battle simulation, the observation methods explored earlier are applied in the field study realm. Further, the observational nature of this type of evaluation is a deviation from the current, self-assessment methodology. Based on the literature reviewed, the following research questions have been formed:

1) How can direct observation methods be used to evaluate military command and control teams?
    a) What data types can be collected in the TOC using direct observation?
    b) What are the tradeoffs among the different observation techniques?
    c) How can the potential data be collected and analyzed?

Based on previous research, the following hypothesis has been formulated:

1) Direct observation techniques can be employed to explore the command and control team processes in a quantitative manner with non-expert observers.
    a) The data collected will be reliable across observers.
    b) The data results will be equivalent to the command team's self-assessment.

The questions concerning the data types and tradeoffs have been answered based on the literature findings, and they have been summarized in the Observational Taxonomy. The remaining sub questions and hypotheses are the subject of the remainder of this study.

3.2 Procedure

3.2.1 Battle Simulation

This study made use of a U.S. Army battle simulation exercise to test the proposed methodology. The type of battle simulation used in this study was a command and control training exercise. In such an exercise, the war-fighting units are simulated, but the command and control processes are kept as real as possible. The following figure illustrates the layout of a typical command and control battle simulation.

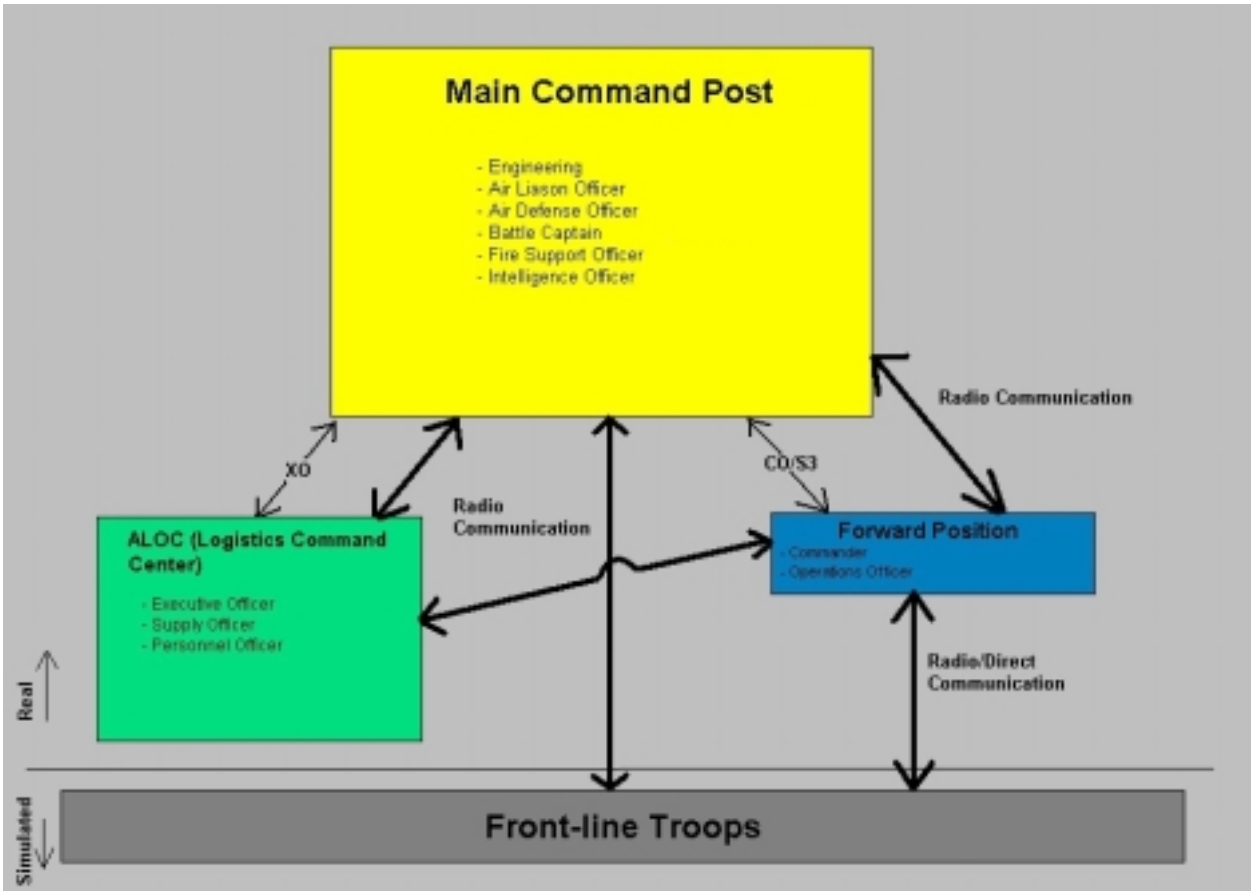


Figure 5 - Battle Simulation Concept

The people participating in each element of the simulation can change, depending on the type of command team being trained and the type of training being performed. The positions listed in this figure are the positions used in this battle simulation. However, some positions were filled by the same person, due to staffing limitations.

Each block in this figure represents a physical location in the battlefield. The front-line troops are manipulated in the software of the simulation by human operators. These operators absorb the artificiality of the simulation. This allows the command team to experience the reality of battle to the highest possible extent. Everything above the front-line troops operates much as it would in real battle.

The main command post is the largest of the brigade command structures on the battlefield. It is here that intelligence information is received and processed and long-term plans are made. This structure also serves as a communications hub, facilitating communication among higher, lower, and adjacent commands and even commands in other service branches, such as the Air Force.

The Army Logistics Command Center (ALOC) typically has a much smaller staff than the main command post. The ALOC is charged with ensuring friendly forces are adequately supplied and staffed. This is accomplished in two ways. First, supplies and/or personnel can be requested, typically by the Commanding Officer (CO), the Operations Officer (S3), or the main command post. Second, the ALOC staff can be proactive by keeping apprised of the battle situation by monitoring radio traffic. By hearing the reports from the battlefield, the staff can anticipate the needs of the friendly forces.

Finally, the CO and the S3 are typically located somewhere on the battlefield. These individuals coordinate the battle on a local level. They have direct access to both the front-line troops and the command structures. In the physical layout of the battlefield, the CO and S3 are mobile, operating out of wheeled vehicles, and are usually located somewhere between the command post and the front line.

There are many forms of communication available to the command team. In this exercise, communication is primarily based on handheld digital radios, which emulate SINCGARS. A standard telephone is used for infrequent communication that would occur over the DNVT tactical telephone. In addition, the CO, S3, and XO occasionally interface directly with the main command post, and the CO and S3 can also interface directly with the front-line troops.

The simulation also simulates two flank brigades. These are two other groups, of equal size to the brigade being trained, that occupy the same battle theatre. In the simulation, these brigades are virtual brigades. Typically, there are one or two people playing each of these brigades, but rarely do these brigades engage in combat in the simulation. To aid in the visualization of the battlefield, the next page contains a hierarchical chart for the U.S. Army. The levels of command that are relevant to this exercise are highlighted. The TOC exists at the brigade level. Communications below the brigade level are to individual unit commanders in the field.

As you travel down the hierarchy, orders transition from high-level strategic orders to more practical and specific objective orders. If an entire numbered army were deployed into a battle theatre, there would be a command center for each level above the battalion level. Each successive command center would facilitate this transition. The brigade TOC represents the last of the transition of the orders before the units actually execute the orders on the battlefield.

3.2.2 Participants

The command and control team was populated with participants from the Battle Simulation Center's (BSC) contractor pool. These contractors train active duty personnel on the use of the simulation, and in most cases are retired military personnel with extensive experience in command and control. While these individuals are considered experts in command and control, they do not often work collectively as a command and control team. Individuals participating as observers were taken from the BSC's personnel and from the Army Research.

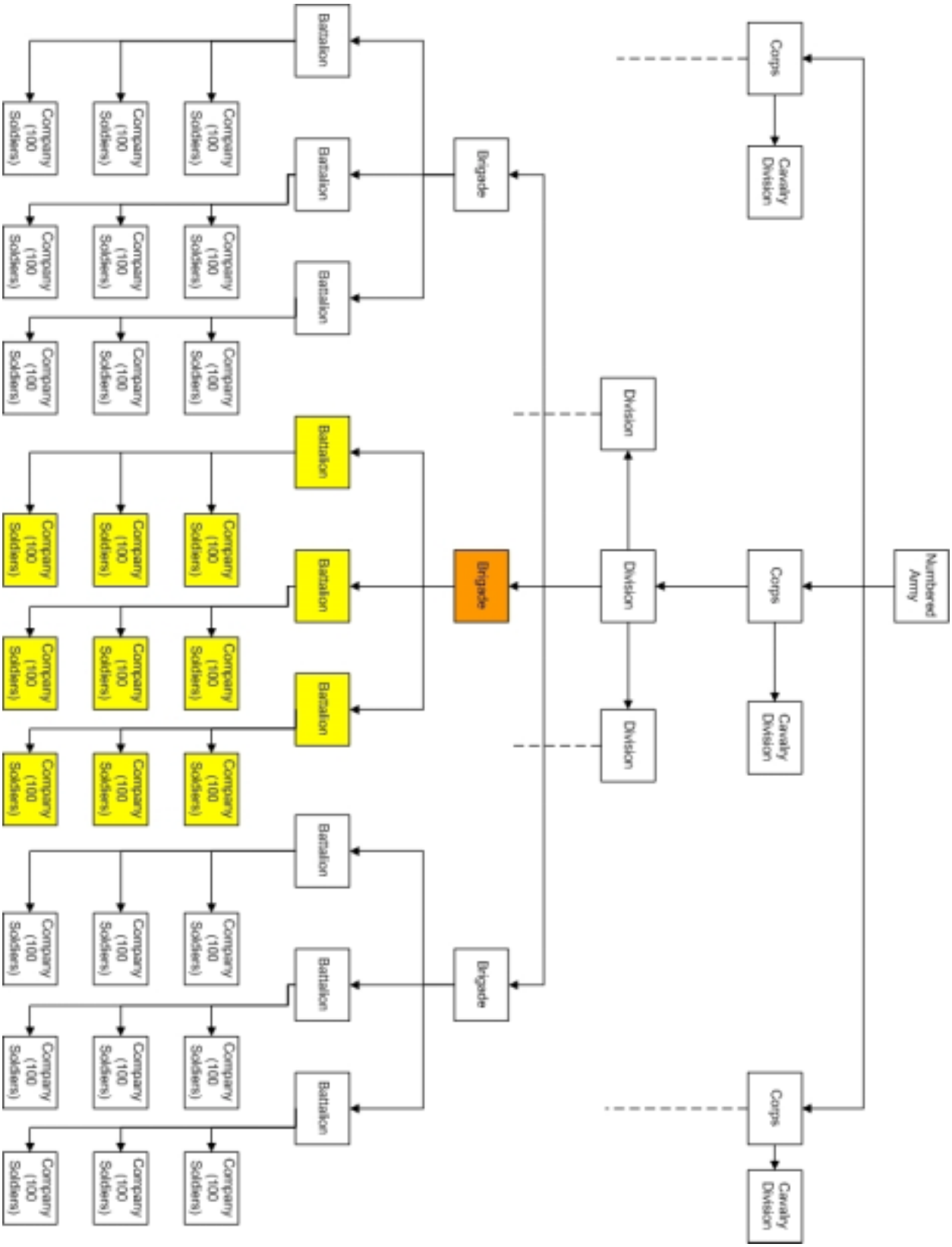


Figure 6 - U.S. Army Force Structure

Lab: Ft. Hood Field Element. Sixteen individuals participated in the command team and three participated as observers.

The observer team bears a direct relevance to this study, because the bulk of the data was collected with these individuals. The three observers used in this study had diverse backgrounds. One observer was a 66-year old, Caucasian male with no active-duty military experience and little experience with TOC operations. Another observer was a 36-year old, African-American female, who was an active-duty E-7 at the time of this study's publication. The final observer was a 24-year old, Caucasian male, with several years of military training as an officer and no experience with Army TOC operations. It should be mentioned that the observers were not selected based on any qualifications, and the three observers were the only three available to participate as observers.

The age of the remaining participants ranged from 28 to 62 years of age (mean = 49.9; sd = 12.5). Since military commanders have historically been male, especially in the generations represented by the participants, all of the command team participants were white males. However, everyone in the contracting staff was given the opportunity to participate in this exercise, regardless of gender. All of the individuals comprising the command team had active-duty military experience. The leader of the contracting staff was the Commanding Officer, and he assigned the other positions based on individual experience. This was an ideal way for the participants to be assigned a position, as the leader of the contracting staff was the best qualified to know the individual strengths and weaknesses of the contracting staff. Since this exercise was used by the Battle Simulation Center as a training exercise for the contracting staff, the participants were not compensated for their time by this researcher. This exercise was considered a part of their employment with the Battle Simulation Center and the participants were considered "on the clock" during the exercise.

The participants were informed of the purposes of this study prior to the first day of the exercise and were required to complete an informed consent form before the exercise began. Since this study was not the only study performed using this exercise and the fact that time was

limited on the exercise days, the informed consent procedures were completed prior to the first day of the exercise.

3.2.3 Observer Training

The observers underwent a three-hour training session on the usage of the data collection tool. This training session was divided over two days and was completed the day before the simulation exercise began. The observers were given a copy of the tool, as well as the manual that accompanies the tool, and trained in the usage of each.

Usage training consisted of going through the items, one by one, and discussing the meaning behind each item and possible observable behaviors for each item. The observable behaviors were listed from examples given in the literature, where they could be found. Other behaviors were listed from the common experience of the researcher and the other two observers, based on the intent of the item. The intent was to relay the overall intent of the item, not to dictate the specific behaviors to look for. The manual was used to keep a record of sample behaviors. The second part of the usage training session required the observers to complete the tool based on a written scenario. Using a written scenario allowed the observers to practice the observation skills necessary to complete the tool in the C^2 environment. This scenario is contained in Appendix C. After the observers had rated the scenario, the ratings were disclosed and discussed. Where discrepancies were detected, the discussion took on a more normalizing tone. The reason for each observer's rating was openly discussed in order to gain insight on the rationale behind the rating so that normalizing actions could be formulated.

This part of the training program accomplished two objectives. First, it provided the opportunity for the observers to acquaint themselves with the tool and the manual. Any questions they had concerning the intent of particular items were discussed until all observers were satisfied they fully understood the item. All questions were noted, so that the manual could be revised to alleviate the uncertainty. The second accomplishment of the training session was the evolution of the manual that accompanied the tool. The manual was designed to accept input

from the observers. Throughout the training session, the observers were asked if there were any example behaviors that should be added or removed from the manual.

3.2.4 Facilities

This field experiment was performed at the III Corps Battle Simulation Center (BSC) at Ft. Hood, Texas. The BSC is a state-of-the-art facility for combat simulations. This field study used the command and control training portion of this facility.

This facility contains a mockup of a field Tactical Operations Center (TOC). The communications equipment within the TOC is linked directly with computer operators in the simulation rooms. As such, the command and control team within the TOC interfaces with those controlling the simulation in the same way they would in actual combat. However, instead of maneuvers being performed in the field, the computer operators execute the commands in the simulation.

3.2.5 Equipment

The major equipment used in this experiment was the simulation equipment and the command and staff equipment in the TOC. The actual specifications on the computer driving the simulation are unknown. However, the interfaces to the simulation were via 14 PC workstations. BSC personnel manned these workstations.

The primary equipment used in the TOC was used to facilitate communication with the forward troops, represented in this exercise by the simulation, and included:

1. Internal digital phones – these phones were used to emulate the DNVT tactical telephone and utilized the internal digital lines of the BSC.
2. Motorola hand-held two-way radios – these radios emulated the SINCGARS tactical radio system. Three separate radio nets were utilized in this exercise to support different command functions.

3. Four VISION workstations – these computers served to emulate MCS, ASAS, and FBCB2 (These systems allow a graphical representation of the battlefield). The fourth workstation had the ability to see all three.
4. Two video projectors – these were used to display the MCS and ASAS emulations in the TOC.

3.2.6 Software

The software utilized in this battle simulation was designed specifically for U.S. Army use. The main battle simulation driver was the Corps Battle Simulation (CBS). In addition to this software, the VISION XXI system interfaced with CBS to provide ASAS, MCS, and FBCB2 emulation for the command team. The data collected was entered into a Microsoft Excel spreadsheet. Finally, analyses were performed in Excel, NCSS 2001, and PASS 2002.

3.2.7 TOC Setup

The TOC used in this exercise was setup in a training room at the III Corps Battle Simulation Center at Fort Hood, Texas. A diagram of the command post is included below. Directly following the illustration is a definition of the relevant acronyms used in the illustration for the major groups in the TOC.

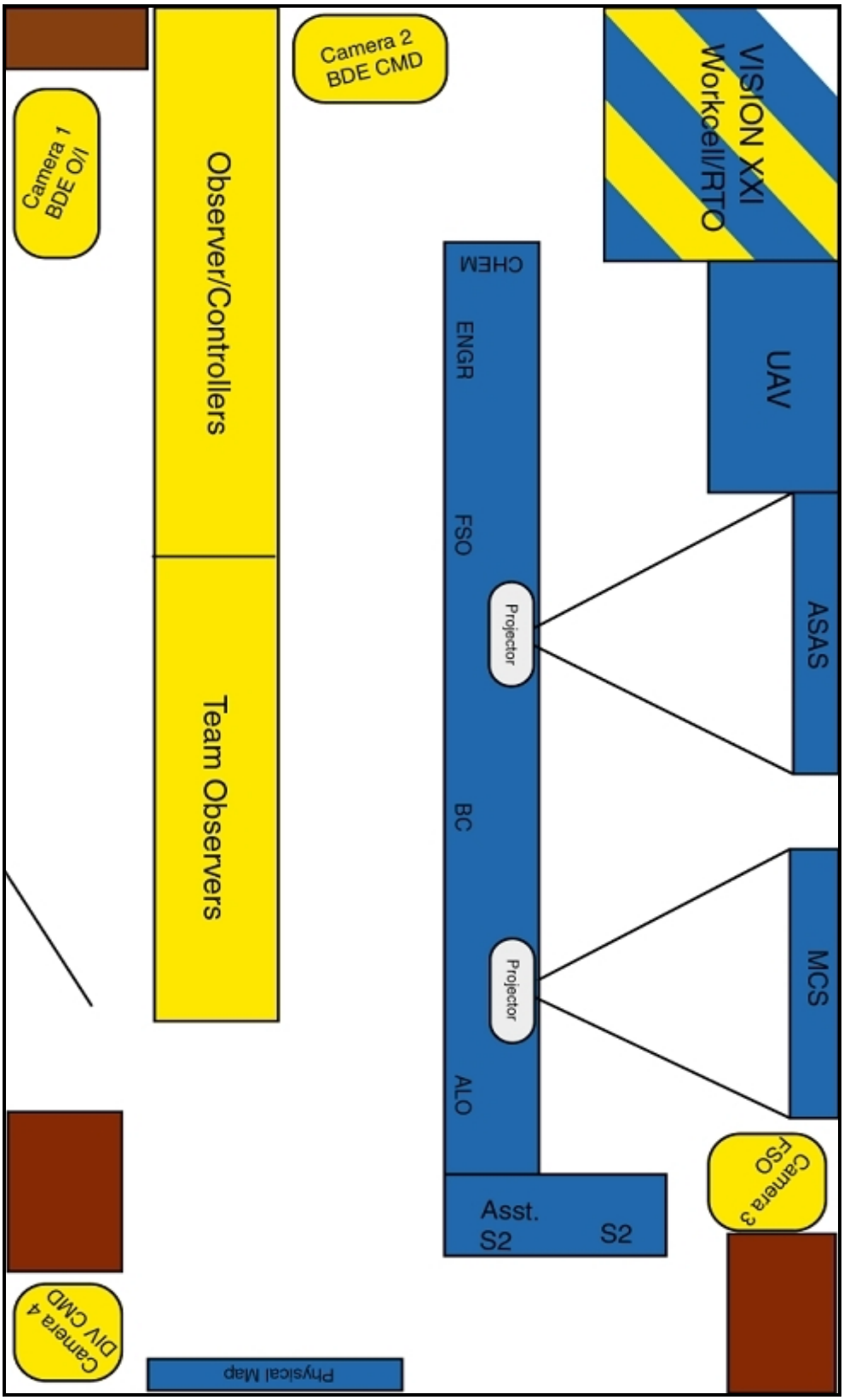


Figure 7 - Omega Challenge Main Command Post Layout

Table 2 - Command Post Group Definitions

Acronym	Definition	Primary Duty
ASAS	All Source Analysis System	A graphical projection consisting of a map of the battle area and the positions of all hostile forces.
ALO	Air Liaison Officer	"The senior Air Force officer at each tactical air control party who provides advice on the capabilities, limitations, and employment of fixed wing aircraft and coordinates close air support missions with the fire support element." (U.S. Army, 1997a)
BC	Battle Captain	The battle captain serves as the leader of the TOC in the absence of the CO and the XO. Since the CO and the XO are rarely in the TOC, the BC is recognized as the leader of TOC operations.
CHEM	Chemical	The chemical officer in the TOC is in charge of ensuring an adequate defense against chemical weapons and authorizing the use of friendly chemical agents against enemy targets.
CO	Commanding Officer	The CO is the leader of the echelon at which the TOC is built. In this study, the CO represents the leader of an Army brigade.
ENGR	Engineering	"Combat engineers are an integral part of the combat arms team. Engineers adapt terrain to enhance the battle effectiveness of fire and maneuver. The orientation of engineers in support of a task force is forward; their efforts are designed to support forward fights." (U.S. Army, 1992)
FSO	Fire Support Officer	"A functional portion of a force tactical operations center that provides centralized targeting, coordination, and integration of fires delivered on surface targets by fire support means under the control of or in support of the force. This element is staffed from the field artillery headquarters or field artillery staff section of the force and representatives of other fire support means." (U.S. Army, 1997a, p.1-67)
MCS	Maneuver Control System	A graphical projection including a map of the battlefield and the positions of all friendly units.
RTO	Radio-Telephone Operator	These individuals receive all communications from outside the TOC and transmit most communications coming from the TOC. This position is charged with converting the information in the message to a usable form for the command team.
S2	Intelligence	"Group tasked with collecting and organizing information and knowledge about an adversary by means of observation, investigation, analysis, or understanding. Also refers to the commander of this group, specifically." (U.S. Army, 1997a, p.1-83)
S3	Operations and Training	"The principal staff officer for all matters concerning training, operations and plans, and force development and modernization and completeness." (U.S. Army, 1997b, p. 4-12)
UAV	Unmanned Aerial Vehicle	An unmanned aerial surveillance vehicle used to report on enemy positions in the battlefield. This vehicle falls under the S2's authority.
XO	Executive Officer	The executive officer is the second in command in the echelon that is being simulated. In this study, the XO is the second in command of an Army brigade.

3.2.8 Experimental Procedure

The simulation progressed through a different phase of combat each day:

1. Day 1: Position troops and perform recon and counter recon actions.
2. Day 2: The troops attack the opposing forces (OPFOR).
3. Day 3: Perform a river crossing.
4. Day 4: Consolidate forces and defend against OPFOR attack.

While the simulation progressed through the different stages of combat, the team observation procedure remained the same for each day of the exercise. The overall timeline of the day's activities is given below:

1. 08:00 – 09:00: Perform technical setup and checks. During this time, the command staff assured that everything was prepared for the day's simulation. The observers were given the assessment tool and took their places within the TOC to observe.
2. 09:00 – 16:00: Play the war game. It is during this phase that the team observation took place.
3. 16:15 – 17:00: Perform After Action Review (AAR), collect data from various sources, and setup for next day's exercise. An AAR is a meeting of the command staff to discuss the team's performance. It is during this time that the commander and section commanders were given the opportunity to complete the assessment tool. The data sheets were then collected from the observers and commanders and stored for analysis.

The observers observed the team's functioning throughout the exercise. At the end of each hour of the exercise, the observers completed the assessment tool based on the previous hour's observation. The researcher and the observers were instructed to not socialize with the command participants during the exercise, in order to reduce observer influence.

3.2.9 Team Assessment Tool

To facilitate the collection of the observation data, an input form has been devised for the observers to use. The instrument has been developed such that the following team process attributes (from Swezey and Llaneras, 1997) are measured:

1. Coordination
2. Communication
3. Conflict Resolution
4. Decision Making
5. Problem Solving
6. Boundary Spanning
7. Joint Knowledge
8. Motivation
9. Personalities
10. Shared Thought Processes

These team process attributes are theorized to capture the elements necessary to study team processes. Unfortunately, one team process, Shared Task Ability, was accidentally omitted from the tool design.

Using the descriptions for each of the team attributes found in the literature, team assessment tool items were devised to collect information regarding each of the attributes. At least two, and not more than seven, items were included in each construct. Theoretically, the greater the number of items in each construct, the better, as the factor analysis and the Cronbach's Alpha analysis will reduce the number of items. However, for simplicity, the number of items in each construct was limited. This simplicity was important since non-expert observers used the tool. Not all constructs have equal numbers of items because several of the constructs overlap. For example, decision-making and problem solving are difficult to separate from a team process point of view, as decision-making is performed in order to solve problems.

A continuous visual analog scale was chosen in an attempt to eliminate bias in the ratings. This scale was developed with anchors only. The observer made a vertical mark on the scale at the point between the two anchors that seemed appropriate. After the data sheets were collected, a ten-point skeleton was overlaid, so the rating could be quantified and recorded. Though the data was analyzed on a ten-point scale, the observers never saw anything other than the continuous visual analog scale.

3.2.10 Analyses

The data collection instrument was given to the three observers in the TOC each day. It is to be completed at the end of each hour of the exercise. The same instrument was given to the team commander and section commanders at the completion of each day's exercise.

Several analyses were performed on the ratings. First, a factor analysis and a Cronbach's Alpha analysis were performed on the team process constructs to determine the internal reliability of the questions comprising each construct. Second, the inter-rater reliability was determined using Kendall's Concordance and a main-effects ANOVA. Finally, the average results from the observers were compared to the results gathered from the team commanders, each day, in order to rate the tool's construct validity. Adjustments are made to the tool as required by the results of the factor analysis. A copy of the data collection instrument is contained in Appendix A.

3.2.10.1 Factor Analysis and Cronbach's Alpha

A factor analysis was performed on the assessment tool results from the observers to ensure internal reliability by ensuring that all factors identified within the construct were factors that directly relate to the construct. Items within factors identified through the factor analysis that did not measure the intended construct were deleted from the tool.

Cronbach's alpha was also used to judge the reliability of each construct. The construct composition was considered reliable only if Cronbach's alpha was at least 0.7 for the items in that construct. Items that reduced the Cronbach's alpha value below 0.7 were removed from the construct. The remaining data analysis proceeded with this reduced dataset. The computational formula for Cronbach's Alpha that is used in this study is as follows:

$$r_\alpha = \frac{N}{N-1}\left(1 - \frac{\sum s_n^2}{s_x^2}\right),$$

Equation 1 - Cronbach's Alpha

where s_n^2 is the variance of each item and s_x^2 is the variance of the total score (Suen and Ary, 1989).

3.2.10.2 Inter-rater Reliability

Each item that survived the item reduction phase of analysis was subjected to an inter-rater reliability test. This test was performed to indicate items that are not consistent across different raters. Items that failed the inter-rater reliability test were not necessarily deleted from the tool, though this is done on occasion. There were a few options concerning what to do with these items. First, the item could have been deleted from the tool. This was done if there were many items remaining in the team construct being measured. Second, the item could have been reworded in an effort to remove the variability. This was done if there were few items measuring the particular team process. Finally, the item could have been noted and the training program adapted in an attempt to minimize the variability.

There was one observation period where one observer missed the opportunity to rate the command team. Since only one rating in 26 possible ratings were missed, the ratings for that hour from all observers were discarded, in order to make the calculations simpler. The inter-rater reliability calculation was further complicated by two factors. First, the data collected was interval in nature, and the researcher did not want to lose the richness that interval data provided. Second, there were more than two observers. The vast majority of inter-rater reliability calculations that have been developed in previous research are designed for nominal or ordinal data collected by two observers. Since no desirable, previously existing inter-rater reliability formula was found, two different methods were used in this study.

The first method used for assessing the inter-rater reliability was Kendall's Concordance (W). This measures the rank-ordered correlations of the ratings. To facilitate this assessment, each observer's ratings were ranked, per item, across the entire exercise. Kendall's Concordance was calculated on an item basis. The formula used in calculating Kendall's Concordance follows (Siegel and Castellan Jr., 1988).

$$W = \frac{12\sum R_i^2 - 3k^2 N(N-1)^2}{k^2 N(N^2-1) - k\sum T_i},$$

Equation 2 - Corrected Kendall's Concordance

where R_i is the sum of the rankings per item, per observer; N is the total number of responses per item, per observer; and k is the number of observers. T_i is the correction factor for ties, which is discussed below.

Equation 2 gives an equation for Kendall's Concordance that is corrected for tied ranks. This correction was necessary because there was a high proportion of tied ranks in the data collected. The proportion may have been abnormally high because the data collection was not designed to collect data on an ordinal level. The correction factor that has been incorporated in Equation 2 is the same correction factor that is used in the Spearman rank-order correlation coefficient for tied ranks (Siegel and Castellan Jr., 1988). Explicitly, the correction is as follows:

$$T_j = \sum_{i=1}^{g_j} \left(t_i^3 - t_i\right),$$

Equation 3 - Correction for Tied Ranks

where t_i is the number of tied ranks in the i^{th} grouping of ties and g_j is the number of groups of ties in the j^{th} set of ranks (Siegel and Castellan Jr., 1988, p. 266).

Once W is calculated for an item and corrected for ties, a chi-square test for independence can be performed on the value. The calculated value of chi-square is as follows:

$$X^2 = k(N-1)W$$

Equation 4 - Calculated Chi-square Value

This value of chi-square is compared to the tabled chi-square value using the calculated Chi-square value and *df* = N – 1. For this test, an alpha level of 0.1 was used to assess significance based on the null hypothesis of independent ratings across observers. For tabled values of Chi-square below this level of alpha, the null hypothesis would be rejected and the rankings would be considered dependent, or sufficiently correlated. A rejection of the null hypothesis would suggest that the inter-rater reliability is sufficiently high for that item.

The second inter-rater reliability analysis performed in this study uses a two-factor, main-effects ANOVA. The two factors are Time Period and Observer. The Time Period factor is in place primarily to distinguish each set of ratings as independent of each other. Significance in this factor could suggest that either the command team or observer opinions changed as time passed. The Observer factor is the factor of interest in this test. Significance in this factor would suggest that the individual observers had a significant effect on the ratings. The ANOVA design is as follows.

Figure 8 - Inter-Rater Reliability ANOVA Design

	Observer		
Time Period	1	2	3
1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$
2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$
…	…	…	…
24	$X_{1,24}$	$X_{2,24}$	$X_{3,24}$
25	$X_{1,25}$	$X_{2,25}$	$X_{3,25}$

Since the purpose of inter-rater reliability is to test whether a tool is stable across multiple users, it is desired that the "Observer" factor not be significant. A lack of significance in this factor infers that the individual observers did not have a significant effect on the ratings. Thus, the item has an acceptable inter-rater reliability.

Since there are no replications per cell, and therefore n = 1, a standard ANOVA cannot be performed, because the error term, subjects, would have zero degrees of freedom. Therefore, the

ANOVA performed was only concerned with main effects, and not interactions. From a results standpoint, this compromise is made without hesitation, because the primary interest is in testing for significance in the observer factor. The interaction mean square error was used as the mean squared error for the main effects by the software performing the analysis, per Montgomery (2001, p. 190). The ANOVA was performed at an alpha level of 0.1.

3.2.10.3 Construct Validity

What is termed here as construct validity is actually somewhere between construct and criterion validity. Ordinarily, a test such as the one conducted here, would be considered criterion validity. However, since there is not a one-to-one comparison possible, due to the fact that the observers completed the tool each hour and the command team only completed the tool at the end of each day, it is not truly criterion validity that is being tested. Instead, this phase of the analysis is termed construct validity because the team process constructs are being compared. The team process construct scores from the observer team are compared to the construct scores from the command team, who produce what is considered more ideal representations of the construct scores.

Construct scores were calculated by taking the average of the scores from the items remaining in that construct. Construct validity was assessed by comparing the results of the assessment tool completed by the observers to those completed by the command team. The average scores for each construct were calculated across the commanders' ratings for comparison. The construct scores on the observers' forms were averaged across all hours of that day's exercise and across both observers. The comparisons were made on a daily basis because the command team only completed the tool once daily. Since the sample sizes in the comparisons are different, the sample size is less than 30, and the interest is in finding a difference in the mean ratings between the observer team and the command team, a two-tailed paired t-test will be used to compare the samples.

After doing a preliminary analysis of the data, it was found that the power of this test was very low, ($\beta > 0.9$). Therefore, a power analysis was performed on the test, using the descriptive

statistics from the comparison of the construct scores. This analysis was used to help determine an acceptable level of alpha to use for the paired t-test. With such a small sample size, a high power is not easily attainable. However, this researcher's opinion is that there should be a balance between alpha and power. The alpha chosen as a result of this analysis is 0.4. The details of the power analysis can be seen in Appendix G.

3.3 Agenda

The facility and the participants were provided for this study by the III Corps Battle Simulation Center and the Army Research Laboratory Field Element at Ft. Hood. They were provided without compensation from this research endeavor because this exercise was scheduled as a training exercise for the contracting staff and as a test for new simulation equipment. There is also another study being performed by ARL in this exercise. The other study being performed by ARL is collecting environmental and equipment usage data and analyzing its relationship with C^2 performance. The study presented in this proposal is an attempt to add measures of the cognitive side of C^2, in the hopes of combining the two studies in the future in order to collect data on command team performance from a more encompassing point of view.

Chapter 4 - Results

This chapter contains the results of the analyses performed in this study. The organization of this chapter follows the order of the analyses performed. First, the number of items was reduced, per construct, using factor analysis and Cronbach's alpha. Second, the inter-rater reliability of each remaining item was tested using Kendall's concordance and main-effects ANOVA. Finally, the constructs that were shown to be reliable from the first two sets of analyses were tested for validity using paired t-tests on the construct scores from the observer team and the command team.

4.1 Item Reduction

It was expected that not every item in the tool would prove to be reliable or adequately measure the construct it was intended to measure. Therefore, the first step in the analysis process for the data collected in this study was to determine the reduced set of data for further analysis. This section details the analyses performed in the data reduction. Item reduction and inter-rater reliability analyses use only the data collected from the observer team. Construct validity analysis is performed using the data collected from both the observer team and the command team.

Once the data were transcribed and corrected for reverse-coded items, a factor analysis (Kline, 1994) and analysis using Cronbach's alpha (Suen and Ary, 1989) was performed on the items in each team process construct. The Cronbach's alpha that is reported in this study is the standardized alpha value. To improve readability, subsections will be dedicated to each team process construct and the complete details of the factor analysis are included only in Appendix H. All factor analyses use a varimax rotation, and factor structures are found using a minimum factor loading of 0.4. Significant factors are factors that have an eigenvalue of at least 1.00. The factor loading of 0.4 was used because lower values produced a bleed-over effect where items commonly appeared in all significant factors. Using this loading criterion clarifies the primary composition of the factors.

4.1.1 Coordination

Items 2 through 7 on the original tool were designed to collect data concerning the team's coordination. The relevant results of the factor analysis follow. Only factors with eigenvalues greater than 1.00 are considered relevant.

Table 3 - Factor Analysis for Coordination

Factor	Eigenvalue	Percent of Variance Explained
1	2.11	99.46

The results of this factor analysis show only one clear factor existing in the items under the Coordination construct. Since there is only one factor present, it is assumed that the factor is, indeed, coordination. Therefore, no items can be automatically deemed irrelevant to the construct.

The next analysis performed utilized Cronbach's alpha. This analysis was performed to determine the reliability of the items within the construct. Since all items in the coordination construct were relevant items, all six items were included in this analysis. The results of the initial Cronbach's alpha analysis follow.

Table 4 – Initial Cronbach's alpha for Coordination

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q2	8.274194	1.590642	0.6147	0.5965
Q3	7.596774	2.198805	0.5943	0.5975
Q4	8.564516	1.313498	0.6302	0.5936
Q5	8.209678	2.041673	0.6517	0.4480
Q6	7.887097	2.240498	0.6609	0.4290
Q7	8.032258	1.924542	0.7653	0.0644
		Total:	0.698	

The initial criterion for rejection was that any item bringing the construct alpha below 0.7 would be removed. The items in this construct do not quite meet this criterion, with an alpha value of 0.698. The number of items in this construct needs to be reduced for two reasons. First, since all items are measuring the same factor, they are, for the most part, redundant. Second, deleting items will raise the Cronbach's alpha to an acceptable level. To achieve the reduced item set, the item that contributed the most to lowering the alpha level was removed and the calculation was repeated. This procedure was repeated until the calculation showed that the further removal of items would reduce the alpha level. The final calculation left Items 3, 4, and 5 in the construct, with an overall Cronbach's alpha of 0.72

Table 5 - Cronbach's Alpha Analysis on Reduced Coordination Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q3	7.596774	2.198805	0.5231	0.6238
Q4	8.564516	1.313498	0.6919	0.5207
Q5	8.209678	2.041673	0.6258	0.5429
		Total:	0.7175	

In order to check the effect of removing Items 2, 6, and 7 on the measurement of the factor, the factor analysis was revisited. The desired items would be those items that have the highest loadings on the factor. The factor loadings are included below.

Table 6 - Factor Loadings for Coordination

Variables	Factor Loading
Q2	0.675629
Q3	0.729684
Q4	0.637104
Q5	0.646895
Q6	0.534292
Q7	0.111070

All items contribute relatively equally to the factor with the exception of Items 6 and 7. Since all other factor loadings are similar and the two lowest loading are to be removed, this factor should be relatively unaffected by the removal of Items 2, 6, and 7. Therefore, only Items 3, 4, and 5

will be retained in the coordination construct. The factor loadings, based on this reduced item set follow.

Table 7 - Factor Loadings for Reduced Coordination Set

Variables	Factor Loading
Q3	-0.828277
Q4	-0.620311
Q5	-0.636695

Item 3 contributes the most to the factor. Therefore, it may be possible, with further testing, to reduce the items in this construct to only Item 3. However, due to the internal reliability constraints set forth in this study, this reduction is not advised.

4.1.2 Communication

Items 8 through 12 on the original tool were designed to collect data concerning the team's communication process. The relevant results of the factor analysis follow. Only factors with eigenvalues greater than 1.00 are considered relevant.

Table 8 - Factor Analysis for Communication

Factor	Eigenvalue	Percent of Variance Explained
1	1.39	99.44

Only one factor emerged from the factor analysis, which is assumed to be a main communication factor. Since only one factor emerged from the factor analysis, no items can be immediately removed from the communication construct. Therefore, the Cronbach's alpha calculation will be used to reduce this item set.

The initial Cronbach's alpha analysis on the items in this construct, shown below, showed that the items were not reliable. The total Cronbach's alpha was 0.49.

Table 9 - Initial Cronbach's Alpha Analysis for Communication

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q8	8.000000	1.342862	0.3909	0.3692
Q9	8.322580	1.627370	0.4107	0.3040
Q10	7.758065	2.386337	0.6063	0.0652
Q11	7.741935	2.071948	0.2952	0.4353
Q12	7.838710	1.651562	0.4321	0.2668
		Total:	0.4880	

To reduce this item set, and increase the reliability to an acceptable level, items were removed one at a time and the Cronbach's alpha was recalculated. The item removed was the item that would effect the greatest increase in the overall alpha value. This procedure was repeated until the alpha value was considered acceptable. At the end of this procedure, only two items remained, resulting in a Cronbach's alpha of 0.73: Items 8 and 12. The final Cronbach's alpha analysis follows.

Table 10 - Cronbach's Alpha Analysis for Reduced Communication Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q8	8.00000	1.342862		0.5839
Q12	7.83871	1.651562		0.5839
		Total:	0.7275	

The factor analysis was revisited to check the effects of removing all but these two items. The factor loadings of the original item set follow.

Table 11 - Factor Loadings for Communication

Variables	Factor Loading
Q8	0.628402
Q9	0.197450
Q10	-0.038069
Q11	0.385510
Q12	0.645025

Items 8 and 12 contributed the most to the overall factor being measured. Therefore, the effect of item removal should not be overly large, and Items 8 and 12 will be retained for measuring the communication construct.

4.1.3 Conflict Resolution

Items 13 through 17 on the original tool were designed to collect data concerning the team's conflict resolution capabilities. The relevant results of the factor analysis follow. Only factors with eigenvalues greater than 1.00 are considered relevant.

Table 12 - Factor Analysis for Conflict Resolution

Factor	Eigenvalue	Percent of Variance Explained
1	2.24	99.52

The factor analysis yielded one main factor measured by the items within the conflict resolution construct. Once again, it is assumed that this single factor is a measurement of the entire conflict resolution construct. Since there is only one factor present, no determination can be made on how to reduce the item set. A Cronbach's alpha analysis will be used to reduce the set. This analysis follows.

Table 13 - Cronbach's Alpha Analysis for Conflict Resolution

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q13	8.322580	1.762768	0.6149	0.7843
Q14	7.709677	1.592885	0.8069	0.2161
Q15	8.145162	2.179237	0.7202	0.5307
Q16	8.693548	1.563820	0.6882	0.6113
Q17	8.290322	1.682963	0.7069	0.5466
		Total:	0.7575	

The overall alpha value for all five items in this construct demonstrates an acceptable level of reliability. Item 14 shows the least reliability, followed by Item 15. In the interest of reducing the number of items in this construct, items were removed one at a time and the Cronbach's alpha was recalculated. The item removed was the item that would effect the greatest increase in the overall alpha value. This procedure was repeated until the alpha value was maximized. The final Cronbach's alpha analysis follows.

Table 14 - Cronbach's Alpha Analysis on Reduced Conflict Resolution Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q13	8.322580	1.762768		0.6906
Q16	8.693548	1.563820		0.6906
		Total:	0.8135	

Removing all but Items 13 and 16 raised the alpha value to 0.8135. To investigate the effect of the item reduction, the factor analysis was revisited. The factor loadings for the items in this construct follow.

Table 15 - Factor Loadings for Conflict Resolution

Variables	Factor Loading
Q13	-0.912738
Q14	-0.251483
Q15	-0.623330
Q16	-0.760670
Q17	-0.616384

Items 13 and 16 were the two highest contributors to the factor being measured. It is assumed by this that the effects of removing the other items in this construct will be minimal. Therefore, only Items 13 and 16 will be retained in this construct.

4.1.4 Decision Making

Items 18 through 21 on the original tool were designed to collect data concerning the team's decision-making abilities. The relevant results of the factor analysis follow. Only factors with eigenvalues greater than 1.00 are considered relevant.

Table 16 - Factor Analysis for Decision Making

Factor	Eigenvalue	Percent of Variance Explained
1	1.54	99.94

A single factor was discovered through the factor analysis. This factor is assumed to be a measurement of the overall decision making construct. Since only one factor was discovered, a Cronbach's alpha analysis was used to reduce this item set. The Cronbach's alpha analysis follows.

Table 17 - Initial Cronbach's Alpha Analysis for Decision Making

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q18	8.258064	1.736168	0.7081	0.2526
Q19	6.290323	2.724432	0.6107	0.4680
Q20	7.354839	2.033303	0.4740	0.6396
Q21	8.241936	1.922136	0.5763	0.4952
		Total:	0.6682	

The overall Cronbach's alpha value was 0.67, which is not considered acceptable in this study. In the interest of increasing the internal reliability of the items in the construct and decreasing the number of items in this construct, items were removed one at a time and the Cronbach's alpha was recalculated. The item removed was the item that would cause the greatest increase in the overall alpha value. This procedure was repeated until the alpha value was maximized. The final Cronbach's alpha analysis follows.

Table 18 - Cronbach's Alpha Analysis on Reduced Decision Making Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q19	6.290323	2.724432		0.5907
Q20	7.354839	2.033303		0.5907
		Total:	0.7231	

The reduction procedure yielded two items: 19 and 20. These two items have a Cronbach's alpha value of 0.72, which is considered acceptable in this study. To investigate the effect of the item reduction, the factor analysis was revisited. The factor loadings for the items in this construct follow.

Table 19 - Factor Loadings for Decision Making

Variables	Factor Loading
Q18	-0.317732
Q19	-0.634439
Q20	-0.846682
Q21	-0.568921

Items 19 and 20 exhibited the largest contribution to the factor being measured. Therefore, the effects of removing Items 18 and 21 are considered minimal, and only Items 19 and 20 will be retained within this construct.

4.1.5 Problem Solving

Items 22, 23, and 24 were designed to collect data on the team's problem solving capabilities. In an initial test for item reliability, it was found that the items in this construct were not sufficiently reliable. The results of the Cronbach's alpha analysis follow.

Table 20 - Preliminary Cronbach's Alpha Analysis on Problem Solving

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q22	8.516129	1.468153	0.3349	0.5383
Q23	8.225806	2.485495	0.3592	0.5317
Q24	8.596774	1.193695	0.6478	0.2760
		Total:	0.5888	

This analysis indicated that the reliability of the items in this construct could not be increased to a sufficient level by removing an item. However, removing Item 24 increases the alpha value to a sufficient level to consider modifications to Items 22 and 23 to increase the

reliability. As a purely exploratory analysis, a factor analysis was performed on three items. This analysis was performed not to determine the number of factors present in the data for these two items, but to examine the factor loadings on a common factor. The results of this analysis follow.

Table 21 - Exploratory Factor Analysis on Problem Solving

Factor	Eigenvalue	Individual Percent of Variance Explained	Cumulative Percent of Variance Explained
1	1.20	99.98	99.98
2	0.01	0.92	100.90
3	-0.01	-0.90	100.00

Table 22 - Factor Loadings for Problem Solving

Variables	Factor1
Q22	-0.65
Q23	-0.70
Q24	-0.30

This exploratory analysis demonstrated that there was only one factor being measured. It also demonstrated that Item 24 had the least bearing on the factor, with a factor loading of -0.30. This coupled with the fact that it had the lowest reliability contribution suggests that attention should be paid to Items 22 and 23 to increase their reliability. Once this reliability is increased, this should be sufficient to collect data on the problem solving aspect of teamwork.

Due to the lack of reliability in the data, all three items were removed from the tool. Future revisions of this tool should attempt to develop other items to measure the problem solving aspect of teamwork, as this tool statistically failed to reliably measure it.

4.1.6 Boundary Spanning

Items 25 and 26 were designed to measure the boundary spanning aspect of teamwork. Initial reliability analysis of these items revealed, as it did in the problem solving section, that the items in this construct were not at all reliable. The results of the Cronbach's alpha analysis follow.

Table 23 - Preliminary Cronbach's Alpha Analysis for Boundary Spanning

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q25	8.112904	2.081179		-0.1352
Q26	8.112904	2.143269		-0.1352
		Total:	-0.3126	

These two items did not demonstrate sufficient reliability, as illustrated by a Cronbach's alpha of -0.31. Due to the lack of reliability in this data set, no further analysis will be performed on these items and they will be removed from the tool. Future revisions of this tool should attempt to develop items to measure the boundary spanning aspect of teamwork. The items developed in this tool focused on occurrence and effectiveness of boundary spanning with the lower commands. No items were constructed to measure boundary spanning with the adjacent and higher commands. The next revision of this tool should attempt to measure these, as well.

4.1.7 Joint Knowledge

Items 27 though 30 were designed to measure the team's joint knowledge. The relevant results of the factor analysis follow. Only factors with eigenvalues greater than 1.00 are considered relevant.

Table 24 - Factor Analysis for Joint Knowledge

Factor	Eigenvalue	Percent of Variance Explained
1	1.59	99.84

Since only one factor was discovered, a Cronbach's alpha analysis was used to reduce this item set. The Cronbach's alpha analysis follows.

Table 25 - Initial Cronbach's Alpha Analysis for Joint Knowledge

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q27	9.193548	0.9723844	0.2348	0.1517
Q28	7.629032	2.9153850	-0.1657	0.3189
Q29	8.290322	1.9534510	-0.4059	0.5764
Q30	7.451613	2.1929650	0.6384	-0.2523
		Total:	0.2662	

The overall reliability of the items within this construct is not acceptable. In the interest of increasing the internal reliability of the items in the construct and decreasing the number of items in this construct, items were removed one at a time and the Cronbach's alpha was recalculated. The item removed was the item that would cause the greatest increase in the overall alpha value. This procedure was repeated until the alpha value was maximized. The final Cronbach's alpha analysis follows.

Table 26 - Cronbach's Alpha Analysis on Reduced Joint Knowledge Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q28	7.629032	2.915385		0.736
Q29	8.290322	1.953451		0.736
		Total:	0.81	

The reduction procedure yielded two items: 28 and 29. These two items have a Cronbach's alpha value of 0.81, which is considered acceptable in this study. To investigate the effect of the item reduction, the factor analysis was revisited. The factor loadings for the items in this construct follow.

Table 27 - Factor Loadings for Decision Making

Variables	Factor Loading
Q27	0.190763
Q28	0.840188
Q29	0.878630
Q30	-0.271512

Items 28 and 29 exhibited the largest contribution to the factor being measured. Therefore, the effects of removing Items 18 and 21 are considered minimal, and only Items 28 and 29 will be retained within this construct.

4.1.8 Motivation

Items 31 though 34 were designed to measure the team's joint knowledge. A preliminary Cronbach's alpha analysis demonstrated that the items within this construct are not reliable. The results of this analysis follow.

Table 28 - Initial Cronbach's Alpha Analysis for Motivation

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q31	8.161290	1.883435	0.2855	0.3750
Q32	7.935484	2.260647	0.2339	0.4024
Q33	6.661290	1.983072	0.5322	0.1313
Q34	8.790322	1.026487	0.4521	0.2214
		Total:	0.4673	

An attempt was made to increase the alpha value by removing items that would cause the largest increase in the alpha value. The highest alpha that could be obtained was 0.54, as shown in the following table.

Table 29 - Final Cronbach's Alpha Analysis for Motivation

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	**Mean**	**Standard Deviation**	**Coef Alpha**	**Corr Total**
Q31	8.161290	1.883435		0.3721
Q32	7.935484	2.260647		0.3721
		Total:	0.5359	

Since the data for this construct are not reliable, no further analysis was performed on these data, and all items within this construct were removed. Future revisions of this tool should develop other items to measure the team's motivation.

4.1.9 Personalities

Items 35 through 39 were designed to measure the team member's personalities, where they apply to teamwork. The relevant results of the factor analysis are shown below. For a factor to be considered relevant, its eigenvalue had to be greater than or equal to 1.00.

Table 30 - Factor Analysis for Personalities

Factor	Eigenvalue	Individual Percent of Variance Explained	Cumulative Percent of Variance Explained
1	2.112629	66.83	66.83
2	1.058294	33.48	100.31

Table 31 - Factor Structure for Personalities

Factor 1	Factor 2
Q35	Q38
Q37	Q39
Q39	

In reviewing the items within each factor, it was apparent that Factor 2 did not measure personality aspects of teamwork. Item 38 measures enthusiasm, which might belong more in the motivation construct than it does in the personality construct. Item 39 measured the organization habits of the team members and might belong the coordination construct. Since Factor 2 does not appear to be relevant to the personality construct, Item 38 was removed. Item 39 was retained for further analysis, since it also appears in Factor 1. Item 36 did not appear to be a major contributor to either factor, so it was removed as well.

The Cronbach's analysis was performed on the remaining items. The results of that analysis are shown below.

Table 32 - Cronbach's Alpha Analysis on Reduced Personalities Set

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q35	8.467742	2.427423	0.7331	0.7952
Q37	7.870968	2.760873	0.7400	0.7964
Q39	7.483871	2.038498	0.8859	0.6332
		Total:	0.8556	

The overall Cronbach's alpha for the remaining items is 0.86, which is considered acceptable. However, since all three of these items are measuring within the same factor and removing Item 39 will increase the alpha level to 0.89, Item 39 will be removed as well. Items 35 and 37 most directly measure personality aspects of teamwork, and they are, therefore, retained in this construct.

4.1.10 Shared Thought Processes

Items 40 through 42 were designed to measure the prevalence of the team's shared thought processes. A preliminary Cronbach's alpha analysis demonstrated that the items within this construct were not reliable and could not be made more reliable by removing an item. The details of this analysis follow.

Table 33 - Preliminary Cronbach's Alpha Analysis for Shared Thought Processes

	--------- Item Values ----------		---- If This Item is Omitted ----	
Item	Mean	Standard Deviation	Coef Alpha	Corr Total
Q40	6.838710	2.741266	0.3318	0.4185
Q41	8.838710	1.332981	0.4697	0.4493
Q42	5.854839	3.415715	0.5270	0.3528
		Total:	0.5428	

Since the highest alpha value that can be achieved is 0.54, which is well below the minimum value for acceptance, all items in this construct were removed and no further analysis

was performed on the data within this construct. Future revisions of this tool should attempt to develop other items to measure this construct.

4.2 Inter-Rater Reliability

The following sections contain the inter-rater reliability analysis performed on each item remaining in the tool. It is organized by item. Within each item, the Kendall's Concordance (W) analysis is performed (Siegel and Castellan Jr., 1988), followed by an ANOVA (Montgomery, 2001) to determine the extent of inter-rater reliability. None of the items' responses satisfied the assumption of normality, but ANOVA is robust to this violation according to Montgomery (2001). Each of these analyses was performed with an alpha level of 0.1. Details of these analyses can be found in Appendix D and Appendix F.

4.2.1 Item 1 – Level of Activity

The Kendall's Concordance for Item 1 was calculated as being 0.4869. The subsequent chi-squared test for independence demonstrated that there was only a 0.068 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 34 - Item 1 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.1)
Time Period	24	132.67	5.53	2.53	0.003035*	0.9974
Observer	2	41.31	20.65	9.47	0.000342*	
Residual	48	104.69	2.18			
Total (Adj)	74	278.67				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability was not sufficient. This discrepancy could potentially be because of the differences in the inputs to each of the two tests. The W analysis suggests that the observers followed similar patterns in their ratings, but the ANOVA suggests that the magnitude of those individual ratings differed significantly. This would suggest that the training program be modified to include an opportunity to normalize the observer ratings or that the ratings should be normalized after collection. The first is preferred, as normalizing the ratings post-hoc would significantly diminish the ability to compare ratings of different teams.

4.2.2 Item 3 – Anticipation of Events and Requests

The Kendall's Concordance for Item 3 was calculated as being 0.6034. The subsequent chi-squared test for independence demonstrated that there was only a 0.009 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 35 - Item 3 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.1)
Time Period	24	202.85	8.45	2.81	0.001150*	0.9990
Observer	2	22.91	11.45	3.81	0.029212*	
Residual	48	144.43	3.01			
Total (Adj)	74	370.19				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is not sufficient.

4.2.3 Item 4 – Timely Presentation of Material

The Kendall's Concordance for Item 4 was calculated as being 0.4734. The subsequent chi-square test for independence demonstrated that there was only a 0.083 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 36 - Item 4 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.1)
Time Period	24	37.39	1.56	1.2	0.2875	0.8603
Observer	2	29.12	14.56	11.23	0.000100*	
Residual	48	62.21	1.30			
Total (Adj)	74	128.72				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is not sufficient.

4.2.4 Item 5 – Level of Confusion

The Kendall's Concordance for Item 5 was calculated as being 0.4128. The subsequent chi-square test for independence demonstrated that there was a 0.194 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 37 - Item 5 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.1)
Time Period	24	162.35	6.76	1.81	0.040095*	0.9727
Observer	2	36.59	18.29	4.89	0.011631*	
Residual	48	179.41	3.74			
Total (Adj)	74	378.35				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficient.

4.2.5 Item 8 – Orderly Communication

The Kendall's Concordance for Item 8 was calculated as being 0.3473. The subsequent chi-square test for independence demonstrated that there was a 0.406 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 38 - Item 8 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.1)
Time Period	24	57.15	2.38	1.16	0.3205	0.8468
Observer	2	9.71	4.85	2.37	0.1043	
Residual	48	98.29	2.05			
Total (Adj)	74	165.15				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is sufficiently high.

4.2.6 Item 12 – Participation in Discussions

The Kendall's Concordance for Item 12 was calculated as being 0.4236. The subsequent chi-square test for independence demonstrated that there was a 0.169 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 39 - Item 12 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	58.59	2.44	1.41	0.1550	0.9163
Observer	2	38.75	19.37	11.17	0.000104*	
Residual	48	83.25	1.73			
Total (Adj)	74	180.59				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.7 Item 13 – Type of Criticism

The Kendall's Concordance for Item 13 was calculated as being 0.3965. The subsequent chi-square test for independence demonstrated that there was a 0.238 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 40 - Item 13 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	96.05	4.00	1.33	0.1944	0.8992
Observer	2	41.39	20.69	6.9	0.002322*	
Residual	48	143.95	3.00			
Total (Adj)	74	281.39				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.8 Item 16 – Amount of Criticism

The Kendall's Concordance for Item 16 was calculated as being 0.4382. The subsequent chi-square test for independence demonstrated that there was a 0.139 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 41 - Item 16 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	50.61	2.11	1.11	0.3705	0.8265
Observer	2	40.03	20.01	10.52	0.000163*	
Residual	48	91.31	1.90			
Total (Adj)	74	181.95				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.9 Item 19 – Amount of Brainstorming

The Kendall's Concordance for Item 19 was calculated as being 0.4217. The subsequent chi-square test for independence demonstrated that there was a 0.173 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 42 - Item 19 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	88.75	3.70	1.14	0.3369	0.8401
Observer	2	302.91	151.45	46.87	0.000000*	
Residual	48	155.09	3.23			
Total (Adj)	74	546.75				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.10 Item 20 – Collaboration in Decision Making

The Kendall's Concordance for Item 20 was calculated as being 0.4026. The subsequent chi-square test for independence demonstrated that there was a 0.221 chance that the agreement among the observers was by random chance. This level of chi-square does not satisfy the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 43 - Item 20 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	80.35	3.35	1.24	0.2548	0.8738
Observer	2	104.83	52.41	19.48	0.000001*	
Residual	48	129.17	2.69			
Total (Adj)	74	314.35				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.11 Item 28 – Surprising Behavior

The Kendall's Concordance for Item 28 was calculated as being 0.4851. The subsequent chi-square test for independence demonstrated that there was only a 0.069 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 44 - Item 28 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	191.28	7.97	1.74	0.051082*	0.9666
Observer	2	188.16	94.08	20.54	0.000000*	
Residual	48	219.84	4.58			
Total (Adj)	74	599.28				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.12 Item 29 – Clear Goals and Objectives

The Kendall's Concordance for Item 29 was calculated as being 0.5423. The subsequent chi-square test for independence demonstrated that there was only a 0.027 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 45 - Item 29 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	109.65	4.57	1.37	0.1741	0.9080
Observer	2	57.31	28.65	8.59	0.000645*	
Residual	48	160.03	3.33			
Total (Adj)	74	326.99				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.13 Item 35 – Distracting Behavior

The Kendall's Concordance for Item 35 was calculated as being 0.3982. The subsequent chi-square test for independence demonstrated that there was a 0.233 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item does not have an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 46 - Item 35 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	113.28	4.72	1.23	0.2648	0.8697
Observer	2	148.59	74.29	19.37	0.000001*	
Residual	48	184.08	3.84			
Total (Adj)	74	445.95				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, as does the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.14 Item 37 – Team Member Demeanor

The Kendall's Concordance for Item 37 was calculated as being 0.5061. The subsequent chi-square test for independence demonstrated that there was only a 0.050 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follows.

Table 47 - Item 37 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	198.00	8.25	1.72	0.055236*	0.9643
Observer	2	126.11	63.05	13.13	0.000028*	
Residual	48	230.56	4.80			
Total (Adj)	74	554.67				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.2.15 Item 43 – Overall Performance

The Kendall's Concordance for Item 43 was calculated as being 0.5442. The subsequent chi-square test for independence demonstrated that there was only a 0.026 chance that the agreement among the observers was by random chance. This level of chi-square satisfies the alpha requirement of 0.1. Therefore, this analysis suggests that this item has an acceptable level of inter-rater reliability.

The ANOVA results summary for this item follow.

Table 48 - Item 43 ANOVA Results Summary

Source	DF	SS	MS	F-Ratio	Prob. Level	Power (Alpha=0.10)
Time Period	24	70.99	2.96	1.41	0.1543	0.9166
Observer	2	19.23	9.61	4.58	0.015134*	
Residual	48	100.77	2.10			
Total (Adj)	74	190.99				
Total	75					

* Term significant at alpha = 0.10

The ANOVA for this item suggests, contrary to the Kendall's Concordance, that inter-rater reliability is not sufficiently high.

4.3 Construct Validity

Construct validity was assessed by comparing the observers' ratings with the self-assessments of the command team members, who are considered command and control experts. A paired t-test was conducted to facilitate this assessment (Montgomery and Runger, 1999). The observers' ratings were averaged for each day, for each construct and paired with the average end of the day ratings made by the command team members. This was repeated for each of the four exercise days, resulting in a paired t-test for each construct with N = 4. The hypothesis test is as follows:

$$H_0 : \mu_1 = \mu_2 \qquad T_0 = \frac{\overline{D}}{S_D / \sqrt{n}}$$
$$H_1 : \mu_1 \neq \mu_2$$
$$\alpha = 0.40$$

Figure 9 - Construct Validity Hypothesis Testing Structure

where D is the difference between the data points within each pair and S_D is the sample standard deviation of the differences. H_0 is rejected if $|T_0| > t_{\alpha/2, n-1}$.

Because of the small sample size, the power of the paired t-test is typically low ($\beta > 0.9$, in some cases). To combat this problem, a balance was found between alpha and beta. This tradeoff resulted in a fairly unorthodox alpha of 0.4. The tests run in this section are at the 0.4 alpha level. Power is also reported for each construct to aid in analysis of the results. For analysis details, please refer to Appendix G.

4.3.1 Level of Activity

There is only one item in the tool, Item 1, which concerns the level of activity. This item asks the user to directly rate the amount of activity experienced by the team during the observation period. This item does not represent a team process construct, but it adds a piece of information that could be found useful in certain analyses and explorations.

The results of the paired t-test follow.

Table 49 - Paired t-test Summary for Level of Activity

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	7.890	0.572	0.286
Team	4	7.538	0.799	0.400
Difference	4	0.353	0.763	0.381
Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	0.924	0.423	Accept Ho	0.553

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the level of activity is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.55. According to this analysis, the measurement of the level of activity appears to be valid.

4.3.2 Coordination

After the item reduction phase of the data analysis, there were three items remaining in the Coordination team process construct. These were Items 3, 4, and 5. Since there were multiple items in this construct, the scores for each of the items were averaged together to form a single construct score. It is this construct score that is tested in the paired t-test for construct validity. The results of the test follow.

Table 50 - Paired t-test Summary for Coordination

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	7.938	0.467	0.234
Team	4	7.978	0.274	0.137
Difference	4	-0.040	0.519	0.259

Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	-0.154	0.887	Accept Ho	0.405

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the coordination is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.41. According to this analysis, the measurement of this construct appears valid.

4.3.3 Communication

Items 8 and 12 remain in the tool to measure the communication team process construct. The item averages were used to construct a single construct score for communication. This overall score was used in the test procedure. The test results follow.

Table 51 - Paired t-test Summary for Communication

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	7.815	0.423	0.212
Team	4	7.600	0.362	0.181
Difference	4	0.215	0.533	0.267

Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	0.806	0.479	Accept Ho	0.521

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the communication is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.52. According to this analysis, the measurement of this construct appears valid.

4.3.4 Conflict Resolution

Items 13 and 16 remain in the tool to measure the conflict resolution team process construct. The item averages were used to construct a single construct score for conflict resolution. This overall score was used in the test procedure. The test results follow.

Table 52 - Paired t-test Summary for Conflict Resolution

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	8.468	0.563	0.281
Team	4	8.163	0.536	0.268
Difference	4	0.305	1.094	0.547
Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	0.558	0.616	Accept Ho	0.851

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the conflict resolution is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.85. According to this analysis, the measurement of this construct appears valid.

4.3.5 Decision Making

Items 19 and 20 remain in the tool to measure the decision making team process construct. The item averages were used to construct a single construct score for decision making. This overall score was used in the test procedure. The test results follow.

Table 53 - Paired t-test Summary for Decision Making

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	6.778	0.462	0.231
Team	4	6.623	0.597	0.298
Difference	4	0.155	0.466	0.233
Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	0.666	0.553	Accept Ho	0.485

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the decision making is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.49. According to this analysis, the measurement of this construct appears valid.

4.3.6 Joint Knowledge

Items 28 and 29 remain in the tool to measure the joint knowledge team process construct. The item averages were used to construct a single construct score for joint knowledge. This overall score was used in the test procedure. The test results follow.

Table 54 - Paired t-test Summary for Joint Knowledge

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	7.925	0.718	0.359
Team	4	7.958	0.115	0.058
Difference	4	-0.033	0.815	0.407
Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	-0.080	0.941	Accept Ho	0.401

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the joint knowledge is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.40. According to this analysis, the measurement of this construct appears valid.

4.3.7 Personalities

Items 35 and 37 remain in the tool to measure the personalities team process construct. The item averages were used to construct a single construct score for personalities. This overall score was used in the test procedure. The test results follow.

Table 55 - Paired t-test Summary for Personalities

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	8.180	0.477	0.238
Team	4	7.848	0.952	0.476
Difference	4	0.333	1.388	0.694

Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	0.479	0.665	Accept Ho	0.446

There is a lack of significant disagreement between the ratings of the observer team and the command team, where the personalities is concerned. This is concluded by the results of the paired t-test at an alpha level of 0.40 and a corresponding power of 0.45. According to this analysis, the measurement of this construct appears valid.

4.3.8 Overall Performance

The final item in the tool, Item 43, asked the rater for an overall rating of the team's effectiveness for the observation period. This item does not represent a team process construct, but it adds a piece of information that could be found useful in certain analyses and explorations. Since this is only a single item in the tool, there were no items to average across to form a construct score. Therefore, the tested pairs consisted of the average of the end of the day ratings made by the command team members and the average ratings made by the observers throughout the day on item 43. The test results follow.

Table 56 - Paired t-test Summary for Overall Performance

Descriptive Statistics				
Variable	***Count***	***Mean***	***Std. Dev.***	***Std. Error***
Observer	4	7.995	0.304	0.152
Team	4	8.655	0.137	0.068
Difference	4	-0.660	0.416	0.208

Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .4)**	**Power**
$\mu 1 \neq \mu 2$	-3.174	0.050	Reject Ho	0.983

At an alpha level of 0.40, the ratings of overall performance were not consistent between the observer team and the command team. The results of the test at the 0.40 level, however, suggested that the null hypothesis would be accepted at the 0.05 level, which is a more traditional value of alpha than the one used in this part of the analysis. As an exploratory measure, the test was rerun with an alpha level of 0.05. The results of that test follow.

Table 57 - Paired t-test for Overall Performance at Alpha = 0.05

Paired t-test Results				
Alternate Hypothesis	**T-Value**	**Prob. Level**	**Decision (alpha = .05)**	**Power**
$\mu 1 \neq \mu 2$	-3.174	0.050	Accept Ho	0.575305

This test demonstrated that the construct was valid at the 0.05 level. Recall that the reason the inflated alpha was used was because the statistical power of the paired t-test was generally very low with an alpha level of 0.05. The statistical power of this test, however, is much higher than the previous tests. In this test, the power at the 0.05 level is higher than that of many of the other constructs' tests at the 0.40 level. This would seem to suggest that the measure is indeed valid. However, since the test demonstrated significant difference at the 0.40 level, the overall performance construct has not passed the requirements for construct validity.

4.3.9 All Constructs

As an exploratory analysis, a principal components analysis was performed on the construct scores from the study to determine the underlying structures according to Jolliffe (2002). Construct scores used in the analysis were calculated as an average of the item scores within that construct, averaged again over the three observers for that hour. Therefore, the data used in the analysis represented the construct score for each hour of the exercise. The analysis was performed with no factor rotation and using the variance-covariance matrix. The minimum loading to be considered a part of the component is 0.4. This component loading value was chosen after looking at the results of the analysis. Loadings much below this value exhibited evidence of item bleed-over, where most items appeared in each of the components. Using a loading factor of 0.4 made the true composition of the components much clearer. The results of the analysis follow. Detailed results of the analysis can be found in Appendix I.

Table 58 - Principal Components Analysis Results for all Constructs

Component	Eigenvalue	Individual Percent of Variance Explained	Cumulative Percent of Variance Explained
1	5.538857	57.06	57.06
2	1.757412	18.10	75.16
3	0.717009	7.39	82.55
4	0.665949	6.86	89.41
5	0.425111	4.38	93.79
6	0.300687	3.10	96.89
7	0.232006	2.39	99.28
8	0.070330	0.72	100.00

Table 59 - Component Loadings for all Constructs

	Components	
Variables	1	2
Activity	-0.3238	-0.9345
Coordination	-0.7469	0.0161
Communication	-0.8360	0.1458
Conflict_Resolution	-0.7701	-0.1086
Decision_Making	-0.5354	-0.1816
Joint_Knowledge	-0.9039	0.0927
Personalities	-0.9017	0.2821
Overall	-0.8208	-0.0372

Table 60 - Component Structure for all Constructs

Component 1	Component 2
Joint_Knowledge	Activity
Personalities	
Communication	
Overall	
Conflict_Resolution	
Coordination	
Decision_Making	

Generally, all of the constructs appear to be measuring the same factor, with the exception of the general activity item, which is the principal component of factor 2. Factor one is assumed to be an overall measure of team performance, since all of the team process constructs that survived the pervious analyses are included in this factor. As such, the remaining constructs appear to confirm the theory behind the Team Effectiveness Model that all the team process attributes are components of overall team performance.

Chapter 5 - Discussion

This study was performed to initiate the development of a methodology for evaluating military teams based on the team process constructs illustrated in the Team Effectiveness Model proposed by Tannenbaum et al. (1992b). In this first stage of development, the vehicle of the methodology was designed, developed, tested, and evaluated. The result was the Team Process Observation Tool. This tool was developed to measure the team's performance on each of the team processes and the team's cognitive process. Due to human error on the part of the researcher, one process, joint task ability, was inadvertently omitted from the tool. This omission was not realized until the field study was completed, and therefore, this cognitive team process did not appear in the study.

The tool developed was tested in a military field study environment. This testing environment removed a great deal of experimental control from the test, but it also added to the external validity of the tool, as the testing environment was performed in an environment similar to that of its intended use. The results of the analysis procedures are discussed in the following sections.

5.1 Command and Control Team Evaluation

The proposed methodology represents a foundation for the development of an evaluation tool for military C^2 teams that is independent of battle outcomes. An assumption is made here that a team that performs well in terms of team process will produce superior output. Therefore, measures of team process performance will be leading indicators for outcome measures. Outcome measures should not be discarded as measures of effectiveness for combat systems; however, they should not be the primary measures for C^2 team performance, because C^2 teams do not fight the battles that determine those outcomes.

To the extent possible, this methodology was designed to be used in naturalistic environments. In C^2 research, however, performing observations in the natural environment, active battle, is not considered a feasible option. Therefore, in order to preserve the tenants of

action research, the next best environment was chosen. The environment chosen, the battle simulation, while not precisely duplicating the conditions of the battle environment, is a safe environment to conduct team observations while maintaining a battle-like environment.

The proposed methodology has also been designed to be a more efficient evaluation methodology for C^2, as compared with some of the other evaluation methods employed by the U.S. Army, such as conversation and protocol analyses. A more efficient evaluation process makes it possible to better use the evaluations to identify and correct team-related problems. The quantitative nature of this methodology also allows for benchmarking teams, which can lead to further improvements in C^2 team performance. This quantification of team performance also allows for this methodology to be applied in current research programs already established in the U.S. Army for C^2 systems.

While this methodology is quantitative, the specific method developed in this research is not truly objective, as it relies on subjective ratings from the observers. Developing the methodology into a more objective, quantitative method for evaluating team performance could have benefits. First, inter-rater reliability would, theoretically, become less of an issue as the objectivity of the method increases. Second, because inter-rater reliability should be increased, the method could more easily be used to compare teams on the team process performance level. Finally, this research endeavor is part of a larger research program that is building a discreet-event simulation of command and control. The methodology presented here is ultimately meant to provide input to this simulation model.

5.2 Observational Methodology

The observational nature of this proposed methodology was built upon many separate observational methodologies, as discussed in Chapter 2. The environment of the battle simulation requires an observer to process inputs from multiple sources, because no one source is sufficient to form a conclusion. These sources are, primarily, the physical locations and reactions of the team members and the content and tone of communication, both within the group

and with outside groups. It is for this reason that multiple observation methods had to be studied to develop an observational methodology to collect the team process data.

The development of the data collection tool used in this study had to serve two separate functions. First, the tool had to be used by the observers in collecting the data produced by observing the command team. Second, for construct validity purposes, it had to be used by the command team to collect self-evaluation data. As such, the tool had to be designed to accommodate both observational and non-observational methods. The results indicated that the tool served its function in each case.

There is an inherent problem with this form of validation, though it was the best choice for this field study. The observers were instructed to give an average rating for the hour. The command team was given instructions to give a rating that represented the average for the entire day. Because of the length of time that an exercise day covered, it cannot be expected that the end of the day self-assessments can represent a true average, due to recency and primacy effects in long-term memory (Lindsey and Norman, 1977) or other factors. More likely, the end of the day ratings were more aligned with the critical incidents that were prominent in the memories of the command team members.

Another issue that became apparent during the battle simulation was the changing nature of what was being observed. The method developed in this study was intended for use with observing physical artifacts produced by the command team. During the first part of the week of this exercise, these physical artifacts, such as walking across the command area to conduct a brief meeting with someone or walking up to a map to assess the current battle situation, were apparent, as the command team was physically active in the TOC. However, as the week progressed, the command team became less physically active. By the end of the week, the command team rarely left their chairs. This change in the nature of the interactions may have been a result of several causes. First, the team may have become increasingly more familiar with working together in a battle situation. As was stated earlier, these individuals typically train active-duty military personnel in the usage of the battle simulations; they have never before participated as an active component in a simulation. Second, the pace of the battle increased

considerably as the week progressed. The lack of physical activity may have been a result of the increased cognitive workload. Finally, it is possible that the participants were growing weary of participating in the exercise, and the motivation began to shift more toward finishing the week, as opposed to doing the best they could during the simulation. Regardless of the reasons for the change, the changes caused the observations to shift from observing physical signs to gathering information from the conversations within the TOC and between the TOC and other entities. All of the observers, without prompting from the researcher, began intently watching the intelligence maps while listening to these conversations in an attempt to obtain the information necessary to complete the observation tool.

As a result of the changes in the observation focus, it is difficult to tell whether the physical observation or the use of conversation produced the best results. The training program included physical artifacts as the primary examples of what to look for in the observation sessions. After observing the changes in this simulation, it may be advisable to add more of the conversational artifacts to the training program. Additionally, some time should be spent training the observers on how to read the intelligence maps. These maps provided many clues that augmented what was heard during the conversations in the latter part of the week.

5.3 Item Reduction

The first stage of the analysis process was to reduce the item set. The tool tested contained several items per team process construct, with 43 items overall. During the testing procedure, it became apparent that the observers were spending a considerable amount of time completing the tool each hour, sometimes taking more than 15 minutes to complete the tool. That means the observers sometimes spent 25% of their time completing the tool, and not observing the team. In order to reduce the reliance on human memory and the averaging effects of recording based on an hour's worth of observation, the time between recordings could be reduced. If this is done, the percentage of time the observer is recording information, and not actively observing, is increased substantially. Reducing the number of items in the tool should decrease the amount of time spent completing the tool. Though the tool takes the form of a tool with items and a rating scale, it is an observational data collection tool, not a questionnaire.

Therefore, the guidelines for questionnaires do not truly apply here. For example, you might not reduce the number of items in a questionnaire if your criterion for correlation is met. In this tool, since completion time is important for the efficiency of this tool, this rule is balanced with the need for a shorter instrument. For this reason, the number of items in each construct were limited to three, wherever possible.

Figure 10, below, represents the observers' performance times for completion of the observation tool over the four days of the exercise. The trend line provided demonstrates the decreasing nature of the completion time, as it is associated with a natural learning curve. From this trend, it is estimated that, for all 43 items, the natural completion time would be between 5 and 6 minutes. If it is assumed that each item requires an equal amount of time to complete, it can be estimated that an observer that is experienced with the tool can complete an item in less than 10 seconds. Once a completion time is calculated, the tradeoff decision can be made to determine the frequency of the recordings.

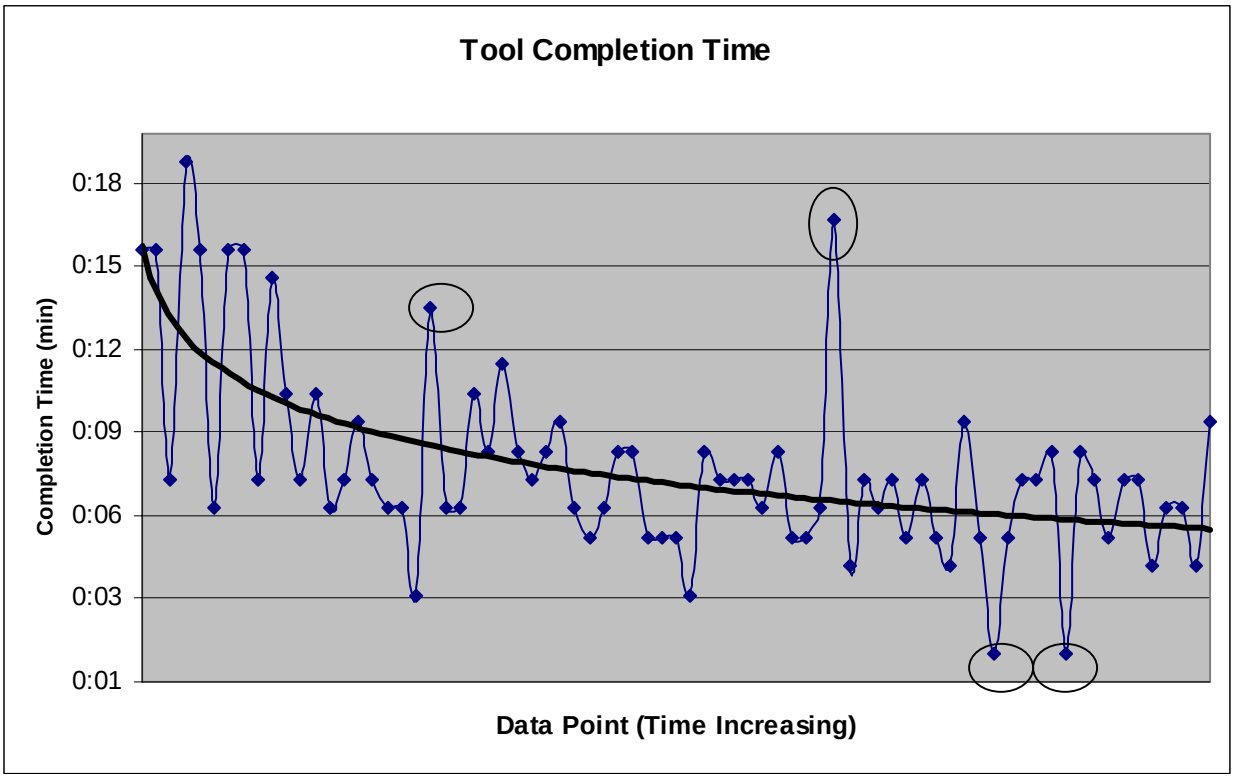


Figure 10 - Tool Completion Times

Data outliers, like the ones circled in the chart above, typically occurred in one of two ways. First, the low outliers were attributed to the researcher. This makes intuitive sense, since the researcher is the person most familiar with the tool. Therefore, it should take him less time to complete the tool. High outliers typically occurred around lunchtime, when the observers were trying to complete the tool and eat lunch at the same time. Another explanation for high outliers is the activity in the TOC at the time of recording the data. When activity was high at the time of recording data, the observers had to devote more attention to the activity, since that activity was part of their next hour's data.

Reducing the time necessary to complete the tool was not the only reason for performing an item reduction. Analysis needed to perform the reduction also illuminated deficiencies in the structure of the tool. When deficiencies are found, they can be remedied or they can be discarded. From a pure reductionism point of view, discarding these "defective" items is preferable. To minimize the time to complete the tool, the discarding of items was considered preferable to reconstructing the items. However, when there were few items in a particular construct, discarding items was not an option.

Two separate analyses were employed in the item reduction phase. One analysis was a factor analysis. The factor analysis was performed to ensure that all items were measuring the intended construct. The second analysis, Cronbach's Alpha, was used to ensure that items were reliable within each construct.

Several constructs did not survive the item reduction analyses. Since the items were designed to collect data on the team processes, a failure of an item is attributed to a mismatch between the construction of the item and the applicability in the TOC. For example, an item could have been constructed to collect a particular aspect of a team process, but when applied in the TOC environment or the specific field study, the item did not apply or could not be observed. The wording of the item, in conjunction with the varied experience of the observers may have caused a disparity in the ratings among the different observers. Again, this is an issue with the design of the tool and not the underlying model. What follows is a discussion of potential causes and remedies for each of these constructs.

The problem solving construct was the first construct to be found unreliable. The analysis of the items in this construct showed that items 22 and 23 were almost reliable enough to include in the tool. Therefore, instead of developing completely new items for this construct, these two should be modified to increase reliability.

There are a number of reasons that could have contributed to the unreliable nature of these items. First, item 22 asked the observer to infer the level of understanding of objectives. Outward signs of understanding can often be difficult to observe. Second, among the observers, there was a great deal of disparity in the amount of knowledge of TOC operations and military command and control in general. Better training for the observers or using observers experienced with TOC operations or command and control observation could diminish both of these potential causes.

Boundary spanning was the next construct that was found to be unreliable. This construct only had two items contained within it, so there was no opportunity to remove items to increase the reliability rating. These items did not cover the range of types of boundary spanning occurring in the TOC. Because of this, this construct should be supported with a more complete set of items and retested. Until a reliable set of items can be developed for this construct, it should not be included in the tool.

The motivation construct showed similar results. It was also found to be unreliable across observers. Since motivation is an intrinsic property, and therefore not observable, the items in this construct were designed to measure factors that influence motivation and the typical outcomes of working with a low motivation. Further research should be performed on measuring work motivation using observation. Once complete, this construct should be reconstructed and tested.

The shared thought processes construct was the last construct to not meet the reliability requirements. Again, this construct represented an attempt to measure something by observation that is intrinsic in nature. This construct is fairly abstract, as well. As with motivation, this team

process should be researched further, and the items composing the construct should be completely reconstructed.

The joint task ability construct was accidentally omitted from the tool during the development stage. The next test should include this construct, along with a new attempt to develop construct measurements for shared thought processes, motivation, boundary spanning, and problem solving. Since the other constructs have met or exceeded the reliability requirements, they would not need to be included in further tests of reliability.

5.4 Inter-Rater Reliability

The inter-rater reliability analysis was complicated by the way the study was performed. Most inter-rater reliability calculations are designed to be used with two observers and nominal or ordinal data. The data collected with this tool was assumed to be interval in nature, because the rating scale is Likert-type, but the best currently established method that could be used in this study was Kendall's Concordance, which requires ordinal data but can be used with more than two observers. According to Siegel and Castellan (1988), a great deal of information is lost when interval data are analyzed as ordinal data, so an ANOVA was performed on the interval data, as well. However, an ANOVA might be better suited to gauge inter-rater reliability once the tool has been further refined. In its present form, the data may not be truly interval in nature. The eventual goal of this observation tool is to provide data that could satisfy the ANOVA criteria for acceptable inter-rater reliability, because the numerical scores will be used in the U.S. Army's command and control simulation model. Since it is the score that is needed for that research, the tool needs to be consistent across raters with rating magnitudes, as well as movement correlations.

Kendall's Concordance has the ability to detect the changes in ratings for a particular item across raters. Therefore, meeting the inter-rater reliability requirement in this measure suggests that the item was designed well enough that the relative changes in ratings across time were uniform across all raters. It does not, however, make any claims about the equivalence of

the magnitudes of the ratings made by the raters for a given time period. The ANOVA was performed in an attempt to gauge the differences in the magnitudes.

Approximately half of the items met the requirements for sufficient inter-rater reliability, based on the Kendall's Concordance calculation. However, only one item was found to have a sufficient level of inter-rater reliability, according to the ANOVA, and this item was found to be insufficient with the Kendall's Concordance calculation.

The overall low inter-rater reliability rating is most likely due to the differences existing among the three observers. This group of observers was a very heterogeneous group. A lack of sufficient observer training prior to the study and the researcher's lack of familiarity with Army command and control environments prevented measures to overcome these differences. The items could be reworded to include specific examples of observable behavior in the TOC to raise inter-rater reliability. However, the biggest benefit to inter-rater reliability would probably come from the normalization effects of training or in the careful selection of observers. However, it is desired that the tool be made such that anyone can use it; therefore, correction of this problem with selection of observers is not considered optimal. If possible, audio/video recordings should be used in future training sessions. This would orient the observers to the C^2 environment as well as the tool.

5.5 Construct Validity

The construct validity portion of this study was designed to compare the perceptions of the observers to the perceptions of the C^2 experts, who were the command team members in this case. In order for a construct to be valid, the observers' ratings of team process performance had to be comparable to the command team's self-assessments. Overall, the daily average of the observers' ratings were not significantly different from the ratings made by the command team, which means that the observers' assessments of the team's performance concurred with the self-assessments of the command team. The only exception was the overall performance rating. In an effort to determine what the cause might be for the discrepancy in the overall performance rating, the data points were graphed, displayed in Figure 11, to see if any patterns were present.

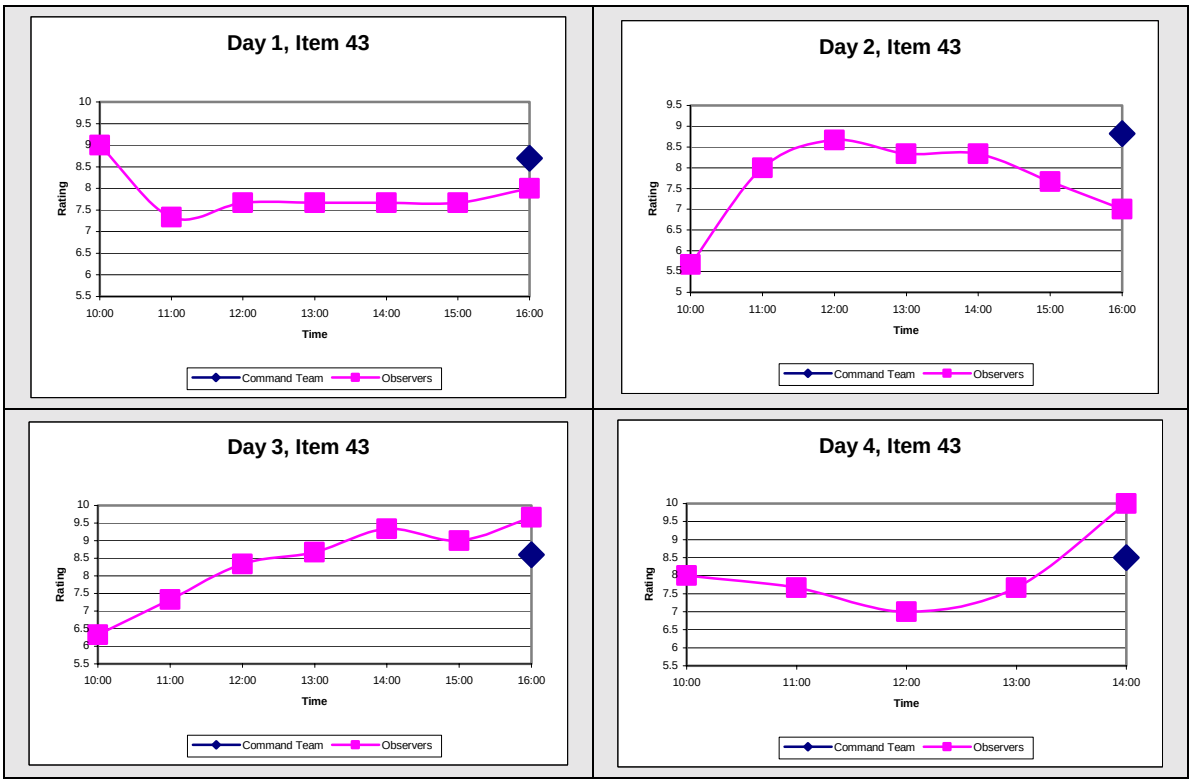


Figure 11 - Results for Overall Performance

Based on the graphs in Figure 11, there are no discernable patterns between the ratings made by the observers and the team members. The battle simulation upon which this study is based did not perform a full After-Action Review (AAR) at the end of each day, due to time constraints. Therefore, it was not possible to obtain an overall sense of the level of performance except through the ratings supplied by the observation tool.

The type of test performed on the overall performance item will not be possible in an actual battle simulation exercise, because the command team will not be completing the tool However, the AARs in these exercises will be more complete, so the team's perception of overall performance should be easily gathered from the AAR. The overall performance rating made by the observer(s) should be compared to the general consensus from the AAR.

Based on the results of this portion of the study, the average ratings made by the observers were equivalent to the self-assessments made by the command team members. This sugges that the average of the observers' ratings represented how the C^2 experts would have

evaluated the team, had this been an active-duty training exercise. Therefore, each of the team process constructs remaining in the tool has passed the construct validity test and should be considered valid and reliable measures of those constructs.

In future exercises, where active-duty personnel are being trained, the construct validity check could be performed in two ways. First, the team's self-assessment could be gleaned from the After Action Review (AAR). This could be accomplished using ethnographic methods, and it would rely on the observer's interpretation of the discussions taking place in the AAR. Second, the evaluator or senior officer engaged in the simulation could be asked to complete the tool, just as the command team did in this study. Most likely, those completing the tool in this manner would do so without training on the usage of the tool. The final decision on which method to employ in future validation of this tool would require a trade-off analysis on the benefits of each method. Some of the trade-offs that must be considered are found in the table below.

Table 61 - Trade-offs in Selecting Type of Data Collection for Overall Performance Rating

Observer, using ethnographic methods	Evaluator or Officer self-assessments
Familiarity with the tool	Lack of familiarity with how to rate using the tool
Third-party observer	Actively involved in the training exercise
More interested in collecting good data	More interested in calling an end to the duty day
Separation from participation in exercise may make for a more objective rating	Ratings may be skewed by training events affecting the person giving the rating

5.6 Other Issues

5.6.1 The Battle Simulation

The battle simulation used in this study was physically setup like a training exercise would have been for active-duty personnel. For example, this command center was setup within the simulation center, rather than in a command tent in the field. One of the largest consequences of this was there was very little physical interaction or movement in the TOC as the week progressed. Almost all observations at the end of the week had to be inferred from the conversations and radio communications in the room. There were also very few barriers to communication that are found in active-duty battle simulations, such as engine noise, weather conditions, etc. These factors may affect the performance of this tool in a more realistic environment.

5.6.2 Inter-Rater Reliability Issues

As stated previously, the widely varied experiences levels within the observer team may have contributed to the lack of inter-rater reliability observed in this study. There are several courses of action that could address this issue. First, the observers could undergo a much more rigorous training session on the usage of the tool. This training would require many hours, but it might be enough to normalize the ratings. Second, expert evaluators could be recruited to perform the observations. This might increase the inter-rater reliability, though this must be tested and not simply assumed. Finally, the tool can be redesigned to be more objective than its current state. Currently, the tool requires a subjective assessment of the team's performance. If a quantitative, objecting method can be applied to this methodology, inter-rater reliability should not be as much of an issue as it is currently.

The observer team was clearly heterogeneous. Some of the inter-rater reliability issues may be resolved by selecting a more homogeneous group of observers, at least with respect to knowledge of TOC operations. While a stated goal of this research was to move away from using C^2 subject matter experts, it may be enlightening to recruit a small group of these experts

in order to explore these inter-rater reliability issues. Doing so should help determine if the inter-rater reliability issues are a result of the tool's design or the selection of the raters.

5.6.3 Surviving Team Process Constructs

The results of the analysis phase of this study demonstrated that six of the ten tested constructs could be reliably and validly measured using the method used in this study. The team process constructs can be found in the following table.

Table 62 - Results on the Team Process Constructs

Individual Team Processes	Shared Team Processes
Coordination	Joint Knowledge
Communication	~~*Joint Task Ability*~~
Conflict Resolution	~~Motivation~~
Decision Making	Personalities
~~Problem Solving~~	~~Shared Thought Processes~~
~~Boundary Spanning~~	

It appears that it was easier to collect data on the individual team processes than it was the shared, more cognitive, team processes. This result could suggest several possibilities. First, the shared team processes represent a higher level of cognition that make inference imperative to measuring them, as opposed to observing the physical artifacts of the lower-level individual team processes. Second, it is possible that these shared team processes do not actually represent input factors to the individual team processes, as suggested. More data would need to be collected in order to determine the actual cause of this, however.

5.6.4 Ordinal vs. Interval Data

There may be differing opinions regarding whether or not the data collected in this study is truly interval data. The assumption is made that the data collected is interval in nature, since the analysis treats the data scale as Likert-type (Likert, 1932). However, the ten-point scale may

possess a level of detail too fine for human discrimination. This would classify the data as ordinal in nature. Interval data is ultimately desired, since there is an interested in producing a numerical score for the performance in each team process that can be compared across teams, rather than just rating the progress of one particular team across time. Reducing the number of gradations on the data scale should make the data more interval in nature than it is currently, as this would bring the scale closer to a Likert scale (Likert, 1932).

While the classification of the data type is of concern, the implications for the analysis portion of this particular study are minimal, since both analysis methods used demonstrated relatively low levels of inter-rater reliability. However, the uncertainty concerning whether this data is truly interval or not might lend more credibility to the Kendall's Concordance rather than the ANOVA for inter-rater reliability analysis. Once the observation tool is refined to use a closer form of the Likert scale, the ANOVA method for testing inter-rater reliability should gain credibility, as the data should become more interval in nature (Likert, 1932).

Chapter 6 - Conclusion

6.1 Implications for Human Factors Engineering

This research endeavor has focused on the military command and control environment. However, the underlying methodology tested in this study would be potentially sound in any team environment. Examples of where this methodology could be applied in a similar manner as was done in this study are trauma teams in emergency rooms, flight crews on commercial aircraft, and M1A2 tank crews in simulated combat. No literature was found during this endeavor to suggest that team performance had ever been measured using the methodology supplied in this research. Aside from the addition of a new team evaluation framework, this method provides a foundation for gathering the type of quantitative data that is necessary in the engineering world.

The specific method of evaluation suggested in this study was tailored to the military command and control environment, and that method may not be portable to all other team environments. However, the researcher suggests that the underlying methodology is portable to all types of teams. This methodology asserts that a key element to team performance is the team's performance on the processes that allow the team to work. This is to say that all teams should, by the assertion of the Team Effectiveness Model, possess all of the team processes, and that it should be possible to develop a specific method to measure these processes. However, that specific method would need to be tailored to the specific type of team that is being studied or evaluated.

Overall, the largest contribution to the field of Human Factors Engineering is the development of a new way of measuring team performance. This methodology for evaluation can be used for much more than team evaluation, however. It can be used in the research context to further examine how teams actually work. One of the more immediate ways of doing this is using this methodology to validate the Team Effectiveness Model. This further examination of teamwork is especially important as more and more of the workforce performs its work in teams and team effectiveness become more important.

6.2 Summary

As a first step in the development of a new methodology for evaluating team effectiveness, it was expected that the results would be somewhat mixed. The purpose of this study could be said to be one of identifying the challenges that need to be overcome with this line of research, such that the methodology can be fully tested in future studies. This study has shown that it is possible to collect data on the team processes using direct observation. It has also demonstrated a challenge in using subjective ratings from individuals with varied levels of experience in TOC operations, as shown by the inter-rater reliability scores.

Several factors existed in the testing of this tool that may have affected the results. Since the testing environment is a field study, there were not many experimental controls placed on the study. A major influence, which was beyond the control of this study, was the fidelity of this particular battle simulation. Another influence that was beyond the researcher's control was the formation of an incomplete command team. This occurred because of staffing limitations unknown prior to the start of the exercise and resulted in several positions being filled by a single individual and a severely understaffed opposing force.

Of the factors that were in the control of the experimenter, observer background and training may have been one of the largest influences. The training program developed was very limited. The development of this program was limited by the time available before the exercise began and the busy schedules of the individuals who volunteered to be observers in this study. It was also a premise of this research that the observers would not need to be experts in either C^2 operations or observation and evaluation. Additionally, there were no readily available examples that could be used in training the observers. To accommodate all these limitations, the training program was devised as an orientation to the data collection tool and focused on the meanings of each of the items in the tool. A simple, written scenario was developed and analyzed by the observers to practice using the tool. Based on the results, this training program should, perhaps, have been more formalized.

In order to make a tool of this type valid and reliable, either the observers must be extensively trained in TOC operations and observation techniques or the tool must be redesigned to further remove subjectivity from the ratings. The latter is preferable because it would increase the usability of the tool. However, such a redesign will require a significant amount of further research to specifically devise observable behaviors for each team process construct. This research may be even more difficult with the cognitive team processes, since outwardly visible signs of these processes may be difficult to detect. Observation of these cognitive processes would require inference from other processes, such as communication patterns, and decision-making activities.

Despite the complexities of pursuing this line of research, it has the potential to vastly increase our understanding of how teams function. The tool, itself, can be developed into both evaluative and diagnostic forms. Therefore, it is possible that not only could it provide performance feedback to a team, it could also identify the specific problems the team needs to correct in order to function better.

6.3 Future Research

This study has laid the foundation for a reliable and valid tool to allow collection of quantitative data on team process performance using direct observation, and it has provided a new methodology for team assessment. This new methodology provides a quantitative evaluation framework that is more suited to the engineering side of Human Factors Engineering. The study has also raised some interesting questions about measuring the team processes in order to adequately evaluate highly cognitive teams using non-invasive techniques.

The next obvious step in this line of research is to refine the tool so that it reliably assesses all of the team process constructs. This includes redesigning the manner in which the those constructs that failed the reliability tests are measured, as well as including the process that was inadvertently omitted from the tool, Joint Task Ability. One challenge that might need to be overcome is how to take the subjectivity out of the ratings. This may be an important challenge to overcome, because removing some of the subjectivity may be a good way to increase the

inter-rater reliability. It is possible that more extensive training could also increase inter-rater reliability, but there is no guarantee that even a very thorough training program will be sufficient with the current form of the tool. Another possibility that should be considered is revisiting the raw data collected from this study and reanalyzing using a different number of gradations on the Likert-type scale. For the types of subjective questions being posed in this tool, a ten-point scale may possess a level of detail that is not truly discernable to human observers. The nature in which the data was collected makes this type of reanalysis possible. It is the opinion of the researcher that there needs to be fewer than ten gradations on the analysis scale, with at least five gradations. Only further analysis can help to narrow down these guidelines.

The other research project that was in progress during this battle simulation produced audio/video recordings of the entire exercise. These recordings could be used to aid in the training of observers for future studies. Additionally, these recordings could be revisited to test any hypotheses generated from the current research study. As a result of these benefits, it may be advantageous to produce similar recordings in future battle simulations.

Aside from training the observer, there may be ways to modify the form of the tool to eliminate a large amount of observer subjectivity. One possibility is devising a system of expected behaviors that map to the process constructs and simply counting the occurrences of these behaviors. Such a system would be objective and quantitative. The training program for such a tool would be relatively succinct, depending on the number of defined behaviors, and would consist of an orientation to the behaviors that should be recorded and some practice with observing those behaviors. Another major benefit to having a tool designed in this manner is that it can be automated to a large degree. In the end, the simpler the tool is for the observer to use, the better.

Once the tool is complete, the next logical step is to compare the construct scores to conventional measures of effectiveness to determine the optimal score for each construct. Once the research progresses to this point, the tool becomes a diagnostic tool as well as an evaluation tool.

With a mature tool in use, it would also be possible to validate the cognitive team processes. It is suggested by Tannenbaum et al. (1992b) that these cognitive processes are input factors into the other team processes. This tool could validate that this is actually the case and determine how the cognitive processes affect the team processes. Otherwise, this tool could be used to determine if these team process attributes need to be modified to better fit team effectiveness. Further, this methodology could be used to validate the Team Effectiveness Model, once the specific method is refined.

In conclusion, further development of this line of research will add significantly to the knowledge base on command and control. Better understanding of command and control will allow us to improve the C^2 processes and equipment, as well as how to adapt each to the specific environment encountered. Once refined, this methodology could be used to evaluate and diagnose team functions in both the military and civilian realms. The possibilities for this line of research are endless.

References

Adler, P. A., and Adler, P. (1994). Observational Techniques. In N.K. Denzin (Ed.), Handbook of Qualitative Research (1 ed., pp. 377-392). Thousand Oaks, CA: SAGE Publications, Inc.

Arnold, K. J. (1999). Team Basics: Practical Strategies for Team Success. Hampton, VA: QPC Press.

Babbitt, B. A., and Nystrom, C. O. (1989). Questionnaire Construction Manual. Fort Hood, TX: U.S. Army Research Institute for the Behavioral and Social Sciences. Research Product 89-20.

Bachman, W. (1988). Nice Guys Finish First: A SYMLOG Analysis of U.S. Naval Commands. In R.B. Polley, A.P. Hare, and P.J. Stone (Eds.), The SYMLOG Practitioner: Applications of Small Group Research (pp. 133-154). New York: Praeger Publishers.

Bainbridge, L., and Sanderson, P. (1995). Verbal protocol analysis. In J.R. Wilson and E.N. Corlett (Eds.), Evaluation of human work (2nd ed., pp. 169-201). Bristol, PA: Taylor & Francis.

Baird, J. E., and Weinberg, S. B. (1981). Group Communication: The Essence of Synergy. (2nd ed.). Dubuque: Wm. C. Brown Company Publishers.

Bales, R. F., Cohen, S. P., and Williamson, S. A. (1979). SYMLOG: a system for the multiple level observation of groups. New York: The Free Press.

Barber, A. E., and Roehling, M. V. (1993). Job Postings and the Decision to Interview: A Verbal Protocol Analysis. Journal of Applied Psychology, 78(5), 845-856.

Barrick, M. R., Stewart, G. L., Neubert, M. J., and Mount, M. K. (1998). Relating Member Ability and Personality to Work-Team Processes and Team Effectiveness. Journal of Applied Psychology, 83(3), 377-391.

Best, P. S., Gentner, F. C., Cunningham, P. H., Tiller, T. C., and Schopper, A. W. (2001). Evaluating Aircrew and Maintainer Warfighter Performance in Aeronautical Systems Using Mission-Oriented Measures of Effectiveness. University of Dayton Research Institute: Human Factors Group. Final Report.

Biringen, Z. (1990). Direct Observation of Maternal Sensitivity and Dyadic Interactions in the Home: Relations to Maternal Thinking. Developmental Psychology, 26(2), 278-284.

Blacker, C. V. R., and Mortimore, C. (1996). Randomized Controlled Trials and Naturalistic Data: Time For a Change? Human Psychopharmacology-Clinical and Experimental, 11(5), 353-363.

Blomberg, J. L. (1995). Ethnography: aligning field studies of work and system design. In A.F. Monk and G.N. Gilbert (Eds.), Perspectives on HCI: Diverse Approaches (pp. 175-198). New York: Academic Press.

Borman, K. M., LeCompte, M. D., and Goetz, J. P. (1986). Ethnographic and qualitative research design and why it doesn't work. American Behavioral Scientist, 30(1), 42-57.

Bowers, C. A., and Jentsch, F. (2002). Using communication analysis to understand team development: An example. In Proceedings of the Human Factors and Ergonomics Society. Baltimore: Human Factors and Ergonomics Society.

Bull, P., and Roger, D. (1989). The social psychological approach to interpersonal communication. In D. Roger and P. Bull (Eds.), Conversation: An interdisciplinary perspective (pp. 9-20). Philadelphia: Multilingual Matters LTD.

Cannon-Bowers, J. A., Salas, E., and Converse, S. (1993). Shared Mental Models in Expert Team Decision Making. In N.J. Caltellan (Ed.), Individual and Group Decision Making: Current Issues (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., and Volpe, C. E. (1995). Defining Competencies and Establishing Team Training Requirements. In I.L. Goldstein (Ed.), Team Effectiveness and Decision Making in Organizations (pp. 333-380). San Francisco, CA: Jossey-Bass Inc.

Cheffers, J. T. F. (1972). The validation of an instrument designed to expand the Flanders System of Interaction Analysis to describe nonverbal interaction, different varieties of teacher behaviour and pupil responses. Unpublished Dissertation, Temple University, Philidelphia.

Christensen, A., and Hazzard, A. (1983). Reactive effects during naturalistic observation of families. Behavioral Assessment, 5(4), Inc.

Darst, P. W., Mancini, V. H., and Zakrajsek, D. B. (1983). Systematic Observation Instrumentation for Physical Education. West Point, NY: Leisure Press.

Dayal, I. (1969). Role Analysis Technique in Job Descriptions. California Management Review, 11(4), 47-50.

Downey, R. G., Duffy, P. J., Shiflett, S., and Helme, W. H. (1979). Criterion Performance Measures of Leadership and Unit Effectiveness in Small Combat Units. Army Research Institute for the Behavioral and Social Sciences. Memorandum. Research Memorandum 75-9.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2), 350-383.

Ericsson, A. K., and Simon, H. A. (1984). Protocol Analysis. Cambridge, MA: The MIT Press.

Festinger, L., Riecken, H. W., and Schachter, S. (1956). When Prophesy Fails. Minneapolis: University of Minnesota Press.

Fisher, S. G., Hunter, T. A., and MacRosson, W. D. K. (2001). A validation study of Belbin's team roles. European Journal of Work & Organizational Psychology, 10(2), 121-144.

Flanders, N. A. (1965). Teacher Influence, Pupil Attitudes, and Achievement. Washington DC: U.S. Department of Health, Education, and Welfare.

Flor, N. V., and Hutchins, E. L. (1993). Analyzing Distributed Cognition in Software Teams: A Case Study of Team Programming During Perfective Software Maintenance. In R.M. Baecker (Ed.), Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration (pp. 272-286). San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Fontana, A., and Frey, J. H. (2000). The Interview: From Structured Questions to Negotiated Text. In N.K. Denzin and Y.S. Lincoln (Eds.), Handbook of Qualitative Research (2nd ed.). Thousand Oaks, CA: Sage Publications.

Francis, D., and Young, D. (1979). Improving Work Groups. La Jolla, CA: University Associates.

Gentner, F. C., Best, P. S., and Cunningham, P. H. (1999). Sources of Measures of Effectiveness (MOEs) for Assessing Human Performance in Aeronautical Systems. Crew System Ergonomics Information Analysis Center (CSERIAC).

Gladstein, D. L. (1984). Groups in Context: A Model of Task Group Effectiveness. Administrative Science Quarterly, 29(4), 499-518.
Graziano, A. M., and Raulin, M. L. (1989). Research Methods: A Process of Inquiry. New York: Harper & Row, Publishers.
Greenbaum, J., and Kyng, M. (1991). Introduction: Situated Design. In J. Greenbaum and M. Kyng (Eds.), Design at work: cooperative design of computer systems (pp. 1-24). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Guzzo, R. A. (1986). Group Decision Making and Group Effectiveness in Organizations. In P.S. Goodman (Ed.), Designing Effective Work Groups (pp. 34-71). San Francisco, CA: Josey-Bass Inc.
Halpin, S. M. (1992). Army Command and Control Evaluation System (ACCES): A Brief Description. U.S. Army Research Institute for the Behavioral and Social Sciences, Ft. Leavenworth FE. Report 22.
Harper, W. R., and Harris, D. H. (1975). The application of link analysis to police intelligence. Human Factors, 17(2), 157-164.
Hendrick, H. W., and Kleiner, B. M. (2001). Macroergonomics: An Introduction to Work Systems Design.
Heritage, J. (1989). Current developments in conversation analysis. In D. Roger and P. Bull (Eds.), Conversation: An interdisciplinary perspective (pp. 21-47). Philadelphia: Multilingual Matters LTD.
Hickson, M., III. (1974). Participation-Observation Technique in Organizational Communication Research. The Journal of Business Communication, 11(3), 37.
Hofstee, W. K. B., de Raad, B., and Goldberg, L. R. (1992). Integration of the Big Five and Circumplex Approaches to Trait Structure. Journal of Personality and Social Psychology, 63(1), 146-163.
House, A. E. (1978). Naturalistic observation: Formal and informal difficulties. Child Study Journal, 8(1), 17-28.
Hutchins, E. (1995). How a Cockpit Remembers Its Speeds. Cognitive Science, 19, 265-288.
Jolliffe, I. T. (2002). Principal Components Analysis. (2nd ed.). New York: Springer.
Kirwan, B., and Ainsworth, L. K. (1992). A Guide to Task Analysis. Philadelphia: Taylor & Francis.
Kirwan, B., and Ainsworth, L. K. (1999). A Guide to Task Analysis. Philadelphia: Taylor & Francis.
Kleiner, B. M., and Drury, C. G. (1998). The use of verbal protocols to understand and design skill-based tasks. Human Factors and Ergonomics in Manufacturing, 8(1), 23-39.
Kline, P. (1994). An easy guide to factor analysis. New York: Routledge.
Layton, R. L., and Feld, P. (1996). Part-Task Performance Measures. Army Research Institute for the Behavioral and Social Sciences. Final Report. Research Note 96-18.
Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology, No. 140, 1-52.
Lindsey, P. H., and Norman, D. A. (1977). An Introduction to Psychology. (2nd ed.). New York: Academic Press.
Lipinski, D., and Nelson, R. (1974). Problems in the use of naturalistic observation as a means of behavioral assessment. Behavior Therapy, 5(3), 341-351.

McCrae, R. R. (1989). Why I Advocate the Five-Factor Model: Joint Factor Analyses of Other Instruments. In D.M. Buss and N. Cantor (Eds.), Personality Psychology: Recent Trends and Emerging Directions (pp. 237-245). New York: Springer-Verlag.

McDaniel, W. C. (1993). Naturalistic Group Decision Making: Overview and Summary. In N.J. Caltellan (Ed.), Individual and Group Decision Making: Current Issues (pp. 293-299). Hillsdale, NJ: Lawrence Erlbaum Associates.

McIntyre, R. M., and Salas, E. (1995). Measuring and Managing for Team Performance: Emerging Principles From Complex Environments. In I.L. Goldstein (Ed.), Team Effectiveness and Decision Making in Organizations (pp. 9-45). San Francisco, CA: Jossey-Bass Inc.

McManus, K. (2001). No feedback, no motivation. IIE Solutions, 33(4), 19.

McNiff, J., and Whitehead, J. (2000). Action Research in Organizations. New York: Taylor & Francis Ltd.

Molseed, M. J. (1994). Naturalistic Observation in the Laboratory. Symbolic Interaction, 17(3), 239-251.

Montgomery, D. C. (2001). Design and Analysis of Experiments. (5th ed.). New York: John Wiley and Sons, Inc.

Montgomery, D. C., and Runger, G. C. (1999). Applied Statistics and Probablility for Engineers. (2nd ed.). New York: John Wiley & Sons, Inc.

Moore, M. M. (2002). Behavioral observation. In M.W. Wiederman and B.E.J. Whitley (Eds.), Handbook for conducting research on human sexuality. (pp. 113-137). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Morgan, B. B., and Bowers, C. A. (1995). Teamwork Stress: Implications for Team Decision Making. In I.L. Goldstein (Ed.), Team Effectiveness and Decision Making in Organizations (pp. 262-290). San Francisco, CA: Jossey-Bass Inc.

Morris, B. (1971). Reflections on role analysis. British Journal of Sociology, 22(4), 395-409.

Nardi, B. A. (1996). Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition. In B.A. Nardi (Ed.), Context and Consciousness (pp. 69-102). Cambridge, MA: The MIT Press.

Neuman, G. A., Wagner, S. H., and Christiansen, N. D. (1999). The Relationship Between Work-Team Personality Composition and the Job Performance of Teams. Group & Organization Management, 24(3), 28-45.

Olmstead, J. A., Christensen, H. E., and Lackey, L. L. (1973). Components of Organizational Competence : Test of a Conceptual Framework. Human Resources Research Organization: Department of the Army, Office of the Chief of Research and Development. HumRRO Technical Report 73-19.

Olson, D. M., and Borman, W. C. (1987). Development and Feild Tests of the Army Work Environment Questionnaire. Army Research Institute for the Behavioral and Social Sciences. Technical Report. ARI Technical Report 737.

Peregrine, P. N., Drews, D. R., North, M., and Slupe, A. (1993). Sampling techniques and sampling error in naturalistic observation: An empirical evaluation with implications for cross-cultural research. Cross Cultural Research: the Journal of Comparative Social Science, 27(3-4), 232-246.

Polley, R. B., and Stone, P. J. (1988). An Introduction to SYMLOG. In R.B. Polley, A.P. Hare, and P.J. Stone (Eds.), The SYMLOG Practitioner: Applications of Small Group Research (pp. 3-14). New York: Praeger Publishers.

Pollock, T. (2001). Performance Insurance: Proper Motivation. Automotive Design & Production, 113(12).
Powell, J. (1969). Why Am I Afraid to Tell You Who I Am? Chicago: Argus Communications.
Prunkl, P. R., and Boyles, W. R. (1968). A preliminary application of the critical incident technique to combat performance of Army aviators. Human Resources Research Organization. HumRRO Professional Paper. No. 24-68.
Roger, D. (1989). Experimental studies of dyadic turn-taking behaviour. In D. Roger and P. Bull (Eds.), Conversation: An interdisciplinary perspective (pp. 75-95). Philadelphia: Multilingual Matters LTD.
Roth, J. T., Hritz, R. J., and McGill, D. W. (1984). Model of Team Organization and Behavior and Team Description Method. Applied Science Associates. Research Note.
Rutter, D. R. (1989). The role of cuelessness in social interaction: An examination of teaching by telephone. In D. Roger and P. Bull (Eds.), Conversation: An interdisciplinary perspective (pp. 294-312). Philadelphia: Multilingual Matters LTD.
Sanson-Fisher, R. W., Poole, A. D., Small, G. A., and Fleming, I. R. (1979). Data acquisition in real time: An improved system for naturalistic observations. Behavior Therapy, 10(4), 543-554.
Scholtes, P. R., Joiner, B. L., and Streibel, B. J. (2002). The Team Handbook. (2nd ed.). Madison: Oriel Incorporated.
Serfaty, D., Volpe, C., and Entin, E. E. (1993). Adaptation to stress in team decision-making and coordination. In Proceedings of the Human Factors and Ergonomics Society (pp. 1228-1232). Seattle, WA.
Shaiken, H. (1996). Experience and the collective nature of skill. In Y. Engestrom and D. Middleton (Eds.), Cognition and communication at work (pp. 279-295). New York: Cambridge University Press.
Shaughnessy, J. J., and Zechmeister, E. B. (1985). Research Methods in Psychology. New York: Alfred A. Knopf, Inc.
Siegel, S., and Castellan Jr., N. J. (1988). Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, Inc.
Sproles, N. (2001). Establishing Measures of Effectiveness for Command and Control: A Systems Engineering Perspective. University of South Australia. Contract Report to the Information Technology Division Electronics and surveillance Research Laboratory. DSTO-GD-0278.
Suen, H. K., and Ary, D. (1989). Analyzing Quantitative Behavioral Observation Data. Hillsdale, NJ: Lawrence Erlbaum Associates.
Sweet, R., Metersky, M., and Sovereign, M. (1985). Command and Control Evaluation Workshop. Naval Postgraduate School. Workshop Summary.
Swezey, R. W., and Llaneras, R. E. (1997). Models in Training and Instruction. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd ed., pp. 514-577). NY, New York: John Wiley & Sons, Inc.
Takemura, K. (1993). Protocol analysis of multistage decision strategies. Perceptual & Motor Skills, 77(2), 459-469.
Tannenbaum, S. I., Beard, R. L., and Salas, E. (1992a). Reprinted from Issues, Theory, and Research in Industrial/Organizational Psychology, Scott I. Tannenbaum, Rebecca L. Beard, and Eduardo Salas, "Team Building and Its Influence on Team Effectiveness: An

Examination of Conceptual and Empirical Developments", pp. 121., Copyright 1992, with permission from Elsevier.
Tannenbaum, S. I., Beard, R. L., and Salas, E. (1992b). Team Building and Its Influence on Team Effectiveness: An Examination of Conceptual and Empirical Developments. In K. Kelley (Ed.), Issues, Theory, and Research in Industrial/Organizational Psychology (pp. 117-153). New York: Elsevier Science Publishing Company.
Thomas, A. P. (1989). Social skills and physical handicap. In D. Roger and P. Bull (Eds.), Conversation: An interdisciplinary perspective (pp. 267-293). Philadelphia: Multilingual Matters LTD.
Tziner, A., and Eden, D. (1985). Effects of Crew Composition on Crew Performance: Does the Whole Equal the Sum of Its Parts? Journal of Applied Psychology, 70(1), 85-93.
U.S. Army. (1992). Tactics and Techniques for Combined Arms Heavy Forces: Armored Brigade, Battalion Task Force, and Company Team. U.S. Army. Field Manual, FM 71-123.
U.S. Army. (1997a). Operational Terms and Graphics. Department of the Army. Field Manual. FM 101-5-1.
U.S. Army. (1997b). Staff Organization and Operations. U.S. Army. Field Manual, FM 101-5.
Upenieks, V. V. (1998). Work sampling: Assessing nursing efficiency. Nursing Management, 29(4), 27.
Uzzell, D. (1995). Ethnographic and Action Research. In G.M. Breakwell, S. Hammond, and C. Fife-Schaw (Eds.), Research Methods in Psychology (pp. 302-313). Thousand Oaks, CA: SAGE Publication Inc.
Whitaker, L. A., Fox, S. L., and Peters, L. J. (1993). Communication between crews: The effects of speech intelligibility on team performance. In Proceedings of the Human Factors and Ergonomics Society (pp. 630-634). Seattle, WA.
Wilson, J. R., and Corlett, E. N. (1995). Evaluation of human work. (2nd ed.). Bristol, PA: Taylor & Francis.
Xiao, Y., Mackenzie, C. F., and Patey, R. (1998). Team coordination and breakdowns in a real-life stressful environment. Proceedings of the Human Factors Society, 1, 186-190.

Appendix A – Team Assessment Tool

Day _____
Observation Hour _____

Observer #_____

Exercise Day #_____

Day _____
Observation Hour _____

Activity

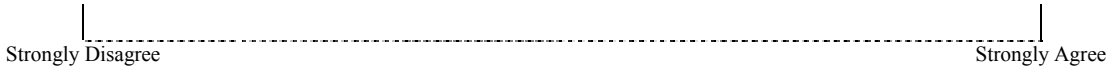
1. Please rate the overall level of activity experienced in this observation period.

Coordination

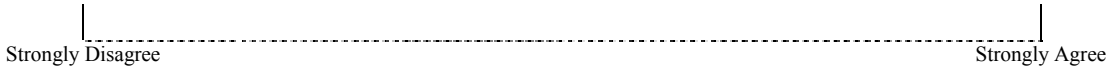
2. The Team appeared to be organized.

3. Team members anticipated events and requests often.

4. Team members presented the required information to the team when it was needed.

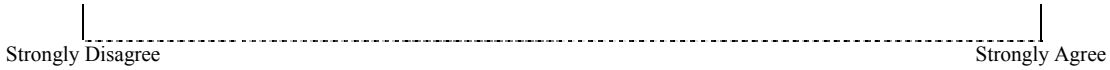


5. One or more team members appeared confused by the order of events.



6. The chain of command was clearly followed.

7. Without being asked, Team members frequently presented information or opinions that facilitated team functions.



Day _____
Observation Hour _____

Communication

8. Team members took turn in team communication.

9. Interruptions were common in team discussions.

10. Team members often repeated themselves when conversing with other team members (due to lack of understanding, noisy environment, not following orders, etc).

11. Team communication was often off-topic.

12. All section commanders participated in team discussions.

Conflict Resolution

13. Criticism given was constructive, rather than destructive.

14. All team members participated equally in team conversations.

15. "Name-calling" or expressions of another team member's incompetence never occurred (in a hostile tone, not in jest).

16. Criticism was the predominant team communication.

17. Overall, Team communication was civil.

Decision Making

18. Adequate information was available when making decisions.

19. The team always developed alternatives before deciding on a course of action.

20. Team meetings were used to make critical decisions.

21. Feedback was given to the team once the decision's results were available.

Day _____
Observation Hour _____

Problem Solving

22. The team members were understood the objectives.

23. One or more team members' actions or communications inhibited the function of the team.

24. Team members spent their time working toward the stated objectives.

Boundary Spanning

25. Troop actions were aligned with the orders given.

26. There was frequent communication between the field units and the command team.

Joint Knowledge

27. The team members understood the roles and responsibilities of the other team members.

28. Team members often showed surprise at the reaction or performance of other team member(s).

29. Team goals and objectives were clear to all team members.

30. Team members assisted other members when necessary.

Motivation

31. The team's morale is high?

32. Feedback is usually given for decisions and actions.

33. Workload is balanced across the team.

34. Team members are able to complete assignments in the time allowed.

Day _____
Observation Hour _____

Personalities

35. One or more team members' actions distracted the team from its mission.

Strongly Disagree Strongly Agree

36. Team members communicate with the other team members often.

Strongly Disagree Strongly Agree

37. One or more team member(s) were irritable or uncooperative.

Strongly Disagree Strongly Agree

38. Team members were enthusiastic about the exercise.

Strongly Disagree Strongly Agree

39. Team member' actions are careful and well organized.

Strongly Disagree Strongly Agree

Shared Thought Processes

40. Events, related to the exercise, took one or more team members by surprise.

Strongly Disagree Strongly Agree

41. Team members followed the strategy set forth by the commander.

Strongly Disagree Strongly Agree

Day _____
Observation Hour _____

42. Team members agreed with decisions, without being fully satisfied with the decision.

Strongly Disagree — Strongly Agree

Overall Performance

43. Overall, this team performed very well.

Strongly Disagree — Strongly Agree

Appendix B – Data Collection Manual

Introduction

This manual has been created to assist the observer in rating the item in the data collection tool. It consists of the items, a brief description of what the item is attempting to measure, and some things to look for in the observation sessions to help make the rating. The observable artifacts are given in tabular format and separated into items that represent the extremes of the rating scale. It is left to the observer to determine what balance exists between these endpoints.

This list of observable artifacts is not exhaustive and should only serve as a guide. If the observer feels another artifact should be added to the list, or taken away, space is provided in the manual for the observer to make comments concerning the manual. The observer is encouraged to provide this type of feedback.

The items are numbered and listed in the order found on the data collection tool for ease of searching.

1. Overall Activity: This is the general level of activity in the TOC. Things to look for in this category:

Decreased Activity	Increased Activity
Few relevant conversations	Many relevant conversations
Team members sitting idle	Most team members standing
Cracking jokes	Most team members working
Laughing	Conversation on phones/radios

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

2. The team appeared to be organized: This is a measure of how smoothly the TOC operated. Things to look for in this category:

Decreased Organization	Increased Organization
Littered work areas/TOC	Tidy work areas
Team members interrupt others	Team moves with purpose
Team members must search the TOC for what they need	Information is provided when or before it is asked for
People bumping into each other	Individual conversations do not interfere with each other
Conversations/Requests are repeated	

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

3. Team members anticipated (were not surprised by) battle events and anticipated team member requests often. This is a measure of the team's understanding of the sequence of battle and the functions of the team members. Things to look for in this category:

Decreased Anticipation	Increased Anticipation
Significant time delay between information request and receipt	Information is provided when or before it is asked for
Expressions of surprise at requests or events	Little notice taken at occurrence of events

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

4. Team members presented the required information to the team when it was needed. This is a measure of the team's understanding of the information that is needed to make decisions and the timing of those decisions. Things to look for in this category:

Less Timely Presentation	More Timely Presentation
Requester must wait for receipt of information Information is never provided Repeated requests for information	The requested information is presented while the requester is present at the provider's station Information provided in first request

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

5. One or more team members appeared confused by the order of events. This is a measure of the team's understanding of the progression of the battle plan. Things to look for in this category:

Less Confusion	More Confusion
Events pass with no special notice	Expressions of disbelief/surprise when events are made known "Why is this happening now?" "What just happened?"

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

6. The chain of command was clearly followed. This is a measure of how well the roles are defined in the team. Things to look for in this category:

Less Clearly Followed	More Clearly Followed
"Just follow orders!" Expressions of uncertainty of who is in charge of what	Orders flow from high to low rank Suggestions flow both directions

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

7. Without being asked, Command Team members frequently presented information or opinions that facilitated team functions. This is a measure of how well teams members contribute to teamwork and how comfortable team members are in the team. Things to look for in this category:

Less Frequent Presentation	More Frequent Presentation
Some team members never speak in meetings Individual opinions must be solicited	During decision making, team members speak without being asked

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

8. Team members took turns in team discussions. This is a measure of how well the team members understand how to communicate in a team setting. Things to look for in this category:

Less Orderly Communication	More Orderly Communication
Team members frequently interrupt during meetings Statements/Actions are made to silence a member	Only one team member speaks at a time in meetings There is an ordered progression of speakers evident

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

9. Interruptions were common in team discussions. This is a measure of the ability of the team to conduct discussions necessary to make decisions. Things to look for in this category:

Fewer Interruptions	More Interruptions
Meetings are conducted without interruption Team members isolate themselves for meetings	People enter and leave the meeting during the discussion Team members attend to other duties while in team meetings During the meeting, major events occur that require the attention of the team

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

10. Team members often repeated themselves when conversing with other team members (due to lack of understanding, noisy environment, not following orders, etc). This is a measure of how easily the team can communicate. Things to look for in this category:

Less Repeating	More Repeating
"Huddles" used for team meetings No repeating necessary for understanding	Raised voices Phrases had to be repeated before understood

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

11. Team communication was often irrelevant to the mission. This is a measure of how well the team could conduct relevant discussion necessary to make decisions and how much of the team's resources are used in making command and control decisions. Off-topic conversation represents team resources that are not being applied to military decision-making, either because of distraction or lack of need for decision-making at the given time. Things to look for in this category:

Less Off-Topic	More Off-Topic
Serious conversation Discussion concerning the mission High level of mission-related activity	Telling jokes Laughing Non mission-centered communication Attempts to communicate with Observers

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

12. A high percentage of section leaders participated in team discussions. This measures the team's ability to conduct team discussions. Things to look for in this category:

Less Participation	More Participation
Some leaders not participating in discussions No team-level discussions	All leaders participating in discussions

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

13. Criticism given was constructive, rather than destructive. This is a measure of how well interpersonal conflict is managed. Things to look for in this category:

Less Constructive Criticism	More Constructive Criticism
Suggestions not given for performance improvement Use of sarcasm in criticism Raised voices when criticism given Criticism given in front of entire team	Performance problem identified and suggestions made to correct Criticism given in private (may not be observed at all) No criticism observed

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

14. Team members who participated in team discussions participated equally. This is a measure of the team's ability to integrate all team members in the relevant discussions. This includes equal participation only from those involved in the conversation. Things to look for in this category:

Less Equal Participation	More Equal Participation
Conversations consisted of one person doing most of the talking Certain members of the conversation did little or no talking Conversations were strictly directive (giving orders) rather than conversations	All conversation members took part in the discussion Suggestions/Opinions were given by all conversation participants

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

15. "Name-calling" or expressions of another team member's incompetence never occurred. This is a measure of the team's ability to manage interpersonal conflict. Distinction should be made between comments made in jest. Such comments based in humor should not be considered here. Things to look for in this category:

More Frequent Expression	Less Frequent Expression
Personal attacks Public expressions of criticism "Humiliating" comments concerning personal performance	All criticism is constructive Criticism is given in "private" No criticism observed

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

16. Criticism was the predominant team communication. This is a measure of the type of communication that occurs in the TOC. Things to look for in this category:

Less Criticism	More Criticism
Mission-focused communication No criticism observed Calm voices	Criticism-focused communication Disciplinary actions were taken frequently Angry voices

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

17. Overall, Team communication was civil. This is a measure of the overall communication patterns experienced in the TOC. Things to look for in this category:

Less Civil	More Civil
Criticism-focused communication Raised voices (in anger) Marked gestures (finger-pointing, furrowed brow, tense facial expressions) during communication Talking about other team members behind his/her back	Mission-focused communication Calm communication Smiles during conversation Laughter

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

18. Adequate information was available when making decisions. This is a measure of the team's ability to collect information necessary to make decisions. Things to look for in this category:

Less Adequate Information	More Adequate Information
Questions unanswered at the time of decision making Expressions of uncertainty at information given for decisions Further information is requested at time of decision making	All questions have been answered by the time the decision is made

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

19. The team always developed alternatives before deciding on a course of action. This is a measure of the team's decision-making process. Things to look for in this category:

Less Alternative Development	More Alternative Development
Single solutions devised and implemented Statements illustrating the lack of options (i.e. 'We do not have time to think about it' or 'I see no other option')	Last step in decision making is choosing from alternative plans Statements illustrating different alternatives (i.e. We can do "this" or "that")

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

20. Team meetings (involving two or more team members) were used to make critical decisions. This measures the team's habit of including all relevant personnel in decision-making. Things to look for in this category:

Fewer Team Meetings	More Team Meetings
Single person makes most decisions Few leaders' meetings	More leaders' meetings to support decision making

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

21. Feedback was given to the team once the decision's results were available. This is the general level of activity in the TOC. Things to look for in this category:

Less Feedback	More Feedback
Little or no feedback given on decision outcomes Results of decisions made are not fully known in the TOC Decisions tend to be "fire and forget"	Feedback always given on decision outcomes Section leaders actively seek the results of their decisions

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

22. The team members fully understood the objectives. This is a measure of how well the objectives were communicated to the team members. Things to look for in this category:

Less Understanding	More Understanding
Team members criticized for taking wrong action Surprise expressed by the team commander (or XO) at the actions taken by friendly forces	Correct actions taken No surprise expressed by anyone as to the actions/movement of friendly forces

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

23. One or more team members' actions or communications inhibited the function of the team. This is a measure of the team's ability to manage the team communication processes. Things to look for in this category:

Less Disruption	More Disruption
Team member contributions are on-topic (i.e. no jokes, other issues) Team members take turns in conversations	Loud talking Team members interrupting conversations Physical expressions of anger/ frustration (i.e. "throwing fits")

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

24. Team members spent their time working toward the stated objectives. This is a measure of the team's dedication to the stated objectives. Things to look for in this category:

Less Productive Time	More Productive Time
Idle personnel Irrelevant conversation (telling jokes, talking about things other than the mission) Personnel attempting to chat with the Observers/Controllers	Personnel steadily working on mission-oriented tasks Team members helping others with mission tasks

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

25. Friendly force actions were aligned with the orders given. This is a measure of the clarity of orders issued from the TOC. Things to look for in this category:

Less Alignment	More Alignment
Surprise expressed by the Operations Officer at the movement/actions of the friendly forces Statements questioning battlefield events (i.e. Why did 'that' happen?), concerning friendly forces.	Battlefield actions go as planned

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

26. There was frequent communication between the field units and the command team. This is a measure of the team's habits in updating intelligence and clarifying orders to the field units. Things to look for in this category:

Less Communication	More Communication
No radio/telephone communication	Extensive use of radio/telephone communication

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

27. The team members understood the roles and responsibilities of the other team members. This is a measure of how familiar the team members are with working in this team. Things to look for in this category:

Less Understanding	More Understanding
Asking who does certain things Being redirected to another individual when requests are made Walking around the room looking for the person needed	Going straight to the person needed Following the correct Chain of Command

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

28. Team members showed surprise at the reaction or performance of other team member(s). This is a measure of how familiar the team members are with each other. Things to look for in this category:

Less Surprise	More Surprise
Normal conversive stance	Facial expressions of surprise (i.e. wide eyes, dropped jaw) Verbal expressions of surprise (i.e. "Chill out") Posture expressions (i.e. backing away from the individual, hand in the air) Disciplinary action taken for behavior or performance

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

29. Team goals and objectives were clear to all team members. This is a measure of the team members' familiarity with how the team should operate. Things to look for in this category:

Less Clarity	More Clarity
Repeating of the goals/objectives Criticism on performance of duties	Mission follows without repeating objectives

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

30. Team members assisted other members when necessary. This is a measure of how well the team members receive and interpret action cues during team operations. Things to look for in this category:

Less Assistance	More Assistance
Everyone only does his/her own job Some members are hard at work, while others are sitting idle for extended periods of time	Team members helping where they can

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

31. The team's morale is high. This measures the morale input to operational motivation. Things to look for in this category:

Low Morale	High Morale
Grumpiness/Irritability Argumentative Low energy	Smiles High energy Celebratory with victories

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

32. Feedback is usually given for decisions and actions. This is a measure of the team's feedback mechanisms necessary for performance improvement. Things to look for in this category:

Less Feedback	More Feedback
No feedback on performance at all	Congratulations Criticism (constructive or destructive)

Please, add any items you believe should be included in this list, but are not currently included:

_____

33. Workload is balanced across the team. This is a measure for how the team manages workload. Things to look for in this category:

Less Balance	More Balance
Some people very busy, others completely idle	Everyone is working Everyone is idle Idle personnel assist those with heavier workloads

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

34. Team members are able to complete assignments in the time allowed. This is a measure of the team members' abilities and the realism of the superiors' requirements. Things to look for in this category:

Less Able	More Able
Decision makers must wait longer than normal for information Personnel requests for more time to complete assignments	Information is available when requested Some idle time apparent

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

35. One or more team members' actions distracted the team from its mission. This is a measure of how individual team members' personalities contribute or distract from pursuing the mission. Things to look for in this category:

Less Distracting	More Distracting
Normal voice levels Calm environment Relevant conversations	Loud voices Arguing "Acting out" Telling jokes/Irrelevant conversations

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

36. Team members communicate with the other team members often. This is a measure of the team members' extraversion. Things to look for in this category:

Less Communication	More Communication
Team members mostly work alone at workstation	High interpersonal interaction High levels of radio/telephone traffic

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

37. One or more team member(s) were irritable or uncooperative. This is a measure of how well the members of the team can interact as a team. Things to look for in this category:

Less Irritable	More Irritable
Calm speech Relaxed expressions Smiles	Grated/Harsh conversation style Tense expressions (i.e. red face) Scowls Certain individuals are avoided, if possible

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

38. Team members were enthusiastic about the exercise. This is a measure of the team members' morale and dedication to the stated objectives of the team. Things to look for in this category:

Less Enthusiasm	More Enthusiasm
Low energy Frowns Napping 'This exercise is pointless'	High energy Smiles Celebratory with victories Certain or all individuals act as morale boosters

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

39. Team members' actions are careful and well organized. This is a measure of the thought and planning processes enlisted in performing the team's mission(s). Things to look for in this category:

Less Conscientious	More Conscientious
Haphazard organization of information Mistakes in information presentation Information was needed that was not presented	Workspaces well organized Information presented clearly and precisely Correct information is found and presented on the first attempt

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

40. Events, related to the exercise, took one or more team members by surprise. This is a measure of the team's understanding of the progression of the battle. Things to look for in this category:

Less Surprising	More Surprising
Team members take little notice of events as they transpire	Flurry of activity around events to determine the cause of the event Questioning the reason for events Questioning, "What just happened?"

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

41. Team members followed the strategy set forth by the Battle Captain. This measures the understanding of the team strategy and the trust the team members have in the leader. Things to look for in this category:

Less Adherence	More Adherence
Criticism from the BC concerning the actions taken by the TOC	Little/No feedback from the BC given concerning how the battle progressed

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

42. Team members agreed with decisions, without being fully satisfied with the decision. This measures the team's habit of carrying out decisions without a consensus. Things to look for in this category:

More Consensus	Less Consensus
All members agree to the course of action	Expressions of frustration with a course of action Decision making discussions end with a harsh directive (i.e. "Just do it!")

Please, add any items you believe should be included in this list, but are not currently included:

_____

_____

43. Overall, this team performed very well. This measures your overall impression of how well the team performed and is based on your opinion from the observations.

Appendix C – Observer Training Materials

Training Outline

1. Introduction

1.1. Record start time
1.2. Personal introductions
1.3. IRB
1.4. What the study is trying to accomplish
1.5. Role of the observers in the study

2. Observer Duties

2.1. Continuously observe the command team in the TOC
2.2. Complete the observation tool each hour, on the hour
2.3. Can move around the TOC
2.4. Can observe other cells, if necessary to facilitate understanding of events in the TOC
2.5. Do not interact with the command team
2.6. Do not obstruct the command team
2.6.1. *The command team should perform the way they would if you were not there*

3. Collection Tool

3.1. Go through the tool, item by item
3.1.1. *Discuss the meanings of each item*
3.1.2. *Use the Manual that has been developed*
3.1.3. *Get input from the observers on attributes, update as necessary*
3.1.4. *Answer any questions*
3.2. Complete collection tool, based on scenario
3.2.1. *Report individual ratings on the items (far left, toward the left, center, toward the right, far right)*
3.2.2. *Discuss ratings and reasons for each rating*
3.2.2.1. *Give specific examples if possible*
3.2.2.2. *Try to achieve consensus on ratings through understanding the reasons for each rating and normalizing the reasons across the observers*
3.2.2.3. *Modify manual as necessary*

4. Conclusion

4.1. Ask for any questions, concerns with the tool
4.2. Details of the exercise
4.2.1. *Walkthrough an exercise*
4.2.2. *Will be given one observation tool at a time*
4.2.3. *I will indicate when it is time to complete the tool (every hour on the hour)*
4.2.4. *09:00 to 17:00 CT each day 10 – 13 June*
4.2.5. *Report 08:45 CT at the latest*
4.2.6. *TOC is in Room 160 and 162 (one big room) in the Battle Simulation Center (BSC) Building 4501*
4.3. Final Questions/Concerns

Observer Guidelines

1. **Please, mark all ratings with a vertical, straight line.**
   - **a.** The rating you make will be read by the point where this line intersects the line of the scale.
2. **Do not interact with the people wearing blue or red name badges.**
   - **a.** This observation is to remain as "naturalistic" as possible. This means that the command team must function as if you were not there. Do not initiate conversations with any of the team members and politely dissuade team members from initiating conversation with you.
3. **Do not obstruct the function of the people wearing blue or red name badges.**
   - **a.** The red and blue players will be actively involved with coordinating the battle. This means that activity will be high at times. Take all precautions to ensure that you do not interfere with this activity. Do not obstruct movement or vision. Be mindful of the placement maps and displays and heavily traveled walkways.
4. **You may travel to other command centers.**
   - **a.** If you need to visit another command center (ALOC, Red Cell, etc.) in order to understand what is going on in the TOC, feel free to travel to that command center. Often, these other command centers supply the stimuli for the TOC, and it may be beneficial to see how this is done.
5. **Remain in the TOC more than travel to other centers.**
   - **a.** The majority of the action of interest in this study will be occurring in the TOC. Please, spend the vast majority of your time in the TOC
6. **If you have any questions regarding the data collection, please ask Brent Fraser. If Brent is not available, ask Dr. Middlebrooks.**

Training Scenario

Imagine you are now participating in observing an operational TOC.

Over the past hour, you have been watching the action. For the most part, the section leaders in the TOC have been busy, either in meetings or organizing data from their staffs. However, the pace is not harried.

The S4 has been chatting about his wife to the S1, who laughs occasionally at what the S4 has to say. The Battle Captain made a comment about how he wishes they would get to work. They respond that there is nothing for them to do right now. The BC walks away, and the S4 and S1 continue their conversation, but with lowered voices.

A report came in that the river crossing came under attack by artillery fire from a wooded hill that was considered secure. This report sends the BC into a bit of a rage, and the XO suggests that he go outside and 'get some air'. Several calls are made to the Recon teams in the area to determine the situation. The XO orders new intelligence and the formulation of a plan to eliminate or circumvent the threat and cross the river. The river crossing is temporarily halted until the situation can be assessed. Casualties were small, but much of the river-crossing equipment was damaged and is still exposed to enemy fire.

Radio traffic become non-stop in the last 15 minutes of the hour

The S3 returns from the field near the end of the hour. The hour's observation ends with the S2, S3, and XO in a huddle.

Based on this fictional scenario, please complete the observation tool. While you complete the tool, please, feel free to make any comments on the tool, the clarity of the questions, the wording of the questions, etc. This exercise is only being used to determine the suitability of this tool for deployment in this exercise.

To aid with determining how to rate each question, a training manual is provided for reference. This manual explains what each question is looking for and specific examples of how phenomenon in the TOC should be rated. If you have any questions that cannot be answered by the manual, please ask.

Thank you for your participation in this exercise!

Appendix D – Kendall's Concordance Calculation Details

Item 1																											Sum(Tj)
Item 1 Observer 1	5.5	22.5	5.5	5.5	5.5	1.0	5.5	5.5	22.5	14.5	14.5	14.5	24.0	25.0	5.5	20.5	5.5	14.5	14.5	14.5	14.5	14.5	14.5	20.5	14.5	1506
Item 1 Observer 2	6.0	25.0	6.0	22.5	6.0	22.5	15.0	6.0	6.0	15.0	20.0	20.0	20.0	15.0	6.0	6.0	15.0	15.0	6.0	15.0	6.0	24.0	15.0	6.0	6.0	1686
Item 1 Observer 3	14.5	19.5	21.0	24.5	14.5	14.5	5.0	9.0	14.5	5.0	14.5	14.5	22.5	24.5	9.0	14.5	14.5	19.5	9.0	5.0	1.5	22.5	5.0	5.0	1.5	672
R	26.0	67.0	32.5	52.5	26.0	38.0	25.5	20.5	43.0	34.5	49.0	49.0	66.5	64.5	20.5	41.0	35.0	49.0	29.5	34.5	22.0	61.0	34.5	31.5	22.0	3864
Rbar	8.7	22.3	10.8	17.5	8.7	12.7	8.5	6.8	14.3	11.5	16.3	16.3	22.2	21.5	6.8	13.7	11.7	16.3	9.8	11.5	7.3	20.3	11.5	10.5	7.3	13
sumRi,squared	43251		N =	25		k =	3																			
W =	0.4869																									
Chi^2,calc =	35.054		v =	24																						
tabled value =	7%																									
	There is a 0.067600901187159 chance that the agreement among the raters was by random chance.																									

Item 3																											Sum(Tj)
Item 3 Observer 1	8.5	19.5	8.5	16.0	8.5	8.5	8.5	16.0	21.5	2.0	21.5	8.5	16.0	2.0	24.0	24.0	24.0	8.5	8.5	8.5	2.0	8.5	19.5	16.0	16.0	1170
Item 3 Observer 2	3.0	13.0	17.5	21.5	21.5	17.5	13.0	13.0	13.0	8.0	21.5	8.0	24.5	13.0	21.5	16.0	24.5	8.0	3.0	3.0	3.0	19.0	8.0	8.0	3.0	432
Item 3 Observer 3	23.0	12.5	12.5	23.0	23.0	4.5	20.0	17.5	20.0	12.5	25.0	12.5	4.5	4.5	20.0	12.5	17.5	12.5	4.5	4.5	4.5	12.5	4.5	12.5	4.5	1062
R	34.5	45.0	38.5	60.5	53.0	30.5	41.5	46.5	54.5	22.5	68.0	29.0	45.0	19.5	65.5	52.5	66.0	29.0	16.0	16.0	9.5	40.0	32.0	36.5	23.5	2664
Rbar	11.5	15.0	12.8	20.2	17.7	10.2	13.8	15.5	18.2	7.5	22.7	9.7	15.0	6.5	21.8	17.5	22.0	9.7	5.3	5.3	3.2	13.3	10.7	12.2	7.8	13
sumRi,squared	44683		N =	25		k =	3																			
W =	0.6034																									
Chi^2,calc =	43.442		v =	24																						
tabled value =	1%																									
	There is a 0.00885189613287642 chance that the agreement among the raters was by random chance.																									

Item 4																											Sum(Tj)
Item 4 Observer 1	10.5	20.5	10.5	20.5	10.5	10.5	10.5	20.5	20.5	10.5	10.5	10.5	10.5	24.5	10.5	10.5	10.5	10.5	10.5	1.5	1.5	10.5	24.5	23.0	10.5	4152
Item 4 Observer 2	4.5	22.0	14.0	22.0	14.0	14.0	14.0	4.5	14.0	14.0	22.0	14.0	22.0	22.0	14.0	14.0	14.0	14.0	4.5	4.5	4.5	25.0	4.5	4.5	4.5	1944
Item 4 Observer 3	15.5	21.5	6.5	24.5	24.5	6.5	21.5	15.5	21.5	15.5	19.0	15.5	6.5	6.5	21.5	15.5	15.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	1992
R	30.5	64.0	31.0	67.0	49.0	31.0	46.0	40.5	56.0	40.0	51.5	40.0	39.0	53.0	46.0	40.0	40.0	31.0	21.5	12.5	12.5	42.0	35.5	34.0	21.5	8088
Rbar	10.2	21.3	10.3	22.3	16.3	10.3	15.3	13.5	18.7	13.3	17.2	13.3	13.0	17.7	15.3	13.3	13.3	10.3	7.2	4.2	4.2	14.0	11.8	11.3	7.2	13
sumRi,squared	42607		N =	25		k =	3																			
W =	0.4734																									
Chi^2,calc =	34.088		v =	24																						
tabled value =	8%																									
	There is a 0.083116604121456 chance that the agreement among the raters was by random chance.																									

Item 5																											Sum(Tj)
Item 5 Observer 1	14.5	18.5	21.5	14.5	25.0	6.0	6.0	21.5	21.5	6.0	14.5	6.0	14.5	6.0	14.5	6.0	6.0	6.0	6.0	21.5	6.0	6.0	24.0	18.5	14.5	1596
Item 5 Observer 2	3.5	13.5	21.0	8.5	23.5	17.5	17.5	13.5	20.0	8.5	23.5	13.5	23.5	13.5	23.5	13.5	8.5	8.5	3.5	3.5	3.5	19.0	13.5	3.5	3.5	546
Item 5 Observer 3	20.5	15.5	7.0	25.0	24.0	20.5	20.5	7.0	20.5	7.0	7.0	7.0	7.0	7.0	15.5	20.5	7.0	7.0	20.5	7.0	7.0	15.5	7.0	15.5	7.0	2454
R	38.5	47.5	49.5	48.0	72.5	44.0	44.0	42.0	62.0	21.5	45.0	26.5	45.0	26.5	53.5	40.0	21.5	21.5	30.0	32.0	16.5	40.5	44.5	37.5	25.0	4596
Rbar	12.8	15.8	16.5	16.0	24.2	14.7	14.7	14.0	20.7	7.2	15.0	8.8	15.0	8.8	17.8	13.3	7.2	7.2	10.0	10.7	5.5	13.5	14.8	12.5	8.3	13
sumRi,squared	42381		N =	25		k =	3																			
W =	0.4128																									
Chi^2,calc =	29.722		v =	24																						
tabled value =	19%																									
	There is a 0.194139821065836 chance that the agreement among the raters was by random chance.																									

Item 8																											Sum(Tj)
Item 8 Observer 1	10.5	10.5	10.5	10.5	10.5	2.0	10.5	25.0	20.5	20.5	20.5	10.5	10.5	24.0	10.5	10.5	10.5	10.5	10.5	2.0	2.0	10.5	20.5	20.5	20.5	2964
Item 8 Observer 2	4.0	22.5	24.5	24.5	21.0	18.5	18.5	12.0	12.0	18.5	12.0	12.0	18.5	12.0	12.0	12.0	12.0	12.0	4.0	4.0	4.0	22.5	4.0	4.0	4.0	1128
Item 8 Observer 3	5.5	5.5	25.0	11.5	5.5	5.5	18.0	22.5	11.5	18.0	5.5	18.0	11.5	18.0	11.5	1.5	24.0	22.5	11.5	5.5	18.0	1.5	18.0	18.0	11.5	764
R	20.0	38.5	60.0	46.5	37.0	26.0	47.0	59.5	44.0	57.0	38.0	40.5	40.5	54.0	34.0	24.0	46.5	45.0	26.0	11.5	24.0	34.5	42.5	42.5	36.0	4856
Rbar	6.7	12.8	20.0	15.5	12.3	8.7	15.7	19.8	14.7	19.0	12.7	13.5	13.5	18.0	11.3	8.0	15.5	15.0	8.7	3.8	8.0	11.5	14.2	14.2	12.0	13
sumRi,squared	41667		N =	25		k =	3																			
W =	0.3473																									
Chi^2,calc =	25.004		v =	24																						
tabled value =	41%																									
	There is a 0.405563777659616 chance that the agreement among the raters was by random chance.																									

Item 12																											Sum(Tj)
Item 12 Observer 1	20.0	20.0	3.5	11.5	11.5	11.5	11.5	11.5	24.0	11.5	20.0	11.5	20.0	25.0	3.5	11.5	3.5	11.5	3.5	3.5	3.5	11.5	20.0	20.0	20.0	1536
Item 12 Observer 2	12.0	23.0	19.5	12.0	23.0	19.5	19.5	12.0	12.0	23.0	19.5	12.0	12.0	12.0	12.0	12.0	12.0	12.0	3.5	3.5	3.5	25.0	3.5	3.5	3.5	1614
Item 12 Observer 3	19.5	19.5	13.0	7.0	7.0	19.5	7.0	7.0	7.0	13.0	3.5	19.5	13.0	19.5	13.0	3.5	19.5	25.0	13.0	1.5	1.5	23.0	24.0	13.0	13.0	678
R	51.5	62.5	36.0	30.5	41.5	50.5	38.0	30.5	43.0	47.5	43.0	43.0	45.0	56.5	28.5	27.0	35.0	48.5	20.0	8.5	8.5	59.5	47.5	36.5	36.5	3828
Rbar	17.2	20.8	12.0	10.2	13.8	16.8	12.7	10.2	14.3	15.8	14.3	14.3	15.0	18.8	9.5	9.0	11.7	16.2	6.7	2.8	2.8	19.8	15.8	12.2	12.2	13
sumRi,squared	42576		N =	25		k =	3																			
W =	0.4236																									
Chi^2,calc =	30.501		v =	24																						
tabled value =	17%																									
	There is a 0.168697975495063 chance that the agreement among the raters was by random chance.																									

Item 13																										Sum(Tj)
Item 13 Observer 1	4.0	22.0	14.0	14.0	14.0	14.0	14.0	25.0	22.0	4.0	14.0	14.0	22.0	24.0	14.0	4.0	4.0	4.0	14.0	4.0	4.0	14.0	14.0	14.0	14.0	2544
Item 13 Observer 2	18.5	22.0	22.0	25.0	14.5	22.0	14.5	22.0	5.5	5.5	5.5	22.0	18.5	12.0	5.5	12.0	16.5	12.0	5.5	5.5	5.5	16.5	5.5	5.5	5.5	1152
Item 13 Observer 3	9.0	9.0	9.0	9.0	9.0	9.0	9.0	25.0	22.5	20.0	22.5	9.0	18.5	9.0	22.5	9.0	22.5	9.0	9.0	9.0	9.0	9.0	9.0	18.5	9.0	4962
R	31.5	53.0	45.0	48.0	37.5	45.0	37.5	72.0	50.0	29.5	42.0	45.0	59.0	45.0	42.0	25.0	43.0	25.0	28.5	18.5	18.5	39.5	28.5	38.0	28.5	8658
Rbar	10.5	17.7	15.0	16.0	12.5	15.0	12.5	24.0	16.7	9.8	14.0	15.0	19.7	15.0	14.0	8.3	14.3	8.3	9.5	6.2	6.2	13.2	9.5	12.7	9.5	13
sumRi,squared	41806		N =	25		k =	3																			
W =	0.3965																									
Chi^2,calc =	28.546		v =	24																						
tabled value =	24%																									
	There is a 0.237704133406529 chance that the agreement among the raters was by random chance.																									

Item 16																										Sum(Tj)
Item 16 Observer 1	11.5	11.5	11.5	11.5	11.5	24.0	11.5	11.5	24.0	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	24.0	11.5	10650
Item 16 Observer 2	8.5	23.0	14.0	14.0	18.5	24.5	18.5	24.5	14.0	8.5	14.0	14.0	20.5	14.0	14.0	8.5	20.5	8.5	3.5	3.5	3.5	22.0	3.5	3.5	3.5	624
Item 16 Observer 3	10.0	10.0	10.0	10.0	10.0	10.0	10.0	20.5	22.5	10.0	25.0	10.0	10.0	10.0	20.5	10.0	22.5	10.0	10.0	10.0	10.0	10.0	10.0	24.0	10.0	6852
R	30.0	44.5	35.5	35.5	40.0	58.5	40.0	56.5	60.5	30.0	50.5	35.5	42.0	35.5	46.0	30.0	54.5	30.0	25.0	25.0	25.0	43.5	25.0	51.5	25.0	18126
Rbar	10.0	14.8	11.8	11.8	13.3	19.5	13.3	18.8	20.2	10.0	16.8	11.8	14.0	11.8	15.3	10.0	18.2	10.0	8.3	8.3	8.3	14.5	8.3	17.2	8.3	13
sumRi,squared	41166		N =	25		k =	3																			
W =	0.4382																									
Chi^2,calc =	31.548		v =	24																						
tabled value =	14%																									
	There is a 0.138601855897341 chance that the agreement among the raters was by random chance.																									

Item 19																										Sum(Tj)
Item 19 Observer 1	22.5	9.5	17.0	20.0	22.5	17.0	22.5	25.0	9.5	3.5	17.0	9.5	9.5	17.0	13.5	3.5	22.5	3.5	3.5	9.5	3.5	3.5	17.0	13.5	9.5	606
Item 19 Observer 2	20.0	12.0	20.0	20.0	24.0	20.0	20.0	12.0	12.0	12.0	12.0	12.0	20.0	20.0	12.0	12.0	4.0	12.0	4.0	4.0	4.0	25.0	4.0	4.0	4.0	1398
Item 19 Observer 3	23.0	19.0	1.0	3.0	14.5	19.0	7.0	7.0	14.5	7.0	10.5	7.0	23.0	19.0	23.0	14.5	23.0	14.5	7.0	23.0	3.0	10.5	3.0	14.5	14.5	504
R	65.5	40.5	38.0	43.0	61.0	56.0	49.5	44.0	36.0	22.5	39.5	28.5	52.5	56.0	48.5	30.0	49.5	30.0	14.5	36.5	10.5	39.0	24.0	32.0	28.0	2508
Rbar	21.8	13.5	12.7	14.3	20.3	18.7	16.5	14.7	12.0	7.5	13.2	9.5	17.5	18.7	16.2	10.0	16.5	10.0	4.8	12.2	3.5	13.0	8.0	10.7	9.3	13
sumRi,squared	42694		N =	25		k =	3																			
W =	0.4217																									
Chi^2,calc =	30.359		v =	24																						
tabled value =	17%																									
	There is a 0.17312852843567 chance that the agreement among the raters was by random chance.																									

Item 20																										Sum(Tj)
Item 20 Observer 1	7.0	14.5	3.0	7.0	1.0	22.0	14.5	25.0	14.5	22.0	24.0	14.5	14.5	22.0	14.5	3.0	7.0	7.0	3.0	7.0	14.5	14.5	14.5	20.0	14.5	1158
Item 20 Observer 2	5.5	21.5	5.5	21.5	21.5	21.5	21.5	14.5	14.5	5.5	14.5	14.5	21.5	5.5	14.5	14.5	14.5	14.5	5.5	5.5	5.5	25.0	5.5	5.5	5.5	1704
Item 20 Observer 3	24.0	19.0	2.5	11.5	11.5	5.0	17.5	20.5	22.0	11.5	17.5	2.5	24.0	11.5	11.5	1.0	11.5	5.0	11.5	11.5	5.0	11.5	20.5	11.5	24.0	1056
R	36.5	55.0	11.0	40.0	34.0	48.5	53.5	60.0	51.0	39.0	56.0	31.5	60.0	39.0	40.5	18.5	33.0	26.5	20.0	24.0	25.0	51.0	40.5	37.0	44.0	3918
Rbar	12.2	18.3	3.7	13.3	11.3	16.2	17.8	20.0	17.0	13.0	18.7	10.5	20.0	13.0	13.5	6.2	11.0	8.8	6.7	8.0	8.3	17.0	13.5	12.3	14.7	13
sumRi,squared	42341		N =	25		k =	3																			
W =	0.4026																									
Chi^2,calc =	28.987		v =	24																						
tabled value =	22%																									
	There is a 0.220631521478554 chance that the agreement among the raters was by random chance.																									

Item 28																										Sum(Tj)
Item 28 Observer 1	14.5	14.5	5.0	22.0	5.0	5.0	14.5	22.0	5.0	14.5	5.0	14.5	14.5	5.0	25.0	22.0	14.5	14.5	5.0	5.0	5.0	22.0	22.0	14.5	14.5	1119
Item 28 Observer 2	8.5	8.5	11.0	16.0	20.5	16.0	16.0	24.0	13.0	8.5	16.0	20.5	20.5	24.0	20.5	16.0	24.0	8.5	3.5	3.5	3.5	12.0	3.5	3.5	3.5	474
Item 28 Observer 3	10.5	10.5	10.5	10.5	24.0	10.5	10.5	21.0	10.5	10.5	22.5	10.5	10.5	10.5	22.5	25.0	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	7986
R	33.5	33.5	26.5	48.5	49.5	31.5	41.0	67.0	28.5	33.5	43.5	45.5	45.5	39.5	68.0	63.0	49.0	33.5	19.0	19.0	19.0	44.5	36.0	28.5	28.5	9579
Rbar	11.2	11.2	8.8	16.2	16.5	10.5	13.7	22.3	9.5	11.2	14.5	15.2	15.2	13.2	22.7	21.0	16.3	11.2	6.3	6.3	6.3	14.8	12.0	9.5	9.5	13
sumRi,squared	42539		N =	25		k =	3																			
W =	0.4851																									
Chi^2,calc =	34.927		v =	24																						
tabled value =	7%																									
	There is a 0.0694849263976181 chance that the agreement among the raters was by random chance.																									

Item 29																										Sum(Tj)
Item 29 Observer 1	15.0	15.0	15.0	23.5	4.5	15.0	4.5	23.5	15.0	15.0	15.0	15.0	15.0	15.0	15.0	4.5	4.5	4.5	4.5	4.5	4.5	15.0	23.5	23.5	15.0	2748
Item 29 Observer 2	9.0	15.5	19.5	19.5	19.5	19.5	11.5	23.5	13.5	4.0	13.5	23.5	19.5	11.5	25.0	19.5	9.0	9.0	4.0	4.0	4.0	15.5	4.0	4.0	4.0	594
Item 29 Observer 3	25.0	21.5	21.5	16.0	16.0	24.0	21.5	16.0	16.0	6.5	16.0	6.5	16.0	16.0	21.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	2112
R	49.0	52.0	56.0	59.0	40.0	58.5	37.5	63.0	44.5	25.5	44.5	45.0	50.5	42.5	61.5	30.5	20.0	20.0	15.0	15.0	15.0	37.0	34.0	34.0	25.5	5454
Rbar	16.3	17.3	18.7	19.7	13.3	19.5	12.5	21.0	14.8	8.5	14.8	15.0	16.8	14.2	20.5	10.2	6.7	6.7	5.0	5.0	5.0	12.3	11.3	11.3	8.5	13
sumRi,squared	43631		N =	25		k =	3																			
W =	0.5423																									
Chi^2,calc =	39.046		v =	24																						
tabled value =	3%																									
	There is a 0.0270090636825304 chance that the agreement among the raters was by random chance.																									

Item 35																										Sum(Tj)
Item 35 Observer 1	12.0	23.5	12.0	12.0	12.0	12.0	12.0	12.0	12.0	23.5	12.0	12.0	1.5	12.0	12.0	12.0	12.0	12.0	12.0	12.0	1.5	23.5	23.5	12.0	12.0	6906
Item 35 Observer 2	9.5	17.0	17.0	17.0	17.0	20.5	17.0	22.5	4.5	9.5	12.0	12.0	24.5	12.0	22.5	20.5	24.5	4.5	4.5	4.5	4.5	14.0	4.5	4.5	4.5	672
Item 35 Observer 3	23.5	10.5	10.5	10.5	10.5	10.5	10.5	21.5	10.5	10.5	10.5	10.5	10.5	10.5	21.5	23.5	25.0	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	7992
R	45.0	51.0	39.5	39.5	39.5	43.0	39.5	56.0	27.0	43.5	34.5	34.5	36.5	34.5	56.0	56.0	61.5	27.0	27.0	27.0	16.5	48.0	38.5	27.0	27.0	15570
Rbar	15.0	17.0	13.2	13.2	13.2	14.3	13.2	18.7	9.0	14.5	11.5	11.5	12.2	11.5	18.7	18.7	20.5	9.0	9.0	9.0	5.5	16.0	12.8	9.0	9.0	13
sumRi,squared	41134		N =	25		k =	3																			
W =	0.3982																									
Chi^2,calc =	28.671		v =	24																						
tabled value =	23%																									
	There is a 0.232763047066956 chance that the agreement among the raters was by random chance.																									

Item 37																										Sum(Tj)
Item 37 Observer 1	13.5	13.5	13.5	13.5	13.5	2.0	13.5	24.0	13.5	13.5	13.5	13.5	2.0	13.5	13.5	13.5	13.5	13.5	13.5	13.5	2.0	13.5	25.0	13.5	13.5	8004
Item 37 Observer 2	4.0	11.0	16.0	16.0	22.0	24.5	16.0	22.0	11.0	4.0	19.5	11.0	19.5	16.0	22.0	16.0	24.5	4.0	4.0	8.5	4.0	13.0	8.5	4.0	4.0	522
Item 37 Observer 3	8.0	8.0	8.0	8.0	21.5	8.0	8.0	25.0	18.5	18.5	16.5	8.0	8.0	8.0	24.0	21.5	21.5	8.0	8.0	8.0	8.0	8.0	21.5	16.5	8.0	3432
R	25.5	32.5	37.5	37.5	57.0	34.5	37.5	71.0	43.0	36.0	49.5	32.5	29.5	37.5	59.5	51.0	59.5	25.5	25.5	30.0	14.0	34.5	55.0	34.0	25.5	11958
Rbar	8.5	10.8	12.5	12.5	19.0	11.5	12.5	23.7	14.3	12.0	16.5	10.8	9.8	12.5	19.8	17.0	19.8	8.5	8.5	10.0	4.7	11.5	18.3	11.3	8.5	13
sumRi,squared	42433		N =	25		k =	3																			
W =	0.5061																									
Chi^2,calc =	36.436		v =	24																						
tabled value =	5%																									
	There is a 0.0497631882886097 chance that the agreement among the raters was by random chance.																									

Item 43																										Sum(Tj)
Item 43 Observer 1	12.0	20.0	12.0	12.0	12.0	12.0	12.0	12.0	20.0	12.0	12.0	12.0	20.0	24.0	12.0	3.0	3.0	12.0	3.0	12.0	3.0	20.0	25.0	24.0	3.0	2370
Item 43 Observer 2	4.0	19.0	19.0	16.0	16.0	22.0	10.0	25.0	10.0	4.0	10.0	10.0	19.0	16.0	24.0	23.0	16.0	10.0	4.0	10.0	4.0	22.0	4.0	10.0	4.0	636
Item 43 Observer 3	8.0	16.0	16.0	22.0	22.0	8.0	22.0	25.0	16.0	16.0	16.0	16.0	8.0	16.0	22.0	16.0	16.0	8.0	8.0	3.0	3.0	3.0	16.0	8.0	1.0	1284
R	24.0	55.0	47.0	50.0	50.0	42.0	44.0	62.0	46.0	32.0	38.0	38.0	47.0	56.0	58.0	42.0	35.0	30.0	15.0	25.0	10.0	45.0	45.0	42.0	8.0	4290
Rbar	8.0	18.3	15.7	16.7	16.7	14.0	14.7	20.7	15.3	10.7	12.7	12.7	15.7	18.7	19.3	14.0	11.7	10.0	5.0	8.3	3.3	15.0	15.0	14.0	2.7	13
sumRi,squared	43808		N =	25		k =	3																			
W =	0.5442																									
Chi^2,calc =	39.179		v =	24																						
tabled value =	3%																									
	There is a 0.0261501326116328 chance that the agreement among the raters was by random chance.																									

Appendix E – Item Normality Tests

Descriptive Statistics Report: Item 1

Page/Date/Time	1 7/29/2003 1:57:51 PM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

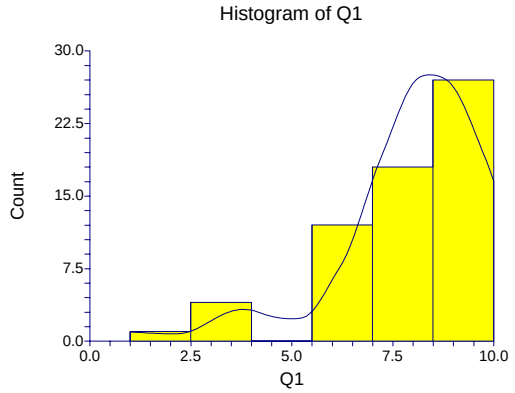
Summary Section of Q1

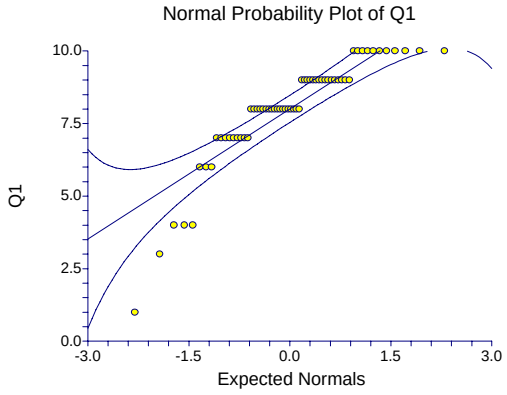
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.983871	1.824172	0.2316701	1	10	9

Normality Test Section of Q1

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.8347189	0.000001			Reject normality
Anderson-Darling	3.005969	0.000000			Reject normality
Martinez-Iglewicz	1.395477		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2132048		0.103	0.112	Reject normality
D'Agostino Skewness	-4.3045	0.000017	1.645	1.960	Reject normality
D'Agostino Kurtosis	3.0814	0.002060	1.645	1.960	Reject normality
D'Agostino Omnibus	28.0239	0.000001	4.605	5.991	Reject normality

Plots Section of Q1





Stem-Leaf Plot Section of Q1

Depth	Stem	Leaves
Low		10,30
5	4*	000
5	.	
5	5*	
5	.	
8	6*	000
8	.	
17	7*	000000000
17	.	
(18)	8*	000000000000000000
27	.	
27	9*	00000000000000000
11	.	
11	10*	00000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 3

Page/Date/Time 2 7/29/2003 1:57:51 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

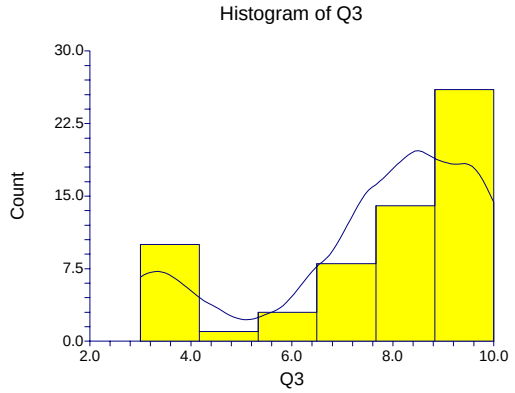
Summary Section of Q3

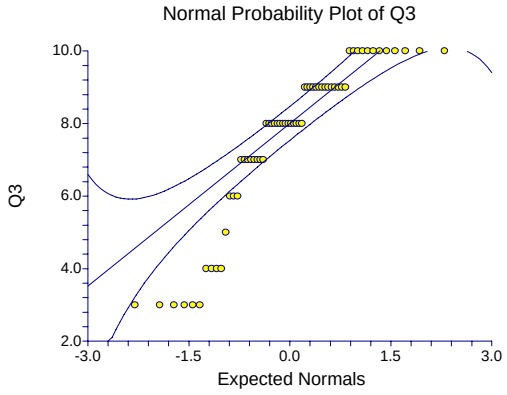
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.596774	2.198805	0.2792485	3	10	7

Normality Test Section of Q3

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.8530619	0.000003			Reject normality
Anderson-Darling	3.125976	0.000000			Reject normality
Martinez-Iglewicz	1.109205		1.077884	1.120109	Can't reject normality
Kolmogorov-Smirnov	0.2017839		0.103	0.112	Reject normality
D'Agostino Skewness	-2.8922	0.003826	1.645	1.960	Reject normality
D'Agostino Kurtosis	-0.0922	0.926512	1.645	1.960	Can't reject normality
D'Agostino Omnibus	8.3733	0.015197	4.605	5.991	Reject normality

Plots Section of Q3





Stem-Leaf Plot Section of Q3

Depth	Stem	Leaves
Low		30,30,30,30,30,30
10	4*	0000
10	.	
11	5*	0
11	.	
14	6*	000
14	.	
22	7*	00000000
22	.	
(14)	8*	00000000000000
26	.	
26	9*	00000000000000
12	.	
12	10*	000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 4

Page/Date/Time 3 7/29/2003 1:57:52 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

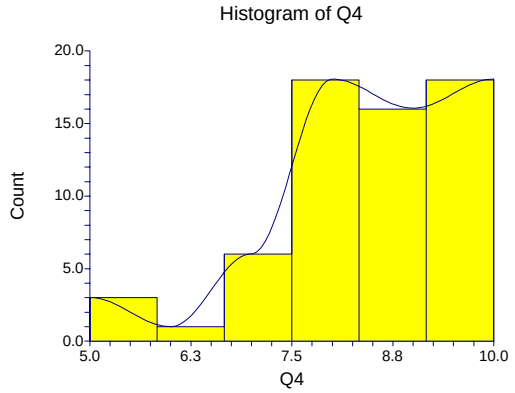
Summary Section of Q4

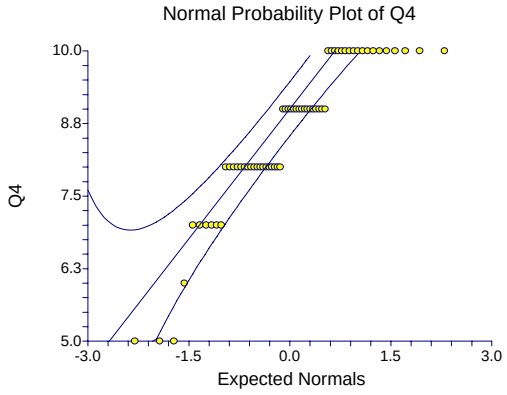
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.564516	1.313498	0.1668144	5	10	5

Normality Test Section of Q4

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.86365	0.000006			Reject normality
Anderson-Darling	2.636913	0.000001			Reject normality
Martinez-Iglewicz	1.139874		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.1621417		0.103	0.112	Reject normality
D'Agostino Skewness	-2.8498	0.004374	1.645	1.960	Reject normality
D'Agostino Kurtosis	1.2679	0.204819	1.645	1.960	Can't reject normality
D'Agostino Omnibus	9.7292	0.007715	4.605	5.991	Reject normality

Plots Section of Q4





Stem-Leaf Plot Section of Q4

Depth	Stem	Leaves
3	5* |	000
3	. |	
4	6* |	0
4	. |	
10	7* |	000000
10	. |	
28	8* |	0000000000000000000
28	. |	
(16)	9* |	0000000000000000
18	. |	
18	10* |	00000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 5

Page/Date/Time 4 7/29/2003 1:57:52 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

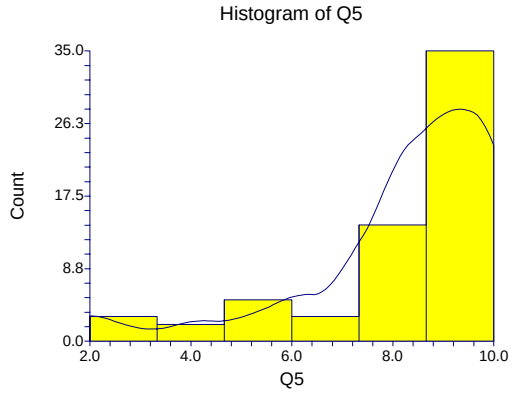
Summary Section of Q5

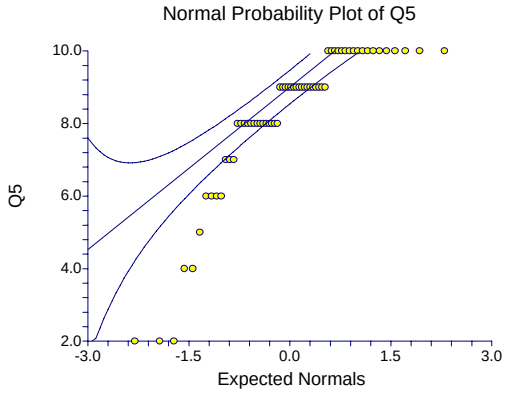
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.209678	2.041673	0.2592928	2	10	8

Normality Test Section of Q5

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.7835066	0.000000			Reject normality
Anderson-Darling	4.288754	0.000000			Reject normality
Martinez-Iglewicz	2.04947		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2332946		0.103	0.112	Reject normality
D'Agostino Skewness	-4.3943	0.000011	1.645	1.960	Reject normality
D'Agostino Kurtosis	2.6756	0.007460	1.645	1.960	Reject normality
D'Agostino Omnibus	26.4687	0.000002	4.605	5.991	Reject normality

Plots Section of Q5





Stem-Leaf Plot Section of Q5

Depth	Stem	Leaves
Low		20,20,20,40,40
6	5*	0
6	.	
10	6*	0000
10	.	
13	7*	000
13	.	
27	8*	00000000000000
27	.	
(17)	9*	00000000000000000
18	.	
18	10*	000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 8

Page/Date/Time 5 7/29/2003 1:57:52 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

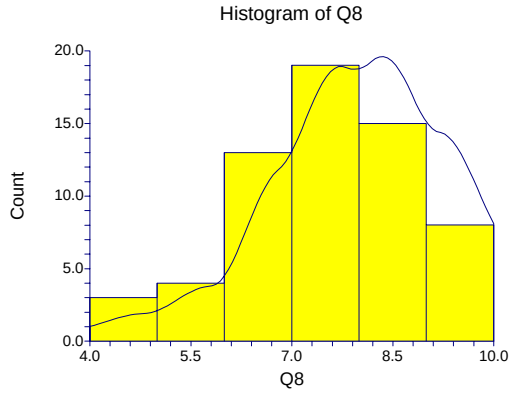
Summary Section of Q8

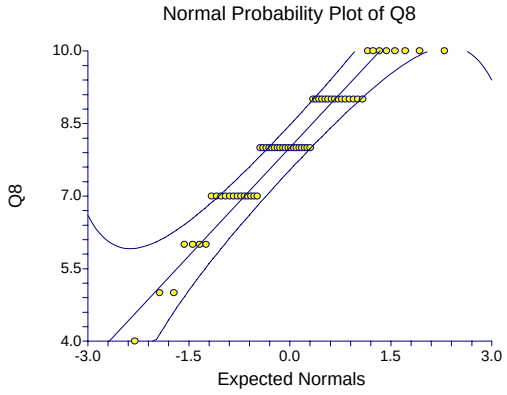
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8	1.342862	0.1705437	4	10	6

Normality Test Section of Q8

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.9255394	0.001049			Reject normality
Anderson-Darling	1.68537	0.000255			Reject normality
Martinez-Iglewicz	1.025858		1.077884	1.120109	Can't reject normality
Kolmogorov-Smirnov	0.1612903		0.103	0.112	Reject normality
D'Agostino Skewness	-1.9170	0.055236	1.645	1.960	Can't reject normality
D'Agostino Kurtosis	0.8282	0.407556	1.645	1.960	Can't reject normality
D'Agostino Omnibus	4.3609	0.112993	4.605	5.991	Can't reject normality

Plots Section of Q8





Stem-Leaf Plot Section of Q8

Depth	Stem	Leaves
1	4*	0
1	.	
3	5*	00
3	.	
7	6*	0000
7	.	
20	7*	0000000000000
20	.	
(19)	8*	00000000000000000000
23	.	
23	9*	000000000000000
8	.	
8	10*	00000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 12

Page/Date/Time 6 7/29/2003 1:57:52 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

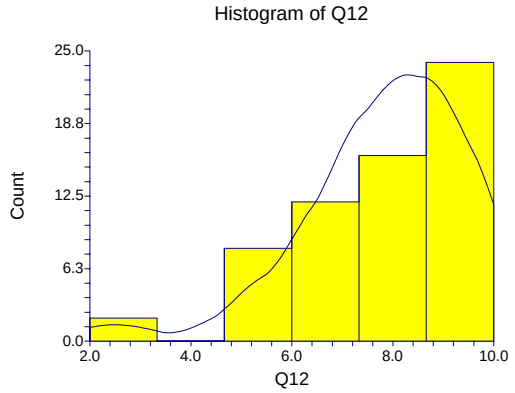
Summary Section of Q12

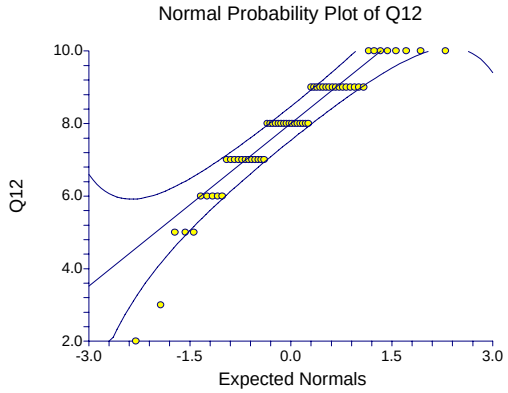
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.83871	1.651562	0.2097486	2	10	8

Normality Test Section of Q12

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.892743	0.000056			Reject normality
Anderson-Darling	1.879632	0.000085			Reject normality
Martinez-Iglewicz	1.233258		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.1679308		0.103	0.112	Reject normality
D'Agostino Skewness	-3.4329	0.000597	1.645	1.960	Reject normality
D'Agostino Kurtosis	2.4204	0.015504	1.645	1.960	Reject normality
D'Agostino Omnibus	17.6432	0.000148	4.605	5.991	Reject normality

Plots Section of Q12





Stem-Leaf Plot Section of Q12

Depth	Stem	Leaves
Low		20,30
5	5*	000
5	.	
10	6*	00000
10	.	
22	7*	000000000000
22	.	
(16)	8*	0000000000000000
24	.	
24	9*	0000000000000000
8	.	
8	10*	00000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 13

Page/Date/Time 7 7/29/2003 1:57:52 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

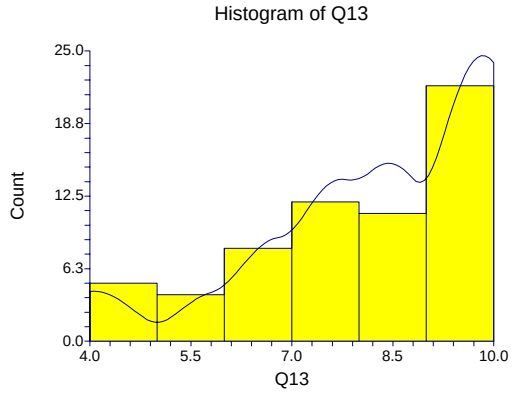
Summary Section of Q13

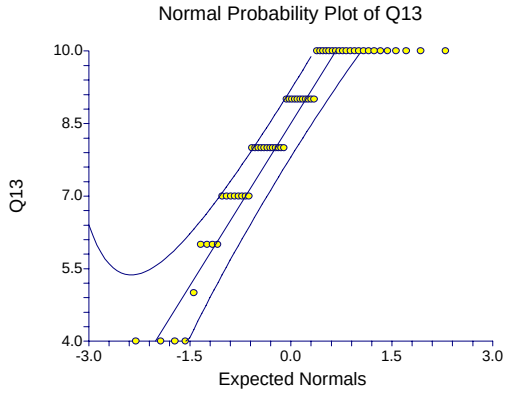
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.32258	1.762768	0.2238717	4	10	6

Normality Test Section of Q13

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.8454422	0.000002			Reject normality
Anderson-Darling	3.041413	0.000000			Reject normality
Martinez-Iglewicz	1.218736		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.1706543		0.103	0.112	Reject normality
D'Agostino Skewness	-3.0086	0.002625	1.645	1.960	Reject normality
D'Agostino Kurtosis	0.6725	0.501256	1.645	1.960	Can't reject normality
D'Agostino Omnibus	9.5039	0.008635	4.605	5.991	Reject normality

Plots Section of Q13





Stem-Leaf Plot Section of Q13

Depth	Stem	Leaves
4	4*	0000
4	.	
5	5*	0
5	.	
9	6*	0000
9	.	
17	7*	00000000
17	.	
29	8*	000000000000
29	.	
(11)	9*	00000000000
22	.	
22	10*	00000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 16

Page/Date/Time 8 7/29/2003 1:57:53 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

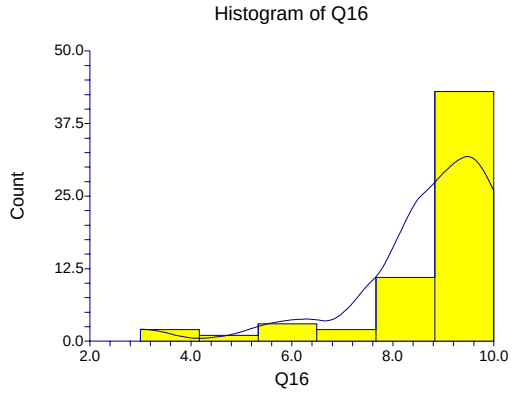
Summary Section of Q16

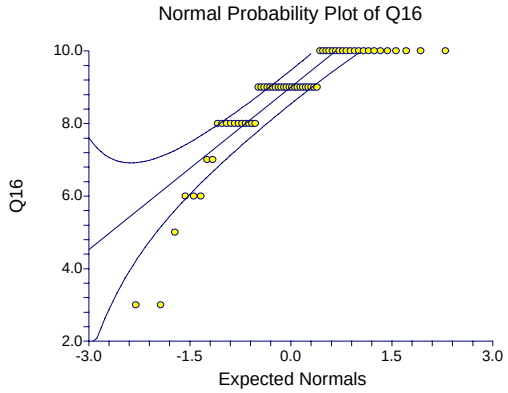
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.693548	1.56382	0.1986053	3	10	7

Normality Test Section of Q16

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.7535166	0.000000			Reject normality
Anderson-Darling	4.707731	0.000000			Reject normality
Martinez-Iglewicz	1.704825		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2551		0.103	0.112	Reject normality
D'Agostino Skewness	-4.8867	0.000001	1.645	1.960	Reject normality
D'Agostino Kurtosis	3.5019	0.000462	1.645	1.960	Reject normality
D'Agostino Omnibus	36.1427	0.000000	4.605	5.991	Reject normality

Plots Section of Q16





Stem-Leaf Plot Section of Q16

Depth	Stem	Leaves
Low		30,30
3	5*	0
3	.	
6	6*	000
6	.	
8	7*	00
8	.	
19	8*	00000000000
19	.	
(22)	9*	0000000000000000000000
21	.	
21	10*	00000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 19

Page/Date/Time 9 7/29/2003 1:57:53 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

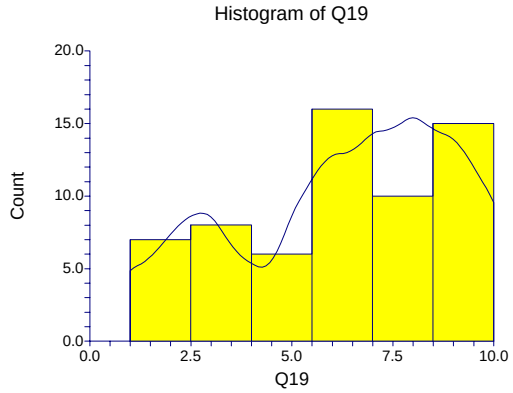
Summary Section of Q19

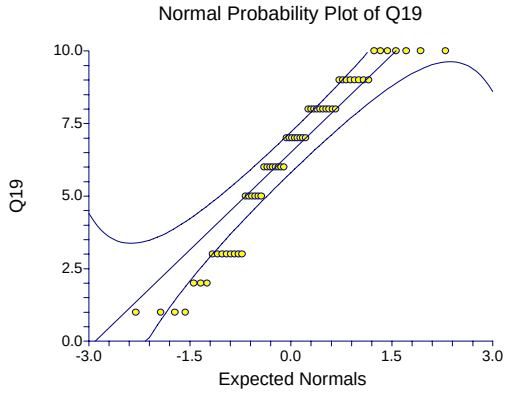
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	6.290323	2.724432	0.3460032	1	10	9

Normality Test Section of Q19

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.9238653	0.000892			Reject normality
Anderson-Darling	1.472957	0.000846			Reject normality
Martinez-Iglewicz	0.9740819		1.077884	1.120109	Can't reject normality
Kolmogorov-Smirnov	0.128356		0.103	0.112	Reject normality
D'Agostino Skewness	-1.5523	0.120587	1.645	1.960	Can't reject normality
D'Agostino Kurtosis	-2.1716	0.029888	1.645	1.960	Reject normality
D'Agostino Omnibus	7.1254	0.028362	4.605	5.991	Reject normality

Plots Section of Q19





Stem-Leaf Plot Section of Q19

Depth	Stem	Leaves
4	1	0000
7	2	000
15	3	00000000
15	4	
21	5	000000
29	6	00000000
(8)	7	00000000
25	8	0000000000
15	9	00000000
7	10	0000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 20

Page/Date/Time 10 7/29/2003 1:57:53 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

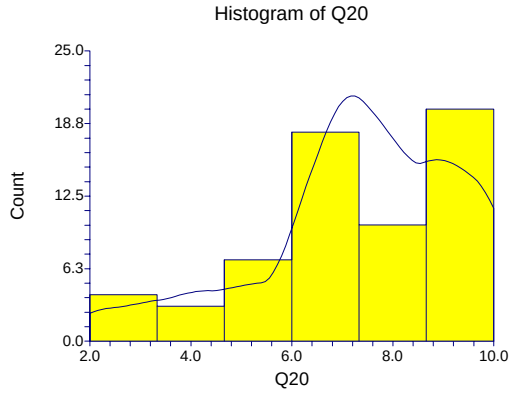
Summary Section of Q20

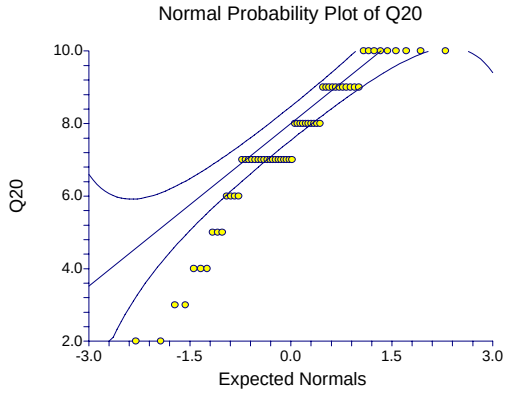
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.354839	2.033303	0.2582297	2	10	8

Normality Test Section of Q20

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.9084474	0.000214			Reject normality
Anderson-Darling	1.800761	0.000133			Reject normality
Martinez-Iglewicz	0.9211341		1.077884	1.120109	Can't reject normality
Kolmogorov-Smirnov	0.1887955		0.103	0.112	Reject normality
D'Agostino Skewness	-2.6220	0.008740	1.645	1.960	Reject normality
D'Agostino Kurtosis	0.8558	0.392088	1.645	1.960	Can't reject normality
D'Agostino Omnibus	7.6076	0.022286	4.605	5.991	Reject normality

Plots Section of Q20





Stem-Leaf Plot Section of Q20

Depth	Stem	Leaves
Low		20,20,30,30
7	4*	000
7	.	
10	5*	000
10	.	
14	6*	0000
14	.	
(18)	7*	000000000000000000
30	.	
30	8*	0000000000
20	.	
20	9*	00000000000
9	.	
9	10*	000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 28

Page/Date/Time 11 7/29/2003 1:57:53 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

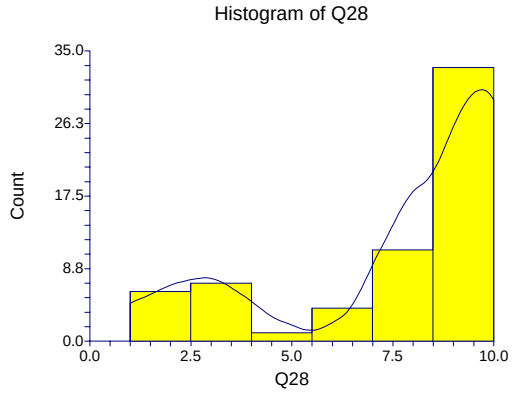
Summary Section of Q28

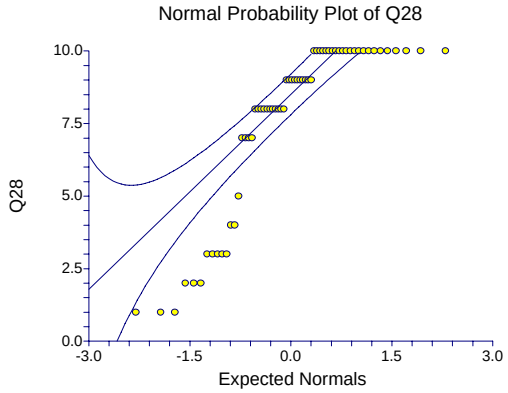
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.629032	2.915385	0.3702543	1	10	9

Normality Test Section of Q28

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.7714096	0.000000			Reject normality
Anderson-Darling	5.612566	0.000000			Reject normality
Martinez-Iglewicz	2.569584		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2441751		0.103	0.112	Reject normality
D'Agostino Skewness	-3.3781	0.000730	1.645	1.960	Reject normality
D'Agostino Kurtosis	0.0553	0.955928	1.645	1.960	Can't reject normality
D'Agostino Omnibus	11.4147	0.003321	4.605	5.991	Reject normality

Plots Section of Q28





Stem-Leaf Plot Section of Q28

Depth	Stem	Leaves
Low		10,10,10,20,20,20
11	3*	00000
11	.	
13	4*	00
13	.	
14	5*	0
14	.	
14	6*	
14	.	
18	7*	0000
18	.	
29	8*	00000000000
29	.	
(10)	9*	0000000000
23	.	
23	10*	00000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 29

Page/Date/Time 12 7/29/2003 1:57:54 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

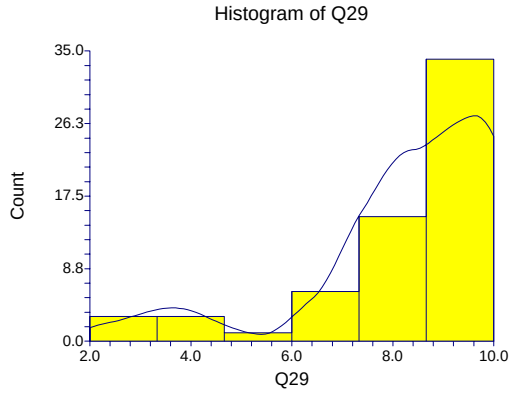
Summary Section of Q29

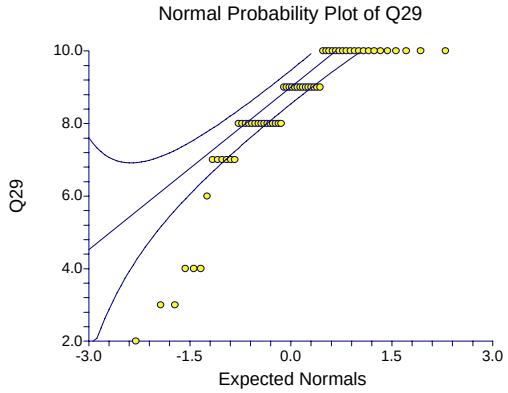
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.290322	1.953451	0.2480885	2	10	8

Normality Test Section of Q29

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.7880917	0.000000			Reject normality
Anderson-Darling	4.200835	0.000000			Reject normality
Martinez-Iglewicz	1.908524		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2151202		0.103	0.112	Reject normality
D'Agostino Skewness	-4.2677	0.000020	1.645	1.960	Reject normality
D'Agostino Kurtosis	2.5097	0.012082	1.645	1.960	Reject normality
D'Agostino Omnibus	24.5125	0.000005	4.605	5.991	Reject normality

Plots Section of Q29





Stem-Leaf Plot Section of Q29

Depth	Stem	Leaves
Low		20,30,30,40,40,40
7	6*	0
7	.	
13	7*	000000
13	.	
28	8*	0000000000000000
28	.	
(14)	9*	00000000000000
20	.	
20	10*	00000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 35

Page/Date/Time	13 7/29/2003 1:57:54 PM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

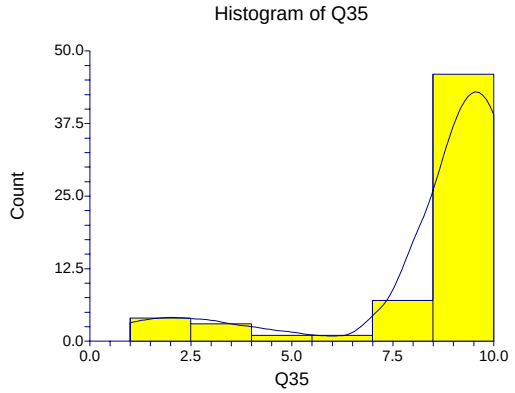
Summary Section of Q35

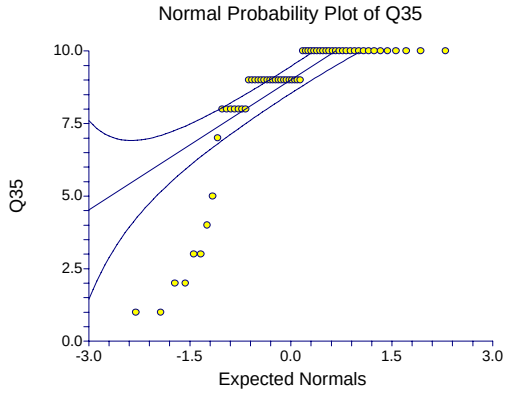
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.467742	2.427423	0.3082831	1	10	9

Normality Test Section of Q35

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.6420614	0.000000			Reject normality
Anderson-Darling	8.746274	0.000000			Reject normality
Martinez-Iglewicz	3.777254		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.3125861		0.103	0.112	Reject normality
D'Agostino Skewness	-5.0209	0.000001	1.645	1.960	Reject normality
D'Agostino Kurtosis	2.9992	0.002707	1.645	1.960	Reject normality
D'Agostino Omnibus	34.2045	0.000000	4.605	5.991	Reject normality

Plots Section of Q35





Stem-Leaf Plot Section of Q35

Depth	Stem	Leaves
Low		10,10,20,20,30,30,40
8	5*	0
8	.	
8	6*	
8	.	
9	7*	0
9	.	
16	8*	0000000
16	.	
(19)	9*	0000000000000000000
27	.	
27	10*	0000000000000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 37

Page/Date/Time 14 7/29/2003 1:57:54 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

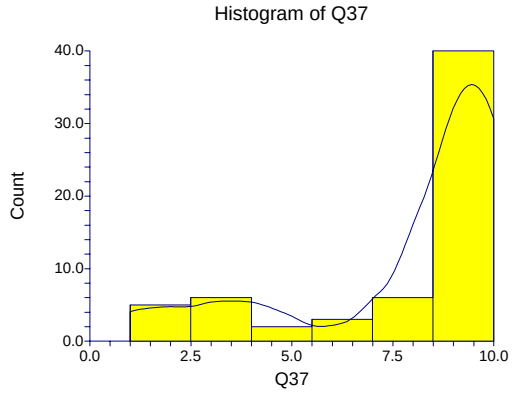
Summary Section of Q37

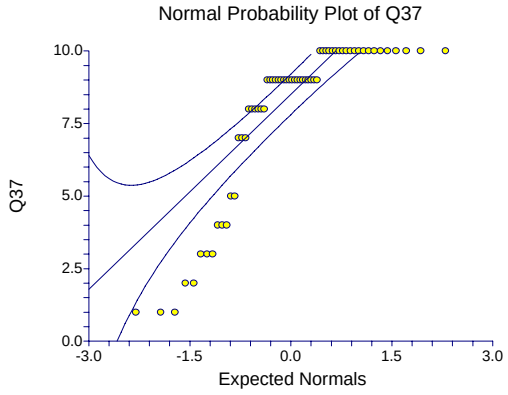
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	7.870968	2.760873	0.3506313	1	10	9

Normality Test Section of Q37

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.7419664	0.000000			Reject normality
Anderson-Darling	6.52459	0.000000			Reject normality
Martinez-Iglewicz	2.852513		1.077884	1.120109	Reject normality
Kolmogorov-Smirnov	0.2877405		0.103	0.112	Reject normality
D'Agostino Skewness	-3.8631	0.000112	1.645	1.960	Reject normality
D'Agostino Kurtosis	1.0805	0.279919	1.645	1.960	Can't reject normality
D'Agostino Omnibus	16.0913	0.000320	4.605	5.991	Reject normality

Plots Section of Q37





Stem-Leaf Plot Section of Q37

Depth	Stem	Leaves
Low		10,10,10,20,20
8	3*	000
8	.	
11	4*	000
11	.	
13	5*	00
13	.	
13	6*	
13	.	
16	7*	000
16	.	
22	8*	000000
22	.	
(19)	9*	0000000000000000000
21	.	
21	10*	00000000000000000000

Unit = .1 Example: 1 |2 Represents 1.2

Descriptive Statistics Report: Item 43

Page/Date/Time 15 7/29/2003 1:57:54 PM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

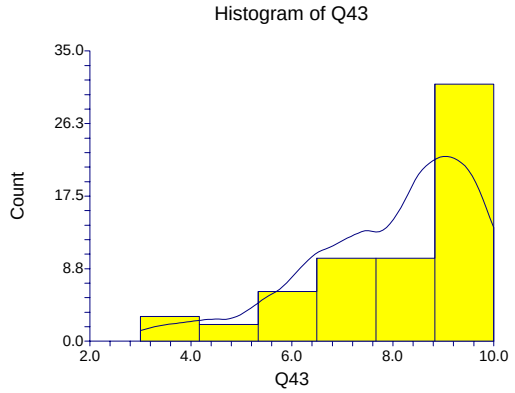
Summary Section of Q43

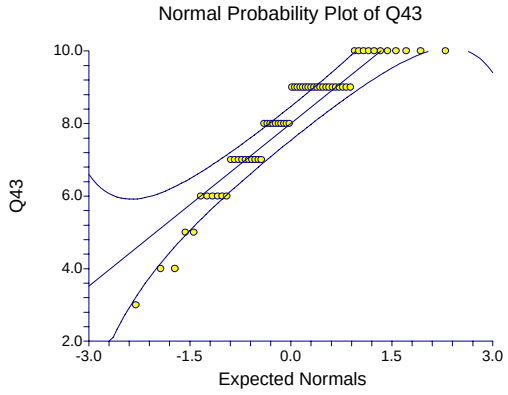
Count	Mean	Standard Deviation	Standard Error	Minimum	Maximum	Range
62	8.016129	1.683984	0.2138661	3	10	7

Normality Test Section of Q43

Test Name	Test Value	Prob Level	10% Critical Value	5% Critical Value	Decision (5%)
Shapiro-Wilk W	0.8916646	0.000051			Reject normality
Anderson-Darling	2.28449	0.000009			Reject normality
Martinez-Iglewicz	1.090605		1.077884	1.120109	Can't reject normality
Kolmogorov-Smirnov	0.2043456		0.103	0.112	Reject normality
D'Agostino Skewness	-2.8786	0.003994	1.645	1.960	Reject normality
D'Agostino Kurtosis	0.9485	0.342860	1.645	1.960	Can't reject normality
D'Agostino Omnibus	9.1861	0.010122	4.605	5.991	Reject normality

Plots Section of Q43





Stem-Leaf Plot Section of Q43

Depth	Stem	Leaves
Low		30
3	4*	00
3	.	
5	5*	00
5	.	
11	6*	000000
11	.	
21	7*	0000000000
21	.	
31	8*	0000000000
31	.	
31	9*	00000000000000000000
11	.	
11	10*	00000000000

Unit = .1 Example: 1 |2 Represents 1.2

Appendix F – Inter-rater Reliability ANOVA Details

The analyses performed in this appendix were executed on NCSS 2002. The output shown in this Appendix Js a modified version of the output generated by this software package.

Item 1

Analysis of Variance Report: Item 1
Page/Date/Time 1 6/30/2003 10:54:29 AM
Database C:\Documents and Settings\Br ... rrater Reliability\Item 1.S0
Response Item_1

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

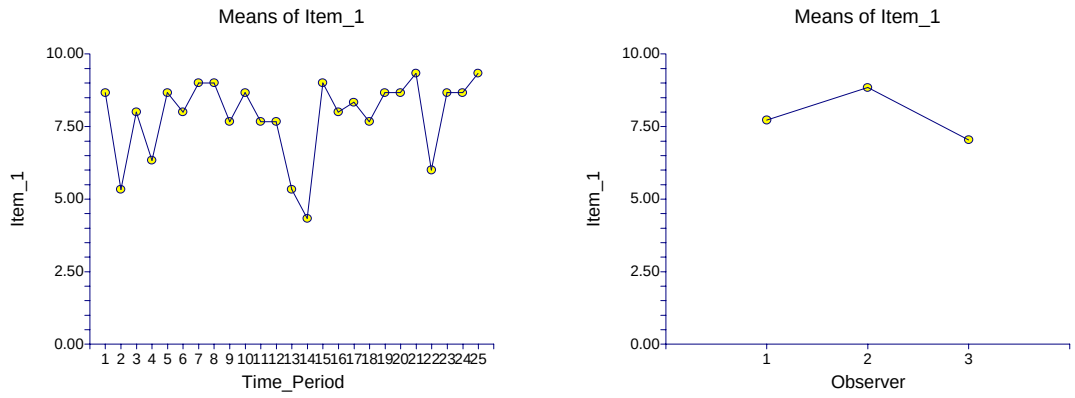
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	132.6667	5.527778	2.53	0.003035*	0.997375
B: Observer	2	41.30667	20.65333	9.47	0.000342*	
r: Residual	48	104.6933	2.181111			
Total (Adjusted)	74	278.6667				
Total	75					

* Term significant at alpha = 0.10

Plots Section



Item 3

Analysis of Variance Report: Item 3
Page/Date/Time 1 6/30/2003 10:57:24 AM
Database C:\Documents and Settings\Br ... rrater Reliability\Item 3.S0
Response Item_3

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

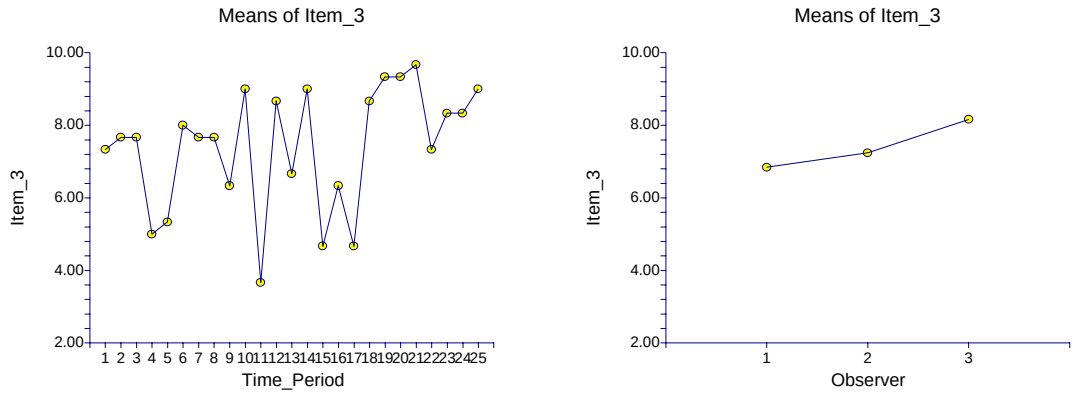
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	202.8533	8.452222	2.81	0.001150*	0.999009
B: Observer	2	22.90667	11.45333	3.81	0.029212*	
r: Residual	48	144.4267	3.008889			
Total (Adjusted)	74	370.1867				
Total	75					

* Term significant at alpha = 0.10

Plots Section



Item 4

Analysis of Variance Report: Item 4
Page/Date/Time 1 6/30/2003 10:59:11 AM
Database C:\Documents and Settings\Br ... rrater Reliability\Item 4.S0
Response Item_4

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

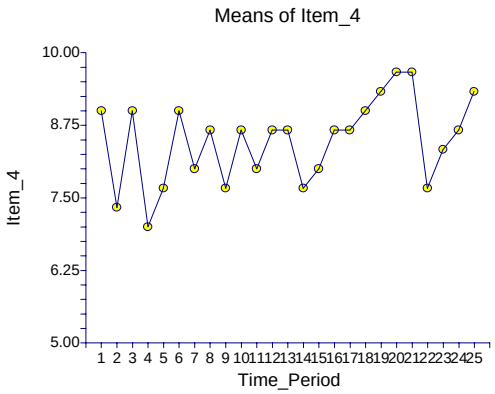
Note: Expected Mean Squares are for the balanced cell-frequency case.

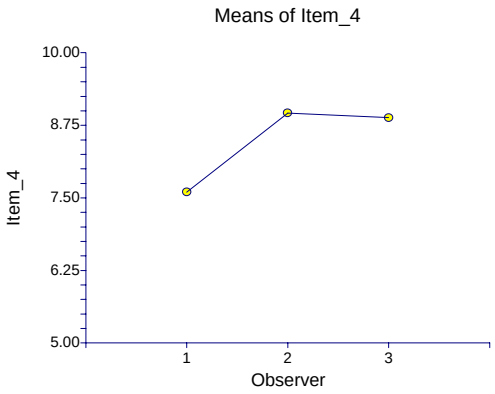
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	37.38667	1.557778	1.20	0.287479	0.860342
B: Observer	2	29.12	14.56	11.23	0.000100*	
r: Residual	48	62.21333	1.296111			
Total (Adjusted)	74	128.72				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 5

Analysis of Variance Report: Item 5
Page/Date/Time 1 6/30/2003 11:01:06 AM
Database C:\Documents and Settings\Br ... rrater Reliability\Item 5.S0
Response Item_5

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

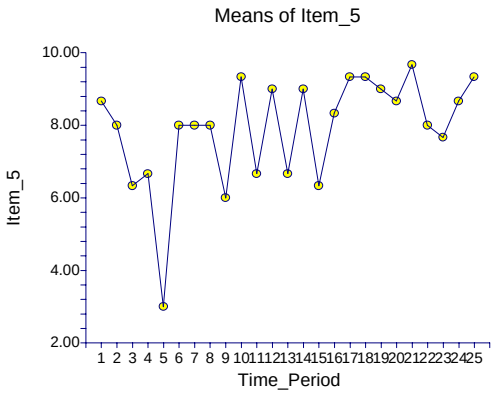
Note: Expected Mean Squares are for the balanced cell-frequency case.

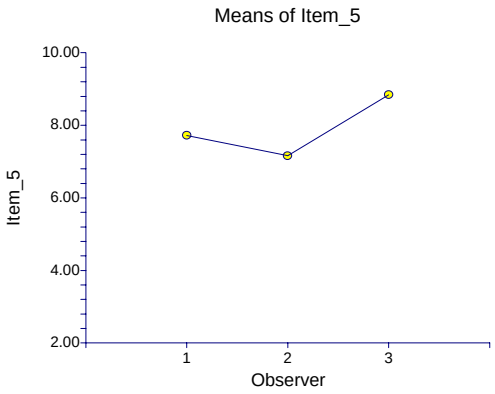
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	162.3467	6.764444	1.81	0.040095*	0.972745
B: Observer	2	36.58667	18.29333	4.89	0.011631*	
r: Residual	48	179.4133	3.737778			
Total (Adjusted)	74	378.3467				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 8

Analysis of Variance Report: Item 8
Page/Date/Time 1 6/30/2003 10:47:45 AM
Database C:\Documents and Settings\Br ... rrater Reliability\Item 8.S0
Response Item_8

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

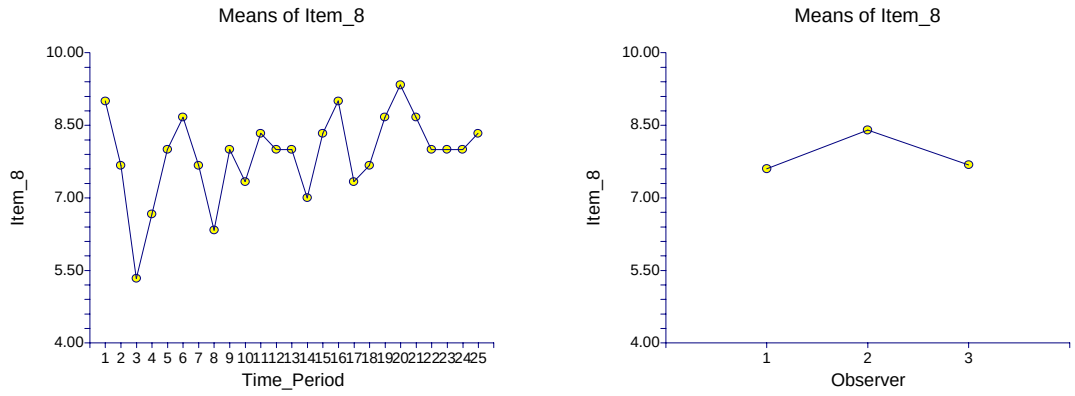
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	57.14667	2.381111	1.16	0.320505	0.741065
B: Observer	2	9.706667	4.853333	2.37	0.104330	
r: Residual	48	98.29333	2.047778			
Total (Adjusted)	74	165.1467				
Total	75					

* Term significant at alpha = 0.05

Plots Section



Item 12

Analysis of Variance Report: Item 12
Page/Date/Time 1 6/30/2003 11:05:39 AM
Database C:\Documents and Settings\Br ... rater Reliability\Item 12.S0
Response Item_12

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

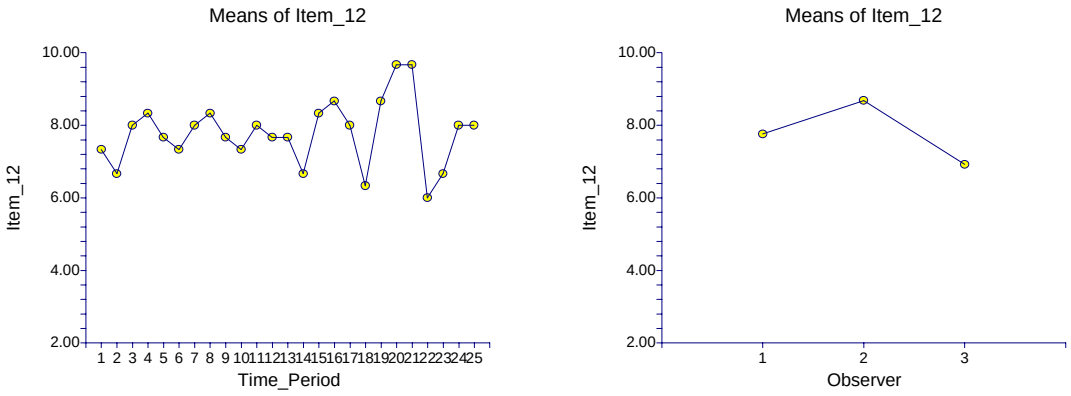
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	58.58667	2.441111	1.41	0.154963	0.916343
B: Observer	2	38.74667	19.37333	11.17	0.000104*	
r: Residual	48	83.25333	1.734444			
Total (Adjusted)	74	180.5867				
Total	75					

* Term significant at alpha = 0.10

Plots Section



Item 13

Analysis of Variance Report: Item 13
Page/Date/Time 1 6/30/2003 11:17:52 AM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_13

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

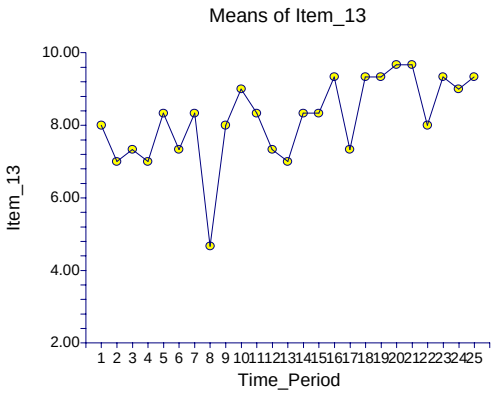
Note: Expected Mean Squares are for the balanced cell-frequency case.

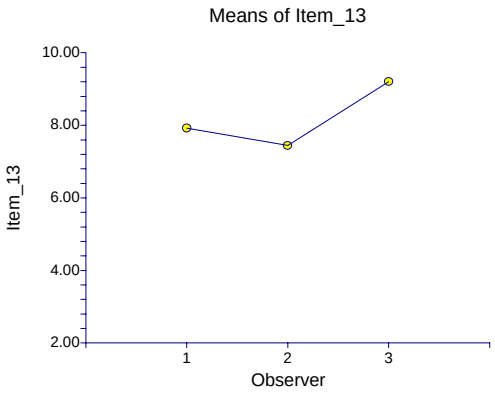
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	96.05334	4.002222	1.33	0.194438	0.899167
B: Observer	2	41.38667	20.69333	6.90	0.002322*	
r: Residual	48	143.9467	2.998889			
Total (Adjusted)	74	281.3867				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 16

Analysis of Variance Report: Item 16
Page/Date/Time 1 6/30/2003 11:20:38 AM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_16

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

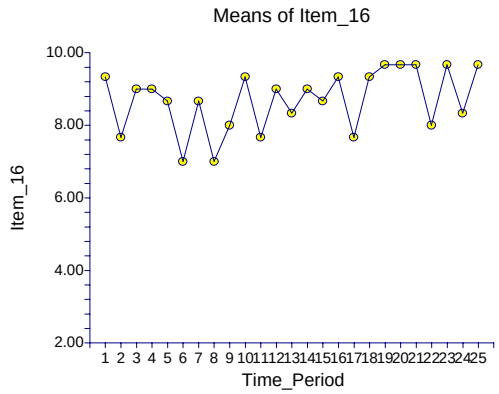
Note: Expected Mean Squares are for the balanced cell-frequency case.

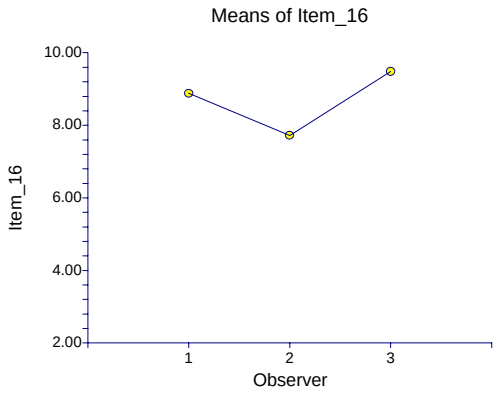
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	50.61333	2.108889	1.11	0.370454	0.826453
B: Observer	2	40.02667	20.01333	10.52	0.000163*	
r: Residual	48	91.30666	1.902222			
Total (Adjusted)	74	181.9467				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 19

Analysis of Variance Report: Item 19
Page/Date/Time 1 6/30/2003 11:25:50 AM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_19

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

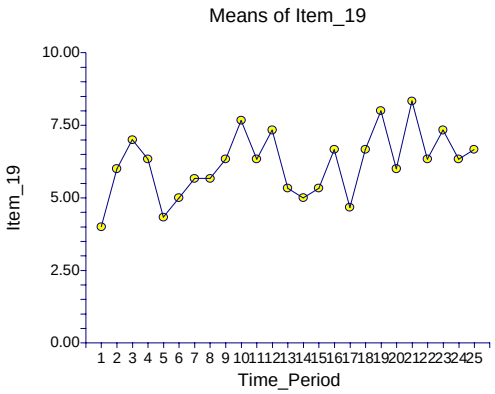
Note: Expected Mean Squares are for the balanced cell-frequency case.

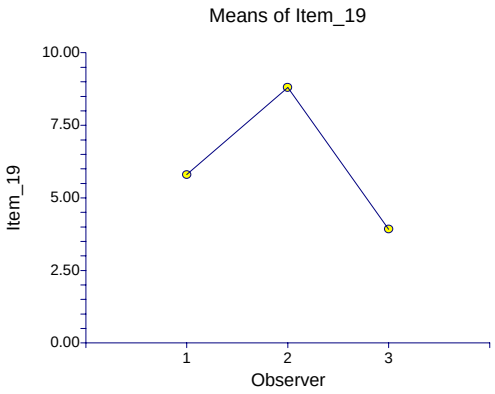
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	88.74667	3.697778	1.14	0.336891	0.840149
B: Observer	2	302.9067	151.4533	46.87	0.000000*	
r: Residual	48	155.0933	3.231111			
Total (Adjusted)	74	546.7466				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 20

Analysis of Variance Report: Item 20
Page/Date/Time 1 6/30/2003 11:29:01 AM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_20

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

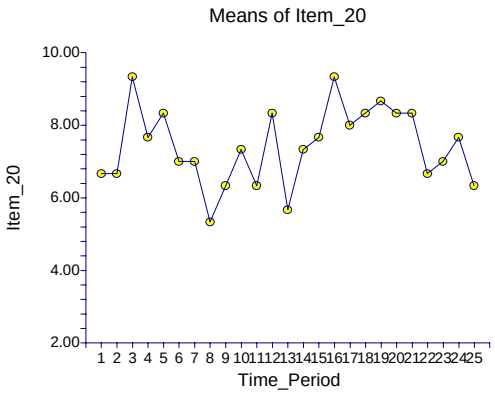
Note: Expected Mean Squares are for the balanced cell-frequency case.

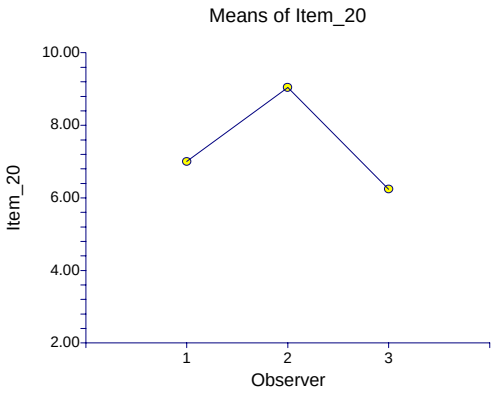
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	80.34666	3.347778	1.24	0.254801	0.873805
B: Observer	2	104.8267	52.41333	19.48	0.000001*	
r: Residual	48	129.1733	2.691111			
Total (Adjusted)	74	314.3467				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 28

Analysis of Variance Report: Item 28
Page/Date/Time 1 6/30/2003 1:08:13 PM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_28

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

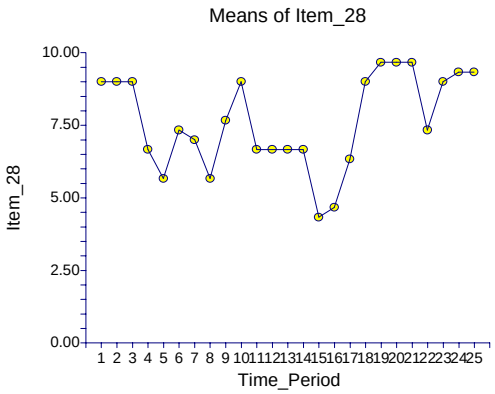
Note: Expected Mean Squares are for the balanced cell-frequency case.

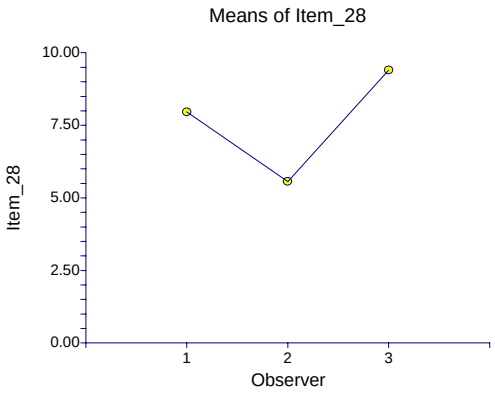
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	191.28	7.97	1.74	0.051082*	0.966571
B: Observer	2	188.16	94.08	20.54	0.000000*	
r: Residual	48	219.84	4.58			
Total (Adjusted)	74	599.28				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 29

Analysis of Variance Report: Item 29
Page/Date/Time 1 6/30/2003 1:16:11 PM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_29

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

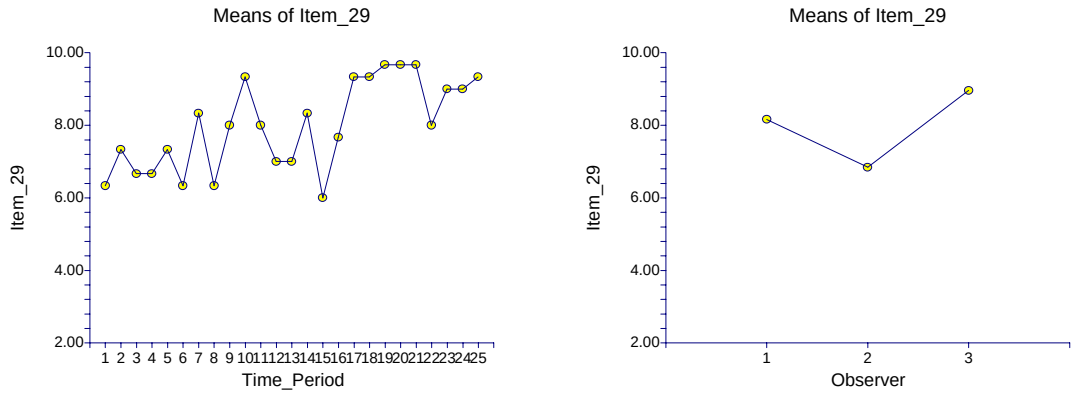
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	109.6533	4.568889	1.37	0.174051	0.907961
B: Observer	2	57.30667	28.65333	8.59	0.000645*	
r: Residual	48	160.0267	3.333889			
Total (Adjusted)	74	326.9867				
Total	75					

* Term significant at alpha = 0.10

Plots Section



Item 35

Analysis of Variance Report: Item 35
Page/Date/Time 1 6/30/2003 1:53:45 PM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_35

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

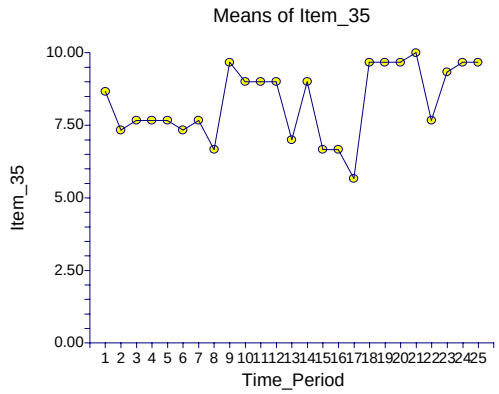
Note: Expected Mean Squares are for the balanced cell-frequency case.

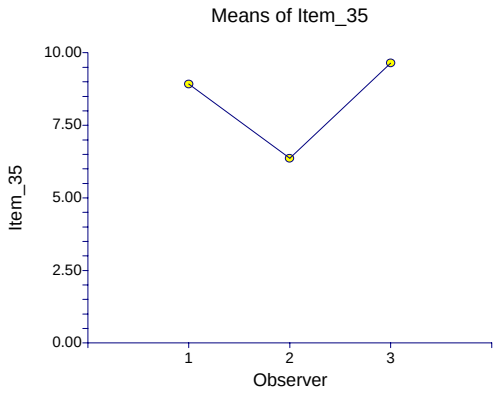
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	113.28	4.72	1.23	0.264752	0.869691
B: Observer	2	148.5867	74.29333	19.37	0.000001*	
r: Residual	48	184.08	3.835			
Total (Adjusted)	74	445.9467				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Item 37

Analysis of Variance Report: Item 37
Page/Date/Time 1 6/30/2003 1:56:00 PM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_37

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

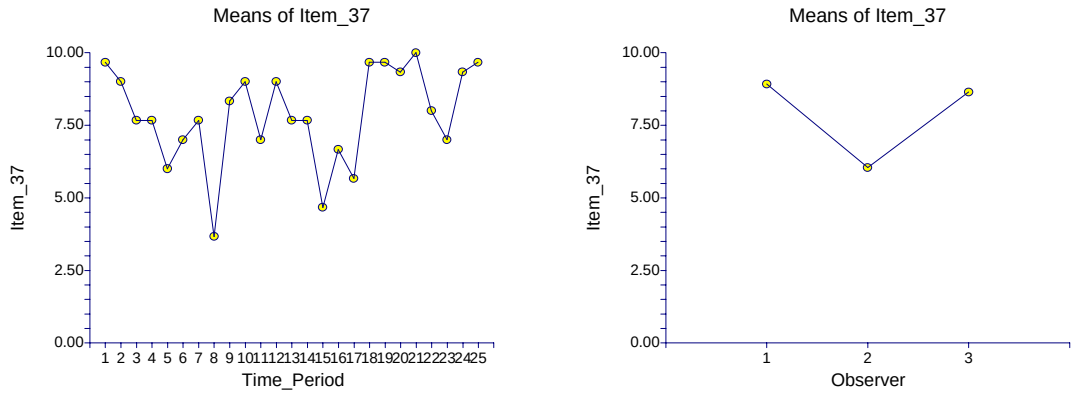
Note: Expected Mean Squares are for the balanced cell-frequency case.

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	198	8.25	1.72	0.055236*	0.964309
B: Observer	2	126.1067	63.05333	13.13	0.000028*	
r: Residual	48	230.56	4.803333			
Total (Adjusted)	74	554.6667				
Total	75					

* Term significant at alpha = 0.10

Plots Section



Item 43

Analysis of Variance Report: Item 43
Page/Date/Time 1 6/30/2003 2:01:16 PM
Database C:\Documents and Settings\Br ... ter Reliability\All Items.S0
Response Item_43

Expected Mean Squares Section

Source Term	DF	Term Fixed?	Denominator Term	Expected Mean Square
A: Time_Period	24	Yes	r(AB)	r+bsA
B: Observer	2	No	r(AB)	r+asB
r(AB)	48	No		r

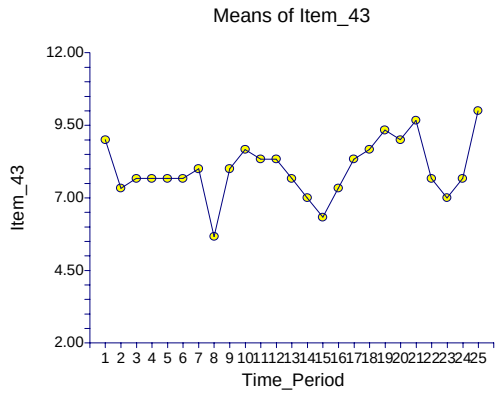
Note: Expected Mean Squares are for the balanced cell-frequency case.

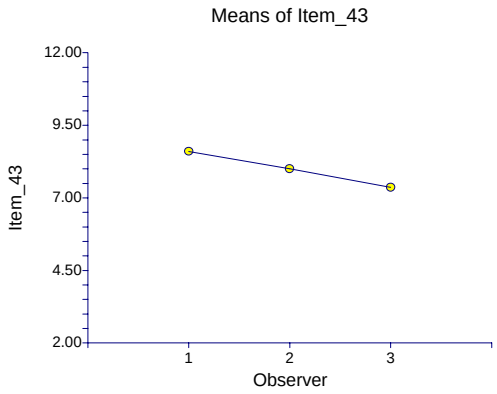
Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.10)
A: Time_Period	24	70.98666	2.957778	1.41	0.154274	0.916648
B: Observer	2	19.22667	9.613334	4.58	0.015134*	
r: Residual	48	100.7733	2.099444			
Total (Adjusted)	74	190.9867				
Total	75					

* Term significant at alpha = 0.10

Plots Section





Appendix G – Construct Validity Test Details

The paired t-tests performed in this appendix were executed in NCSS 2002. The power analysis information was obtained from PASS 2002, using the output from the paired t-test. The analysis output contained in this Appendix Js a modified version of the output from each of these software packages.

Activity

Paired t-test

Paired T-Test Report: Activity
Page/Date/Time 1 7/1/2003 11:14:35 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = A1, X2 = A2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
A1	4	7.89	0.5720723	0.2860361	7.610122	8.169878
A2	4	7.5375	0.799015	0.3995075	7.146593	7.928407
Difference	4	0.3525	0.7626435	0.3813217	-2.061274E-02	0.7256128
T for Confidence Limits = 0.9785						

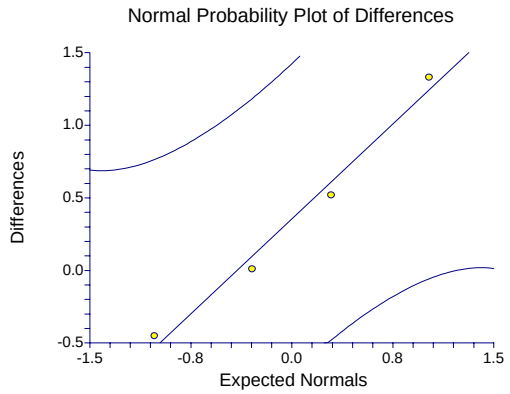
Tests of Assumptions about Differences Section

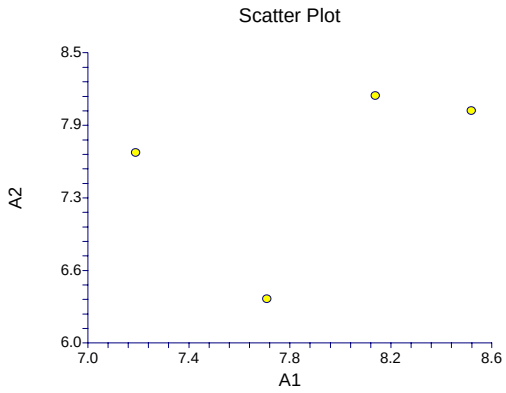
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	0.420118		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
A1-A2<>0	0.9244	0.423451	Accept Ho	0.102363	0.022528
A1-A2<0	0.9244	0.788275	Accept Ho	0.008131	0.001455
A1-A2>0	0.9244	0.211725	Reject Ho	0.180350	0.042289

Plots Section





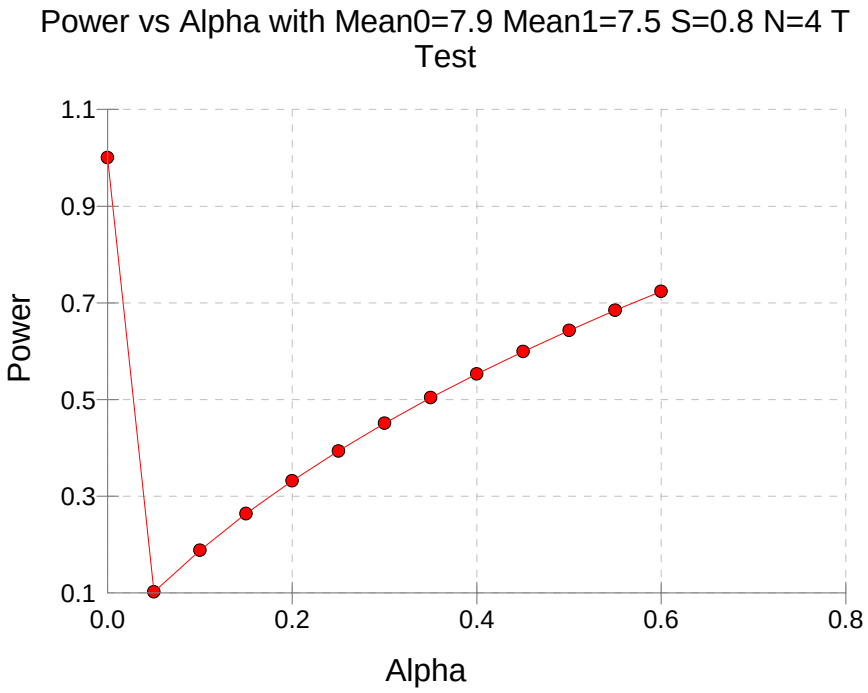
Power Analysis

One-Sample T-Test Power Analysis: Activity
Page/Date/Time 1 7/1/2003 9:53:28 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	7.9	7.5	0.8	0.462
0.10236	4	0.05000	0.89764	7.9	7.5	0.8	0.462
0.18848	4	0.10000	0.81152	7.9	7.5	0.8	0.462
0.26433	4	0.15000	0.73567	7.9	7.5	0.8	0.462
0.33246	4	0.20000	0.66754	7.9	7.5	0.8	0.462
0.39445	4	0.25000	0.60555	7.9	7.5	0.8	0.462
0.45140	4	0.30000	0.54860	7.9	7.5	0.8	0.462
0.50414	4	0.35000	0.49586	7.9	7.5	0.8	0.462
0.55331	4	0.40000	0.44669	7.9	7.5	0.8	0.462
0.59944	4	0.45000	0.40056	7.9	7.5	0.8	0.462
0.64298	4	0.50000	0.35702	7.9	7.5	0.8	0.462
0.68430	4	0.55000	0.31570	7.9	7.5	0.8	0.462
0.72371	4	0.60000	0.27629	7.9	7.5	0.8	0.462

Chart Section



Coordination

Paired t-test

Paired T-Test Report: Coordination
Page/Date/Time 1 7/1/2003 11:16:09 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = Coord1, X2 = Coord2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
Coord1	4	7.9375	0.4672883	0.2336441	7.708886	8.166114
Coord2	4	7.9775	0.2743933	0.1371966	7.843257	8.111743
Difference	4	-0.04	0.5189734	0.2594867	-0.2939005	0.2139005

T for Confidence Limits = 0.9785

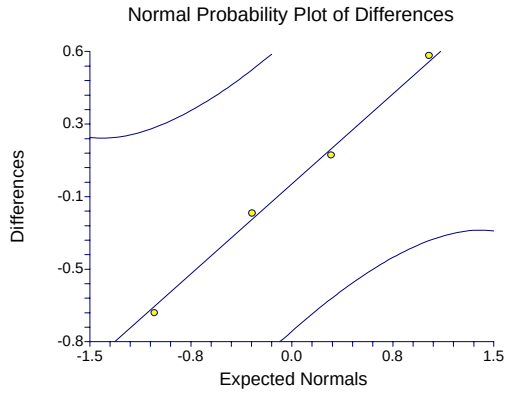
Tests of Assumptions about Differences Section

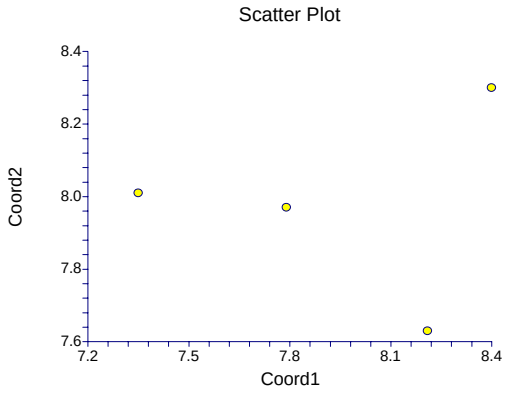
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	0.094823		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
Coord1-Coord2<>0	-0.1542	0.887277	Accept Ho	0.051452	0.010334
Coord1-Coord2<0	-0.1542	0.443639	Accept Ho	0.064123	0.013122
Coord1-Coord2>0	-0.1542	0.556361	Accept Ho	0.038412	0.007518

Plots Section





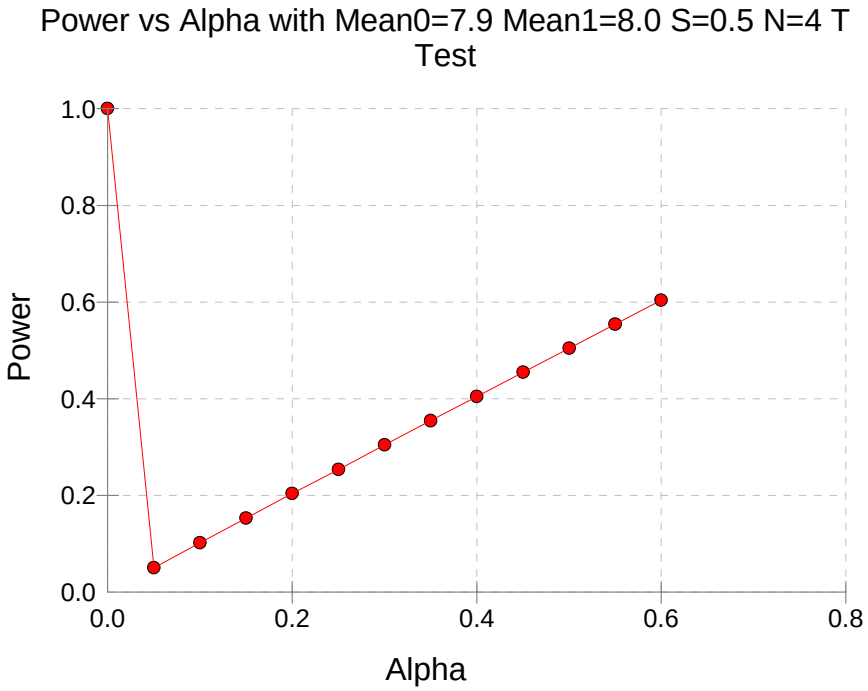
Power Analysis

One-Sample T-Test Power Analysis: Coordination
Page/Date/Time 1 7/1/2003 10:30:51 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	7.9	8.0	0.5	0.077
0.05145	4	0.05000	0.94855	7.9	8.0	0.5	0.077
0.10254	4	0.10000	0.89746	7.9	8.0	0.5	0.077
0.15336	4	0.15000	0.84664	7.9	8.0	0.5	0.077
0.20398	4	0.20000	0.79602	7.9	8.0	0.5	0.077
0.25442	4	0.25000	0.74558	7.9	8.0	0.5	0.077
0.30471	4	0.30000	0.69529	7.9	8.0	0.5	0.077
0.35486	4	0.35000	0.64514	7.9	8.0	0.5	0.077
0.40489	4	0.40000	0.59511	7.9	8.0	0.5	0.077
0.45482	4	0.45000	0.54518	7.9	8.0	0.5	0.077
0.50466	4	0.50000	0.49534	7.9	8.0	0.5	0.077
0.55441	4	0.55000	0.44559	7.9	8.0	0.5	0.077
0.60409	4	0.60000	0.39591	7.9	8.0	0.5	0.077

Chart Section



Communication

Paired t-test

Paired T-Test Report: Communication
Page/Date/Time 1 7/1/2003 11:17:28 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = Comm1, X2 = Comm2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
Comm1	4	7.815	0.4233596	0.2116798	7.607877	8.022122
Comm2	4	7.6	0.3615707	0.1807853	7.423107	7.776894
Difference	4	0.215	0.5332604	0.2666302	-4.589028E-02	0.4758903

T for Confidence Limits = 0.9785

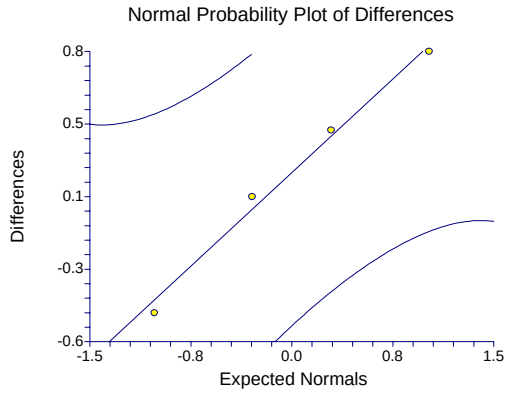
Tests of Assumptions about Differences Section

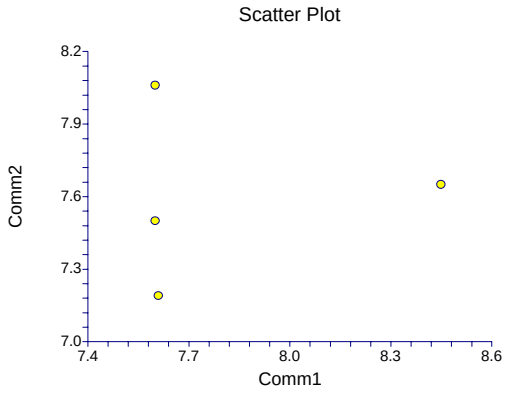
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	0.083619		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
Comm1-Comm2<>0	0.8064	0.479032	Accept Ho	0.089833	0.019438
Comm1-Comm2<0	0.8064	0.760484	Accept Ho	0.010591	0.001918
Comm1-Comm2>0	0.8064	0.239516	Reject Ho	0.157254	0.036025

Plots Section





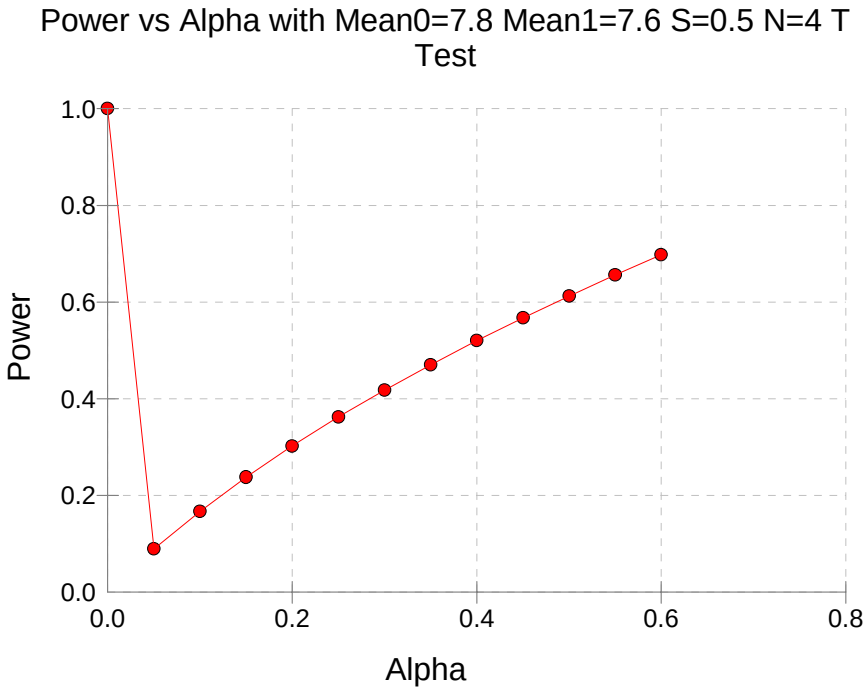
Power Analysis

One-Sample T-Test Power Analysis: Communication
Page/Date/Time 1 7/1/2003 10:34:37 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	7.8	7.6	0.5	0.403
0.08983	4	0.05000	0.91017	7.8	7.6	0.5	0.403
0.16785	4	0.10000	0.83215	7.8	7.6	0.5	0.403
0.23821	4	0.15000	0.76179	7.8	7.6	0.5	0.403
0.30272	4	0.20000	0.69728	7.8	7.6	0.5	0.403
0.36250	4	0.25000	0.63750	7.8	7.6	0.5	0.403
0.41836	4	0.30000	0.58164	7.8	7.6	0.5	0.403
0.47089	4	0.35000	0.52911	7.8	7.6	0.5	0.403
0.52059	4	0.40000	0.47941	7.8	7.6	0.5	0.403
0.56785	4	0.45000	0.43215	7.8	7.6	0.5	0.403
0.61301	4	0.50000	0.38699	7.8	7.6	0.5	0.403
0.65636	4	0.55000	0.34364	7.8	7.6	0.5	0.403
0.69815	4	0.60000	0.30185	7.8	7.6	0.5	0.403

Chart Section



Conflict Resolution

Paired t-test

Paired T-Test Report: Conflict Resolution
Page/Date/Time 1 7/1/2003 11:18:44 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = CR1, X2 = CR2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
CR1	4	8.4675	0.5628721	0.2814361	8.192122	8.742877
CR2	4	8.1625	0.5362446	0.2681223	7.90015	8.42485
Difference	4	0.305	1.094029	0.5470147	-0.2302387	0.8402387

T for Confidence Limits = 0.9785

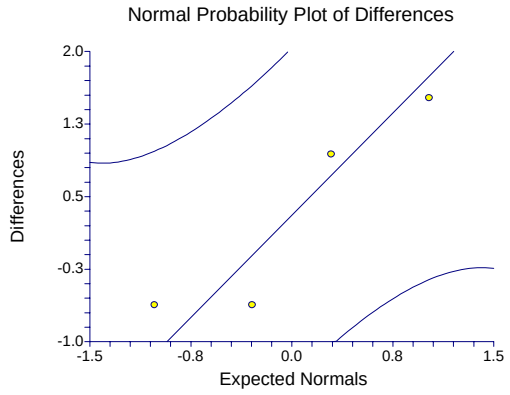
Tests of Assumptions about Differences Section

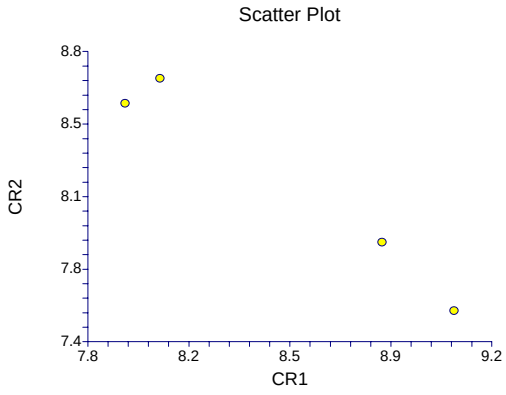
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	-0.981517		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
CR1-CR2<>0	0.5576	0.616038	Accept Ho	0.069025	0.014435
CR1-CR2<0	0.5576	0.691981	Accept Ho	0.017910	0.003333
CR1-CR2>0	0.5576	0.308019	Reject Ho	0.114932	0.025142

Plots Section





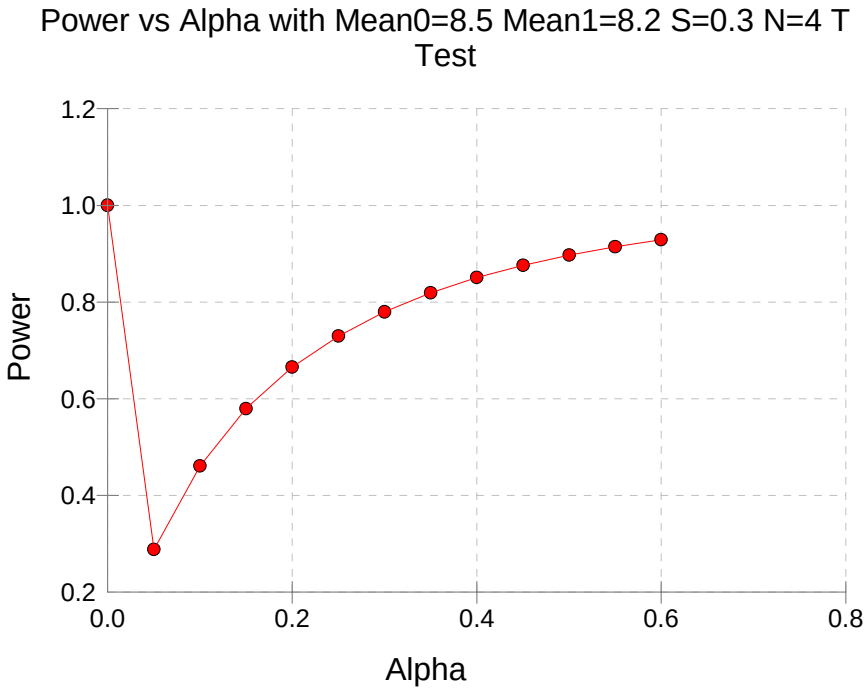
Power Analysis

One-Sample T-Test Power Analysis: Conflict Resolution
Page/Date/Time 1 7/1/2003 10:40:11 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	8.5	8.2	0.3	1.000
0.28875	4	0.05000	0.71125	8.5	8.2	0.3	1.000
0.46124	4	0.10000	0.53876	8.5	8.2	0.3	1.000
0.57966	4	0.15000	0.42034	8.5	8.2	0.3	1.000
0.66565	4	0.20000	0.33435	8.5	8.2	0.3	1.000
0.73032	4	0.25000	0.26968	8.5	8.2	0.3	1.000
0.78019	4	0.30000	0.21981	8.5	8.2	0.3	1.000
0.81943	4	0.35000	0.18057	8.5	8.2	0.3	1.000
0.85082	4	0.40000	0.14918	8.5	8.2	0.3	1.000
0.87629	4	0.45000	0.12371	8.5	8.2	0.3	1.000
0.89724	4	0.50000	0.10276	8.5	8.2	0.3	1.000
0.91468	4	0.55000	0.08532	8.5	8.2	0.3	1.000
0.92940	4	0.60000	0.07060	8.5	8.2	0.3	1.000

Chart Section



Decision Making

Paired t-test

Paired T-Test Report: Decision Making
Page/Date/Time 1 7/1/2003 11:19:28 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = DM1, X2 = DM2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
DM1	4	6.7775	0.462268	0.231134	6.551342	7.003658
DM2	4	6.6225	0.5965107	0.2982554	6.330666	6.914335
Difference	4	0.155	0.4657968	0.2328984	-7.288464E-02	0.3828847

T for Confidence Limits = 0.9785

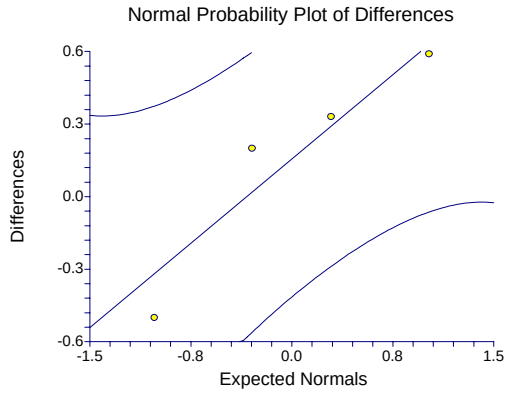
Tests of Assumptions about Differences Section

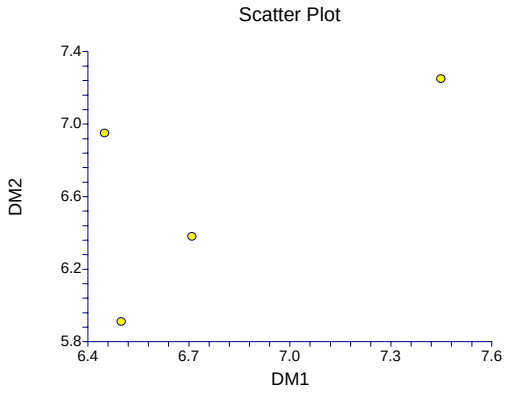
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	0.639262		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
DM1-DM2<>0	0.6655	0.553311	Accept Ho	0.077119	0.016362
DM1-DM2<0	0.6655	0.723345	Accept Ho	0.014334	0.002635
DM1-DM2>0	0.6655	0.276655	Reject Ho	0.132235	0.029498

Plots Section



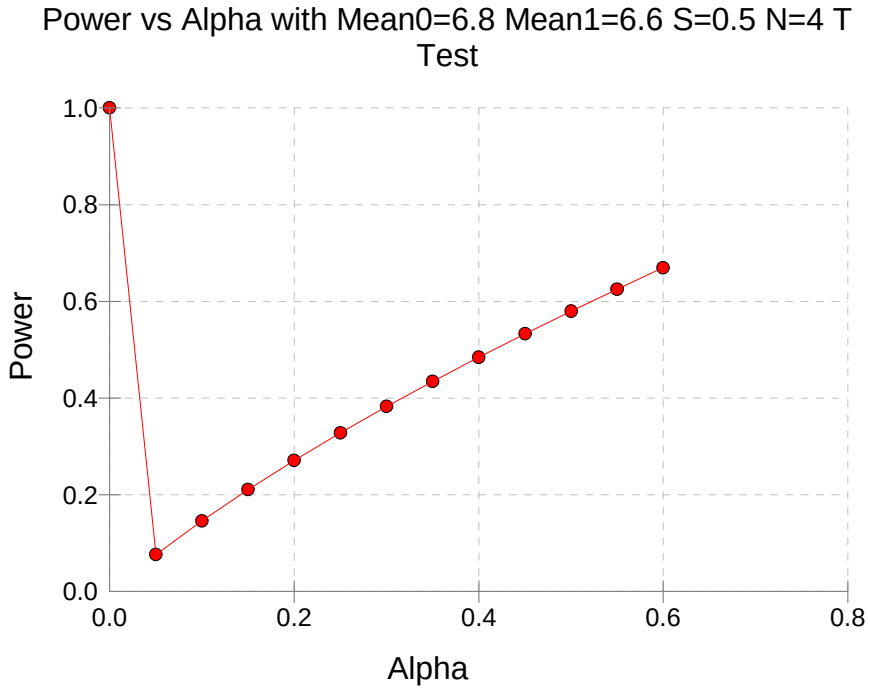


One-Sample T-Test Power Analysis: Decision Making
Page/Date/Time 1 7/1/2003 10:44:43 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	6.8	6.6	0.5	0.333
0.07712	4	0.05000	0.92288	6.8	6.6	0.5	0.333
0.14657	4	0.10000	0.85343	6.8	6.6	0.5	0.333
0.21093	4	0.15000	0.78907	6.8	6.6	0.5	0.333
0.27133	4	0.20000	0.72867	6.8	6.6	0.5	0.333
0.32848	4	0.25000	0.67152	6.8	6.6	0.5	0.333
0.38289	4	0.30000	0.61711	6.8	6.6	0.5	0.333
0.43495	4	0.35000	0.56505	6.8	6.6	0.5	0.333
0.48500	4	0.40000	0.51500	6.8	6.6	0.5	0.333
0.53329	4	0.45000	0.46671	6.8	6.6	0.5	0.333
0.58006	4	0.50000	0.41994	6.8	6.6	0.5	0.333
0.62550	4	0.55000	0.37450	6.8	6.6	0.5	0.333
0.66980	4	0.60000	0.33020	6.8	6.6	0.5	0.333

Chart Section



Joint Knowledge

Paired t-test

Paired T-Test Report: Joint Knowledge
Page/Date/Time 1 7/1/2003 11:20:10 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = JK1, X2 = JK2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
JK1	4	7.925	0.7175189	0.3587594	7.573964	8.276036
JK2	4	7.9575	0.115	0.0575	7.901238	8.013762
Difference	4	-0.0325	0.8147545	0.4073773	-0.4311074	0.3661074

T for Confidence Limits = 0.9785

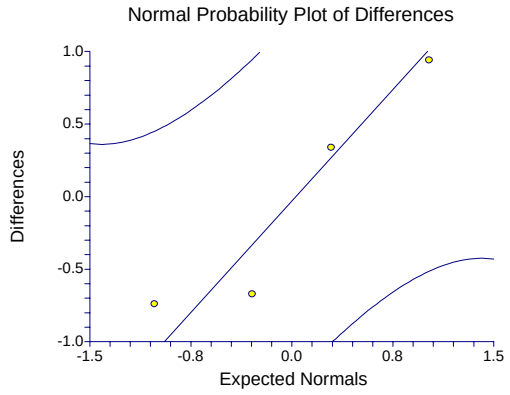
Tests of Assumptions about Differences Section

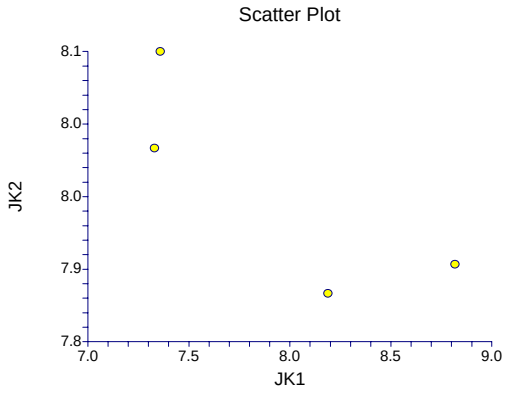
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	-0.822682		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
JK1-JK2<>0	-0.0798	0.941437	Accept Ho	0.050389	0.010089
JK1-JK2<0	-0.0798	0.470719	Accept Ho	0.056975	0.011529
JK1-JK2>0	-0.0798	0.529281	Accept Ho	0.043705	0.008642

Plots Section





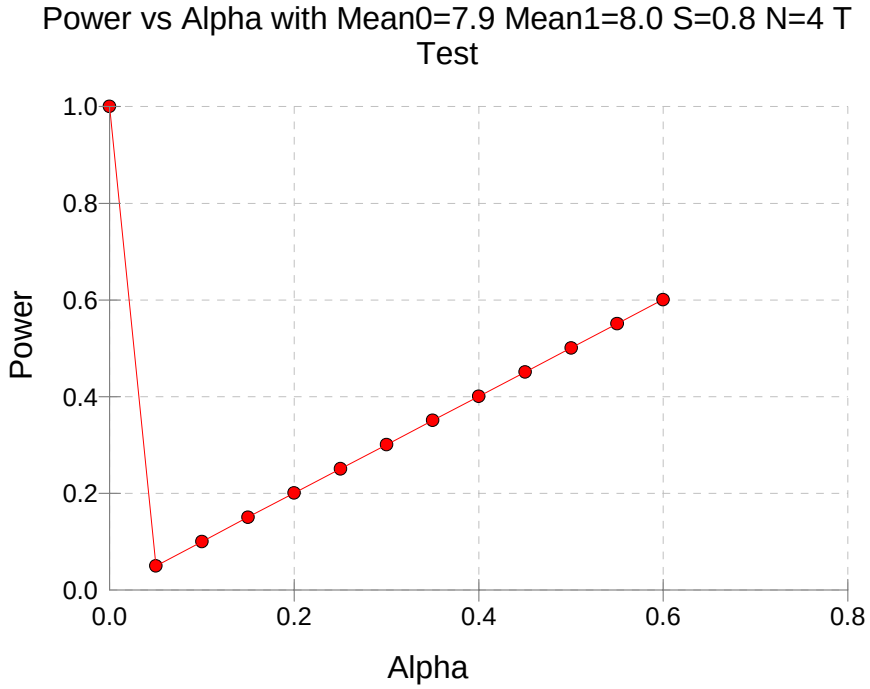
Power Analysis

One-Sample T-Test Power Analysis: Joint Knowledge
Page/Date/Time 1 7/1/2003 10:48:12 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	7.9	8.0	0.8	0.040
0.05039	4	0.05000	0.94961	7.9	8.0	0.8	0.040
0.10068	4	0.10000	0.89932	7.9	8.0	0.8	0.040
0.15090	4	0.15000	0.84910	7.9	8.0	0.8	0.040
0.20107	4	0.20000	0.79893	7.9	8.0	0.8	0.040
0.25119	4	0.25000	0.74881	7.9	8.0	0.8	0.040
0.30126	4	0.30000	0.69874	7.9	8.0	0.8	0.040
0.35130	4	0.35000	0.64870	7.9	8.0	0.8	0.040
0.40131	4	0.40000	0.59869	7.9	8.0	0.8	0.040
0.45130	4	0.45000	0.54870	7.9	8.0	0.8	0.040
0.50125	4	0.50000	0.49875	7.9	8.0	0.8	0.040
0.55119	4	0.55000	0.44881	7.9	8.0	0.8	0.040
0.60110	4	0.60000	0.39890	7.9	8.0	0.8	0.040

Chart Section



Personalities

Paired t-test

Paired T-Test Report: Personalities
Page/Date/Time 1 7/1/2003 11:21:01 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = P1, X2 = P2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
P1	4	8.18	0.4769347	0.2384673	7.946666	8.413334
P2	4	7.8475	0.9517484	0.4758742	7.38187	8.313129
Difference	4	0.3325	1.387501	0.6937504	-0.3463155	1.011316

T for Confidence Limits = 0.9785

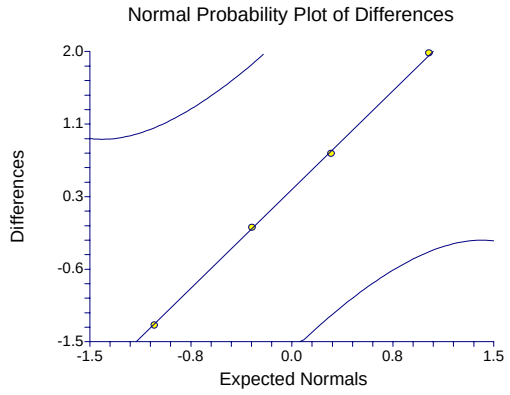
Tests of Assumptions about Differences Section

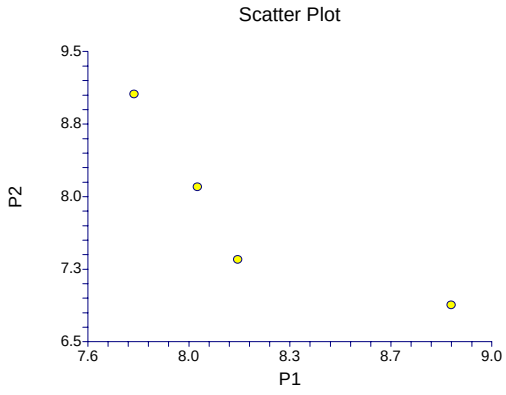
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	-0.872250		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
P1-P2<>0	0.4793	0.664509	Accept Ho	0.064051	0.013263
P1-P2<0	0.4793	0.667746	Accept Ho	0.020945	0.003934
P1-P2>0	0.4793	0.332254	Reject Ho	0.103391	0.022309

Plots Section





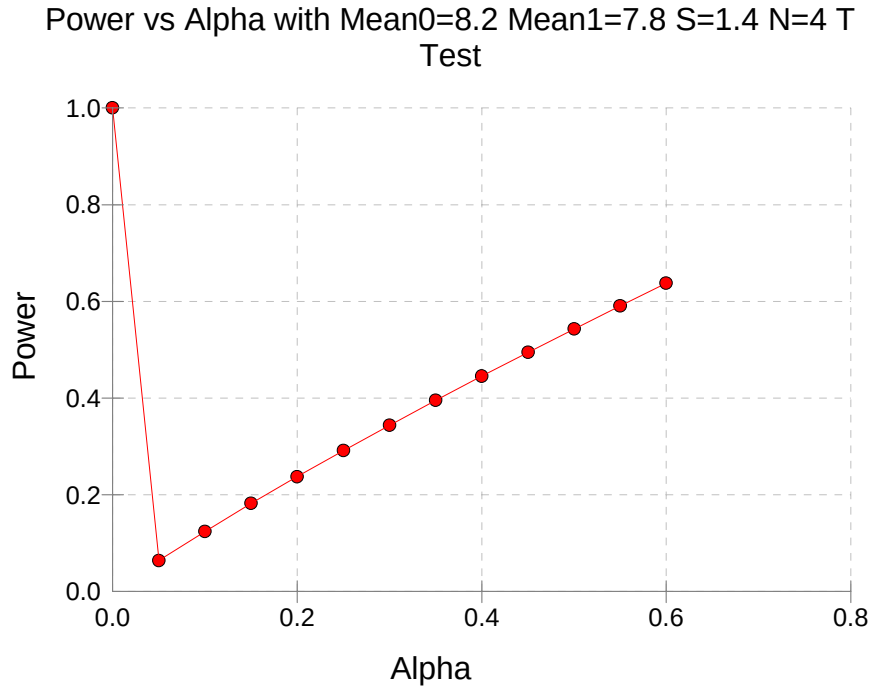
Power Analysis

One-Sample T-Test Power Analysis: Personalities
Page/Date/Time 1 7/1/2003 10:50:12 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	8.2	7.8	1.4	0.240
0.06405	4	0.05000	0.93595	8.2	7.8	1.4	0.240
0.12434	4	0.10000	0.87566	8.2	7.8	1.4	0.240
0.18205	4	0.15000	0.81795	8.2	7.8	1.4	0.240
0.23773	4	0.20000	0.76227	8.2	7.8	1.4	0.240
0.29171	4	0.25000	0.70829	8.2	7.8	1.4	0.240
0.34424	4	0.30000	0.65576	8.2	7.8	1.4	0.240
0.39550	4	0.35000	0.60450	8.2	7.8	1.4	0.240
0.44567	4	0.40000	0.55433	8.2	7.8	1.4	0.240
0.49488	4	0.45000	0.50512	8.2	7.8	1.4	0.240
0.54325	4	0.50000	0.45675	8.2	7.8	1.4	0.240
0.59088	4	0.55000	0.40912	8.2	7.8	1.4	0.240
0.63786	4	0.60000	0.36214	8.2	7.8	1.4	0.240

Chart Section



Overall

Paired t-test

First Test

Paired T-Test Report: Overall (alpha = .4)
Page/Date/Time 1 7/1/2003 11:13:04 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = O1, X2 = O2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
O1	4	7.995	0.3042477	0.1521239	7.846151	8.143849
O2	4	8.655	0.1369915	6.849574E-02	8.587978	8.722021
Difference	4	-0.66	0.4159327	0.2079663	-0.8634893	-0.4565107

T for Confidence Limits = 0.9785

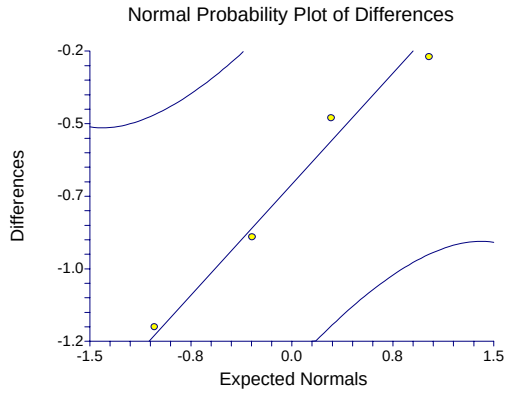
Tests of Assumptions about Differences Section

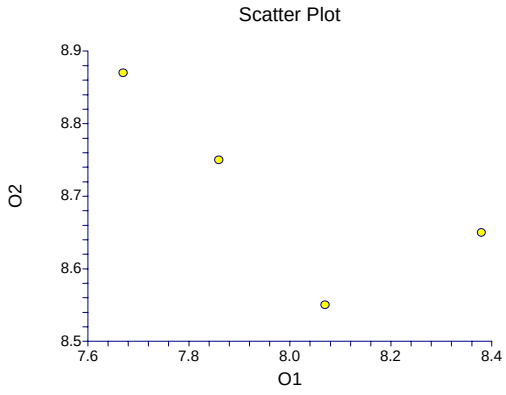
Assumption	**Value**	**Probability**	**Decision(40%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	-0.739775		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (40%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
O1-O2<>0	-3.1736	0.050341	Reject Ho	0.575305	0.192632
O1-O2<0	-3.1736	0.025171	Reject Ho	0.774121	0.326418
O1-O2>0	-3.1736	0.974829	Accept Ho	0.000007	0.000001

Plots Section





Second Test

Paired T-Test Report: Overall (alpha = .05)
Page/Date/Time 1 7/1/2003 11:10:36 AM
Database C:\Documents and Settings\Br ... alysis\Construct Validity.S0
Variable X1 = O1, X2 = O2

Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	**60% LCL of Mean**	**60% UCL of Mean**
O1	4	7.995	0.3042477	0.1521239	7.510874	8.479126
O2	4	8.655	0.1369915	6.849574E-02	8.437016	8.872984
Difference	4	-0.66	0.4159327	0.2079663	-1.321842	1.841721E-03

T for Confidence Limits = 3.1824

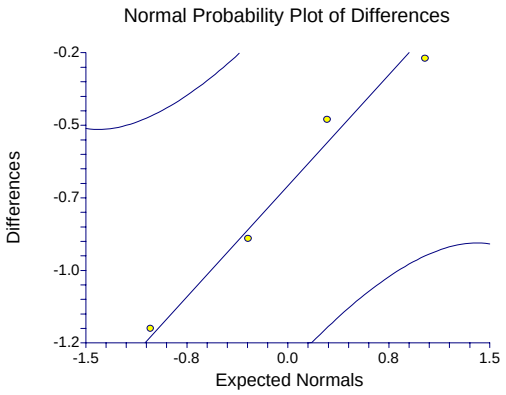
Tests of Assumptions about Differences Section

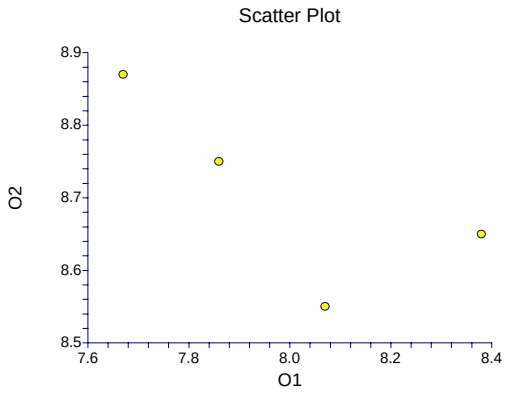
Assumption	**Value**	**Probability**	**Decision(5%)**
Skewness Normality	0.0000		
Kurtosis Normality		1.000000	Cannot reject normality
Omnibus Normality			
Correlation Coefficient	-0.739775		

T-Test For Difference Between Means Section

Alternative Hypothesis	**T-Value**	**Prob Level**	**Decision (5%)**	**Power (Alpha=.05)**	**Power (Alpha=.01)**
O1-O2<>0	-3.1736	0.050341	Accept Ho	0.575305	0.192632
O1-O2<0	-3.1736	0.025171	Reject Ho	0.774121	0.326418
O1-O2>0	-3.1736	0.974829	Accept Ho	0.000007	0.000001

Plots Section





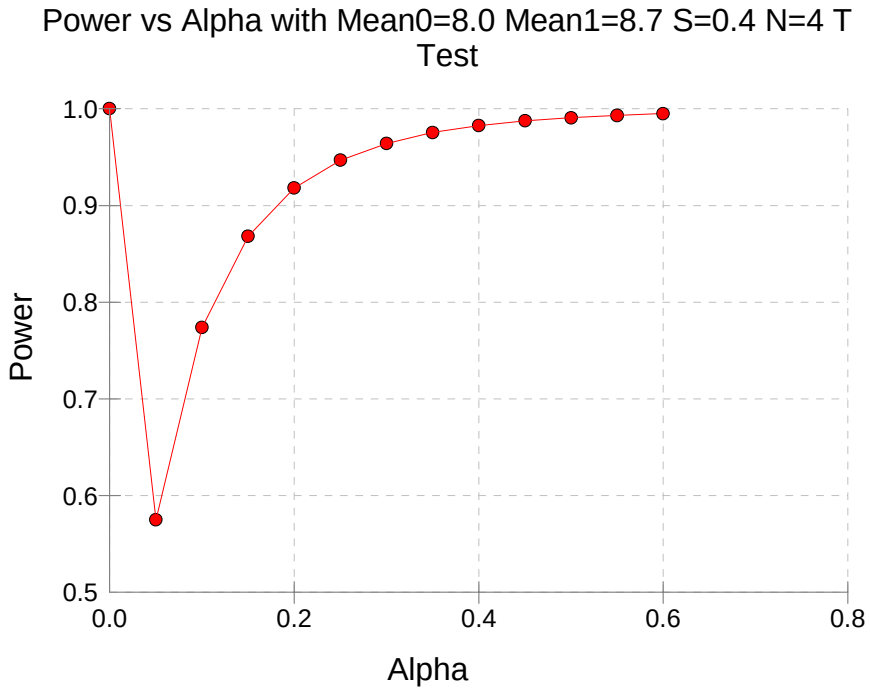
Power Analysis

One-Sample T-Test Power Analysis: Overall
Page/Date/Time 1 7/1/2003 10:55:06 AM

Numeric Results for One-Sample T-Test
Null Hypothesis: Mean0=Mean1 Alternative Hypothesis: Mean0<>Mean1
Unknown standard deviation.

Power	N	Alpha	Beta	Mean0	Mean1	S	Effect Size
1.00000	4	0.00000	0.00000	8.0	8.7	0.4	1.587
0.57530	4	0.05000	0.42470	8.0	8.7	0.4	1.587
0.77413	4	0.10000	0.22587	8.0	8.7	0.4	1.587
0.86802	4	0.15000	0.13198	8.0	8.7	0.4	1.587
0.91810	4	0.20000	0.08190	8.0	8.7	0.4	1.587
0.94689	4	0.25000	0.05311	8.0	8.7	0.4	1.587
0.96437	4	0.30000	0.03563	8.0	8.7	0.4	1.587
0.97543	4	0.35000	0.02457	8.0	8.7	0.4	1.587
0.98267	4	0.40000	0.01733	8.0	8.7	0.4	1.587
0.98754	4	0.45000	0.01246	8.0	8.7	0.4	1.587
0.99090	4	0.50000	0.00910	8.0	8.7	0.4	1.587
0.99327	4	0.55000	0.00673	8.0	8.7	0.4	1.587
0.99498	4	0.60000	0.00502	8.0	8.7	0.4	1.587

Chart Section



Appendix H – Factor Analysis Details

This section lists the NCSS output for the factor analyses performed in this study. It is subdivided by team process construct.

Coordination

Factor Analysis Report: Coordination

Page/Date/Time 1 6/19/2003 7:33:52 AM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q2	62	8.2741935483871	1.59064225443594	0.456475
Q3	62	7.59677419354839	2.1988050213573	0.532439
Q4	62	8.56451612903226	1.31349824167493	0.405902
Q5	62	8.20967741935483	2.04167318697988	0.418473
Q6	62	7.88709677419355	2.24049788604374	0.285468
Q7	62	8.03225806451613	1.92454160550697	0.012337

Correlation Section

	Variables				
Variables	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
Q2	1.000000	0.374293	0.403328	0.491845	0.478024
Q3	0.374293	1.000000	0.517171	0.530383	0.413220
Q4	0.403328	0.517171	1.000000	0.389162	0.283827
Q5	0.491845	0.530383	0.389162	1.000000	0.259707
Q6	0.478024	0.413220	0.283827	0.259707	1.000000
Q7	0.136297	0.068982	0.303961	-0.160290	0.000858

Phi=0.359034 Log(Det|R|)=-1.622843 Bartlett Test=94.40 DF=15 Prob=0.000000

	Variables
Variables	**Q7**
Q2	0.136297
Q3	0.068982
Q4	0.303961
Q5	-0.160290
Q6	0.000858
Q7	1.000000

Phi=0.359034 Log(Det|R|)=-1.622843 Bartlett Test=94.40 DF=15 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	2.111092	99.46	99.46	||||||||||||||||||||
2	0.346817	16.34	115.80	||||
3	0.186263	8.78	124.57	||
4	0.043073	2.03	126.60	|
5	-0.199541	-9.40	117.20	||
6	-0.365117	-17.20	100.00	||||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q2	0.465002
Q3	0.502205
Q4	0.438487
Q5	0.445225
Q6	0.367726
Q7	0.076444

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q2	||||||||
Q3	|||||||||
Q4	|||||||
Q5	||||||||
Q6	|||||||
Q7	||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q2	0.675629
Q3	0.729684
Q4	0.637104
Q5	0.646895
Q6	0.534292
Q7	0.111070

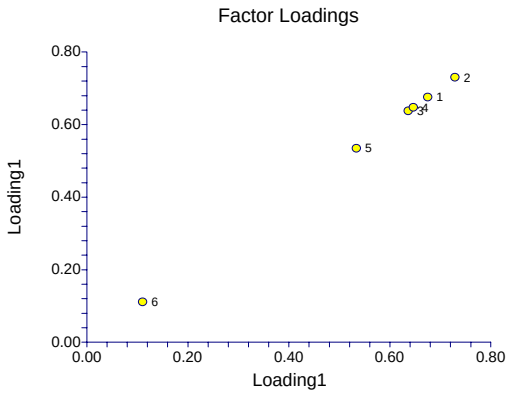
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q2	|||||||||||
Q3	||||||||||||
Q4	||||||||||
Q5	|||||||||||
Q6	||||||||
Q7	|||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q2	||||||||	||||||||
Q3	|||||||||	|||||||||
Q4	|||||||	|||||||
Q5	||||||||	||||||||
Q6	||||||	||||||
Q7	|	|

Plots Section



Communication

Factor Analysis Report: Communication

Page/Date/Time	1 6/19/2003 8:25:19 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q8	62	8	1.34286212565721	0.549098
Q9	62	8.32258064516129	1.62737033729303	0.054211
Q10	62	7.75806451612903	2.3863373668776	0.002015
Q11	62	7.74193548387097	2.07194755828762	0.206655
Q12	62	7.83870967741936	1.65156208607703	0.578533

Correlation Section

	Variables				
Variables	**Q8**	**Q9**	**Q10**	**Q11**	**Q12**
Q8	1.000000	0.247552	-0.081852	0.265138	0.583944
Q9	0.247552	1.000000	0.210387	0.243879	0.031874
Q10	-0.081852	0.210387	1.000000	0.162891	-0.163966
Q11	0.265138	0.243879	0.162891	1.000000	0.380472
Q12	0.583944	0.031874	-0.163966	0.380472	1.000000

Phi=0.279679 Log(Det|R|)=-0.843061 Bartlett Test=49.32 DF=10 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	1.390513	99.44	99.44	|||||||||||||||||||
2	0.423510	30.29	129.73	|||||||
3	0.079475	5.68	135.41	||
4	-0.241071	-17.24	118.17	||||
5	-0.254119	-18.17	100.00	||||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q8	0.628402
Q9	0.197450
Q10	-0.038069
Q11	0.385510
Q12	0.645025

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q8	||||||||||||
Q9	||||
Q10		
Q11	||||||||
Q12	||||||||||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q8	0.741011
Q9	0.232833
Q10	-0.044891
Q11	0.454594
Q12	0.760614

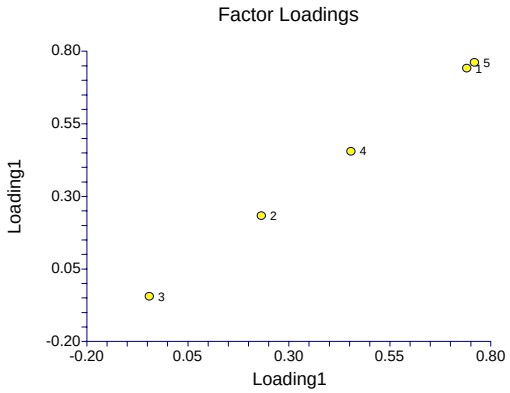
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q8	||||||||||||||
Q9	|||||
Q10	|
Q11	|||||||||
Q12	|||||||||||||||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q8	||||||||||	||||||||||
Q9	||	||
Q10	|	|
Q11	|||||	|||||
Q12	|||||||||||	|||||||||||

Plots Section



Conflict Resolution

Factor Analysis Report: Conflict Resolution

Page/Date/Time	1 6/19/2003 8:47:16 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q13	62	8.32258064516129	1.76276788266176	0.833090
Q14	62	7.70967741935484	1.59288476088164	0.063244
Q15	62	8.14516129032258	2.17923715178567	0.388541
Q16	62	8.69354838709677	1.56381951159123	0.578618
Q17	62	8.29032258064516	1.68296281262008	0.379930

Correlation Section

	Variables				
Variables	**Q13**	**Q14**	**Q15**	**Q16**	**Q17**
Q13	1.000000	0.390040	0.563719	0.690604	0.509449
Q14	0.390040	1.000000	0.102069	0.009766	0.190952
Q15	0.563719	0.102069	1.000000	0.499115	0.386136
Q16	0.690604	0.009766	0.499115	1.000000	0.520211
Q17	0.509449	0.190952	0.386136	0.520211	1.000000

Phi=0.438429 Log(Det|R|)=-1.802591 Bartlett Test=105.45 DF=10 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	2.243422	99.52	99.52	||||||||||||||||||||
2	0.265249	11.77	111.29	|||
3	0.025493	1.13	112.42	|
4	-0.021134	-0.94	111.48	|
5	-0.258864	-11.48	100.00	|||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q13	-0.609383
Q14	-0.167901
Q15	-0.416162
Q16	-0.507856
Q17	-0.411525

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q13	||||||||||||
Q14	||||
Q15	||||||||
Q16	||||||||||
Q17	||||||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q13	-0.912738
Q14	-0.251483
Q15	-0.623330
Q16	-0.760670
Q17	-0.616384

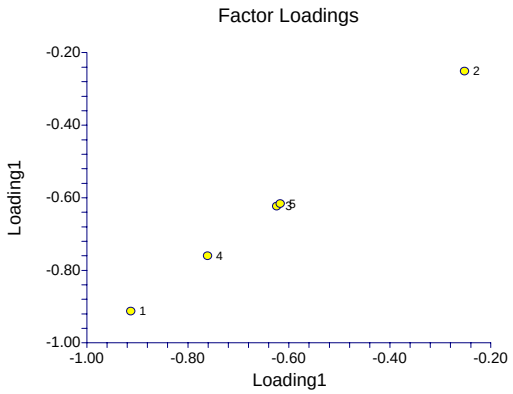
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q13	|||||||||||||||||||
Q14	||||||
Q15	||||||||||||
Q16	||||||||||||||
Q17	||||||||||||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q13	||||||||||||||||	||||||||||||||||
Q14	||	||
Q15	||||||||	||||||||
Q16	||||||||||||	||||||||||||
Q17	||||||||	||||||||

Plots Section



Decision Making

Factor Analysis Report: Decision Making

Page/Date/Time	1 6/19/2003 9:08:47 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q18	62	8.25806451612903	1.73616766462916	0.100954
Q19	62	6.29032258064516	2.72443164037718	0.402513
Q20	62	7.35483870967742	2.03330284644522	0.716871
Q21	62	8.24193548387097	1.92213579867717	0.323671

Correlation Section

	Variables			
Variables	**Q18**	**Q19**	**Q20**	**Q21**
Q18	1.000000	0.067080	0.270840	0.324853
Q19	0.067080	1.000000	0.590719	0.355763
Q20	0.270840	0.590719	1.000000	0.430684
Q21	0.324853	0.355763	0.430684	1.000000

Phi=0.375142 Log(Det|R|)=-0.814872 Bartlett Test=47.94 DF=6 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	1.544008	99.94	99.94	||||||||||||||||
2	0.215102	13.92	113.86	|||
3	-0.008132	-0.53	113.34		
4	-0.206025	-13.34	100.00	|||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q18	-0.255703
Q19	-0.510582
Q20	-0.681390
Q21	-0.457854

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q18	||||||
Q19	||||||||||
Q20	||||||||||||
Q21	|||||||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q18	-0.317732
Q19	-0.634439
Q20	-0.846682
Q21	-0.568921

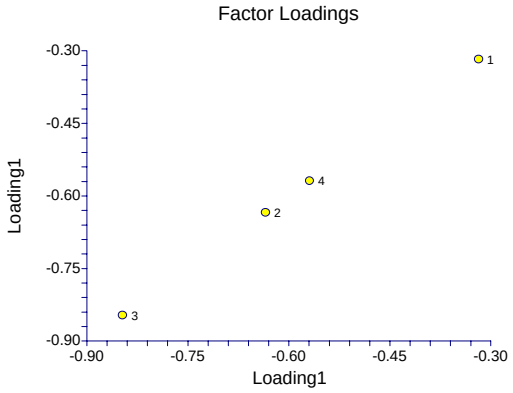
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q18	||||||
Q19	||||||||||||
Q20	||||||||||||||||
Q21	||||||||||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q18	|||	|||
Q19	||||||||	||||||||
Q20	||||||||||||||	||||||||||||||
Q21	|||||||	|||||||

Plots Section



Problem Solving

Factor Analysis Report: Problem Solving

Page/Date/Time	1 6/19/2003 9:24:37 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q22	62	8.51612903225807	1.46815302150462	0.504630
Q23	62	8.2258064516129	2.48549467213342	0.592864
Q24	62	8.59677419354839	1.19369487574182	0.105967

Correlation Section

	Variables		
Variables	**Q22**	**Q23**	**Q24**
Q22	1.000000	0.547068	0.223595
Q23	0.547068	1.000000	0.257733
Q24	0.223595	0.257733	1.000000

Phi=0.372247 Log(Det|R|)=-0.434871 Bartlett Test=25.73 DF=3 Prob=0.000011

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	1.203461	99.98	99.98	||||||||||||||||||||
2	0.011044	0.92	100.90	|
3	-0.010799	-0.90	100.00	|

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q22	-0.647546
Q23	-0.701878
Q24	-0.296735

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q22	||||||||||||
Q23	||||||||||||||
Q24	||||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q22	-0.710373
Q23	-0.769976
Q24	-0.325525

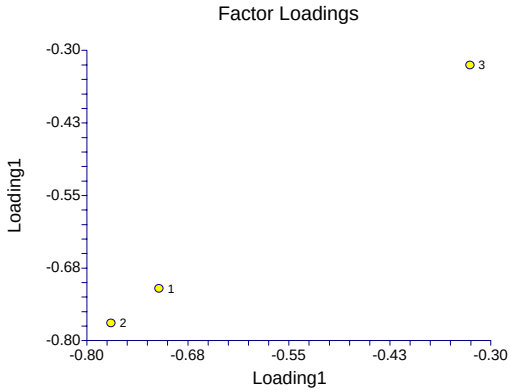
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q22	||||||||||||||
Q23	|||||||||||||||
Q24	|||||||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q22	||||||||||	||||||||||
Q23	|||||||||||	|||||||||||
Q24	|||	|||

Plots Section



Boundary Spanning

Data unreliable. Factor analysis not performed.

Joint Knowledge

Factor Analysis Report: Joint Knowledge

Page/Date/Time	1 6/19/2003 9:51:43 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q27	62	9.19354838709678	0.972384422352303	0.036390
Q28	62	7.62903225806452	2.91538525400507	0.705916
Q29	62	8.29032258064516	1.95345089163745	0.771990
Q30	62	7.45161290322581	2.19296505137498	0.073719

Correlation Section

	Variables			
Variables	**Q27**	**Q28**	**Q29**	**Q30**
Q27	1.000000	0.095136	0.237475	-0.026287
Q28	0.095136	1.000000	0.735977	-0.270807
Q29	0.237475	0.735977	1.000000	-0.203314
Q30	-0.026287	-0.270807	-0.203314	1.000000

Phi=0.347004 Log(Det|R|)=-0.929172 Bartlett Test=54.67 DF=6 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	1.588016	99.84	99.84	|||||||||||||||||||
2	0.124252	7.81	107.65	||
3	-0.023087	-1.45	106.20		
4	-0.098664	-6.20	100.00	||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q27	0.151379
Q28	0.666729
Q29	0.697234
Q30	-0.215457

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q27	||||
Q28	||||||||||||||
Q29	||||||||||||||
Q30	|||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q27	0.190763
Q28	0.840188
Q29	0.878630
Q30	-0.271512

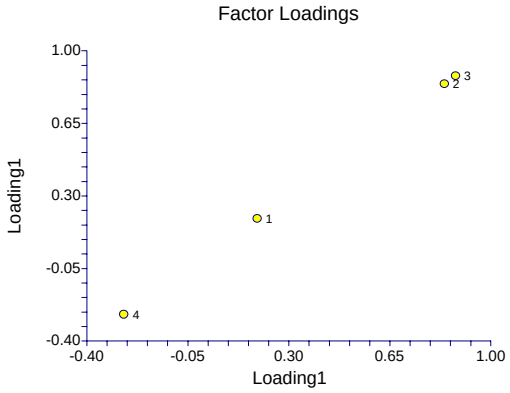
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q27	||||
Q28	||||||||||||||||
Q29	||||||||||||||||
Q30	||||||

Bar Chart of Communalities after Varimax Rotation

Variables	Factors Factor1	Communality
Q27	|	|
Q28	||||||||||||	||||||||||||
Q29	|||||||||||||	|||||||||||||
Q30	||	||

Plots Section



Motivation

Factor Analysis Report: Motivation

Page/Date/Time	1 6/19/2003 10:09:59 AM
Database	C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q31	62	8.16129032258065	1.88343554583496	0.350350
Q32	62	7.93548387096774	2.26064670670197	0.435061
Q33	62	6.66129032258065	1.98307220164925	0.025529
Q34	62	8.79032258064516	1.02648687476846	0.145769

Correlation Section

	Variables			
Variables	**Q31**	**Q32**	**Q33**	**Q34**
Q31	1.000000	0.372106	0.115817	0.246724
Q32	0.372106	1.000000	0.155944	0.262528
Q33	0.115817	0.155944	1.000000	-0.059621
Q34	0.246724	0.262528	-0.059621	1.000000

Phi=0.227135 Log(Det|R|)=-0.291119 Bartlett Test=17.13 DF=6 Prob=0.008826

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	0.956708	100.00	100.00	|||||||||||||||||||
2	0.127571	13.33	113.34	|||
3	0.015138	1.58	114.92		
4	-0.142747	-14.92	100.00	|||

Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q31	-0.605147
Q32	-0.674350
Q33	-0.163352
Q34	-0.390339

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors
Variables	**Factor1**
Q31	||||||||||||
Q32	||||||||||||||
Q33	||||
Q34	||||||||

Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q31	-0.591903
Q32	-0.659592
Q33	-0.159777
Q34	-0.381797

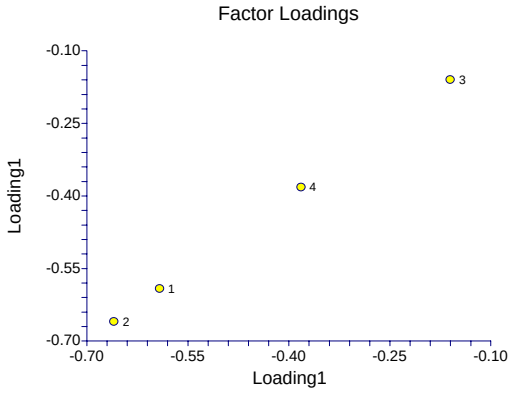
Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors
Variables	**Factor1**
Q31	||||||||||||
Q32	||||||||||||||
Q33	||||
Q34	||||||||

Bar Chart of Communalities after Varimax Rotation

	Factors	
Variables	**Factor1**	**Communality**
Q31	||||||||	||||||||
Q32	|||||||||	|||||||||
Q33	|	|
Q34	|||	|||

Plots Section



Personalities

Factor Analysis Report: Personalities

Page/Date/Time 1 6/19/2003 10:30:35 AM
Database C:\Documents and Settings\Br ... RL\Data Analysis\Data Set.S0

Descriptive Statistics Section

Variables	Count	Mean	Standard Deviation	Communality
Q35	62	8.46774193548387	2.42742352742394	0.939704
Q36	62	8.9516129032258	1.04689109388404	0.070446
Q37	62	7.87096774193549	2.76087341233573	0.716922
Q38	62	7.59677419354839	2.04426168017405	0.699261
Q39	62	7.48387096774193	2.03849780342122	0.744591

Correlation Section

	Variables				
Variables	**Q35**	**Q36**	**Q37**	**Q38**	**Q39**
Q35	1.000000	0.002601	0.801697	-0.073692	0.596223
Q36	0.002601	1.000000	-0.041898	0.235856	0.126377
Q37	0.801697	-0.041898	1.000000	0.147479	0.605490
Q38	-0.073692	0.235856	0.147479	1.000000	0.484251
Q39	0.596223	0.126377	0.605490	0.484251	1.000000

Phi=0.412293 Log(Det|R|)=-2.263237 Bartlett Test=132.40 DF=10 Prob=0.000000

Eigenvalues after Varimax Rotation

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	2.112629	66.83	66.83	|||||||||||||
2	1.058294	33.48	100.31	|||||||
3	0.075275	2.38	102.69	|
4	0.008179	0.26	102.95	|
5	-0.093328	-2.95	100.00	|

Eigenvectors after Varimax Rotation

	Factors	
Variables	**Factor1**	**Factor2**
Q35	-0.591159	-0.417043
Q36	-0.044484	0.262409
Q37	-0.557995	-0.171978
Q38	-0.204838	0.795028
Q39	-0.543356	0.309145

Bar Chart of Absolute Eigenvectors after Varimax Rotation

	Factors	
Variables	**Factor1**	**Factor2**
Q35	|||||||||||	||||||||
Q36	|	||||||
Q37	||||||||||	||||
Q38	|||||	|||||||||||||
Q39	|||||||||	||||||

Factor Loadings after Varimax Rotation

	Factors	
Variables	**Factor1**	**Factor2**
Q35	-0.958460	-0.145111
Q36	0.008718	0.265273
Q37	-0.843695	0.071421
Q38	-0.073582	0.832975
Q39	-0.690414	0.517609

Bar Chart of Absolute Factor Loadings after Varimax Rotation

	Factors	
Variables	**Factor1**	**Factor2**
Q35	|||||||||||||||||||	|||
Q36			||||||
Q37	|||||||||||||||	||
Q38	||	|||||||||||||||
Q39	|||||||||||||	||||||||||

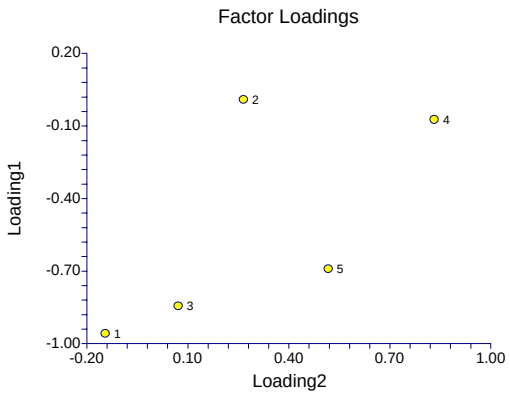
Bar Chart of Communalities after Varimax Rotation

	Factors		
Variables	**Factor1**	**Factor2**	**Communality**
Q35	||||||||||||||||		||||||||||||||||
Q36			||	||
Q37	|||||||||||||		||||||||||||||
Q38			|||||||||||||	|||||||||||||
Q39	|||||||||	||||||	|||||||||||||

Factor Structure Summary after Varimax Rotation

	Factors
Factor1	**Factor2**
Q35	Q38
Q37	Q39
Q39	

Plots Section



Shared Thought Processes

Data unreliable. No factor analysis performed.

Appendix I – Construct Principal Components Analysis

Principal Components Report: Construct Scores

Page/Date/Time 1 7/14/2003 12:58:05 PM
Database C:\Documents and Settings\Br ... ta Analysis\Construct PCA.S0

Eigenvalues

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	5.538857	57.06	57.06	|||||||||||
2	1.757412	18.10	75.16	||||
3	0.717009	7.39	82.55	||
4	0.665949	6.86	89.41	||
5	0.425111	4.38	93.79	|
6	0.300687	3.10	96.89	|
7	0.232006	2.39	99.28	|
8	0.070330	0.72	100.00	|

Eigenvectors

	Factors	
Variables	**Factor1**	**Factor2**
Activity	-0.183031	-0.937694
Coordination	-0.351207	0.013466
Communication	-0.303295	0.093911
Conflict_Resolution	-0.298093	-0.074655
Decision_Making	-0.197518	-0.118960
Joint_Knowledge	-0.487749	0.088794
Personalities	-0.519873	0.288749
Overall	-0.339301	-0.027280

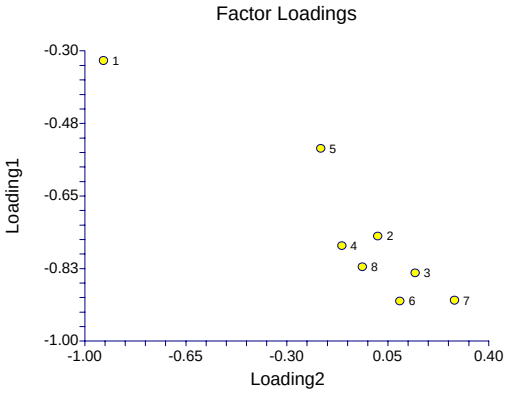
Factor Loadings

	Factors	
Variables	**Factor1**	**Factor2**
Activity	-0.323817	-0.934466
Coordination	-0.746930	0.016132
Communication	-0.836018	0.145813
Conflict_Resolution	-0.770051	-0.108631
Decision_Making	-0.535394	-0.181633
Joint_Knowledge	-0.903934	0.092693
Personalities	-0.901687	0.282102
Overall	-0.820798	-0.037172

Factor Structure Summary

Factor1	Factor2
Joint_Knowledge	Activity
Personalities	
Communication	
Overall	
Conflict_Resolution	
Coordination	
Decision_Making	

Plots Section



Appendix J – Institutional Research Board Protocol

Form 4 - EXPEDITED

IRB Proposal Review #: 03-294

Request for Expedited Approval of Research Involving Human Subjects

RECEIVED
JUN 03 2003
VT Research Compliance

[please print or type responses below]

Principal Investigator (Faculty or Faculty Advisor and primary contact): Dr. Brian Kleiner

Co-Investigators(Faculty or Student) Brent Fraser

Department(s): Industrial and Systems Engineering_ Mail Code: _0118__E-mail: bkleiner@vt.edu Phone 540-231-4926

Project Title: Development of a Methodology for Assessing Military Team Processes__# of Human Subjects 19

Source of Funding Support: ____ Departmental Research _XX_ Sponsored Research (OSP No.: 321539____)

[X] All investigators of this project are qualified through completion of the formal training program or web-based training programs provided by the Virginia Tech Office of Research Compliance.

Note: To qualify for Expedited Approval, the research activities must: (a) present not more than minimal risk to the subjects, (b) not involve any of the special classes of subjects, except children as noted, and (c) involve only procedures listed in one or more of the following categories. The full description may be found in the Expedited Review section of the Virginia Tech *"IRB Protocol Submission Instructions Document"* or 45 CFR 46.110 (http://ohrp.osophs.dhhs.gov/humansubjects/guidance/45cfr46.htm#46.110)

Please mark/check the appropriate category below which qualifies the project for expedited review:

[ ] 1. Clinical studies of drugs and medical devices when proscribed conditions are met [see item (1), page 8 of the "Instructions" document].

[ ] 2. Collection of blood samples by finger, heel or ear stick, or venipuncture subject to proscribed limitations [see item (2), page 9 of the "Instructions" document].

[ ] 3. Prospective collection of biological specimens for research purposes by noninvasive means. Examples: hair and nail clippings, deciduous teeth, permanent teeth, excreta and external secretions, unecannulated saliva, placenta, amniotic fluid, dental plaque, muscosal and skin cells and sputum [see item (3), page 9 of the "Instructions" document].

[ ] 4. Collection of data through noninvasive procedures routinely employed in clinical practice, excluding procedures involving x-rays or microwaves [see item (4), page 9 of the "Instructions"].

[ ] 5. Research involving materials (data, documents, records or specimens) that have been collected or will be collected solely for non-research purposes (such as medical treatment or diagnosis [see item (5), page 10 of the "Instructions" document].

[ ] 6. Collection of data from voice, video, digital, or image recordings made for research purposes [see item (6), page 10 of the "Instructions" document].

[X] 7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language communication, cultural beliefs or practices, social behavior), or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies [see item (7), page 10 of the "Instructions" document].

[redacted]

This project is approved for 12 months from the approval date of the IRB Chair.

rev. 4/27/01

Received 5/10/03 (R)

Contact Information

Dr. Brian Kleiner Industrial and Systems Engineering (0118) 250 Durham Hall Virginia Tech Blacksburg, VA 24061 540-231-4926 bkleiner@vt.edu	Brent Fraser 1907 N Park St Apt 6 Killeen, TX 76541 bfraser@vt.edu

Purpose of Research

The purpose of this study is to develop and demonstrate a methodology for employing naturalistic observation techniques in evaluating team processes. This study will use a military Command and Control (C^2) team for testing and validation purposes. However, the results of this study can be applied with any high-pressure, tactical decision-making team in the military and civilian realms.

The methodology developed in this study will be used to facilitate future understanding of team processes. This study is a first step to developing meaningful, quantitative measures of team process effectiveness. This development path is important because research has yet to develop quantitative measures of effectiveness for teams that is not based solely on outcome-based measures.

At this time, teams are typically evaluated using self-assessment techniques, evaluation by subject matter experts, or based solely on outcomes. Predicting team performance while the team is in action is not currently done, nor is there a clear method for assessing the team's actions and teamwork. This study attempts to form an evaluation methodology that can explore the team's processes.

Currently, if a team is not performing well, it is difficult to determine why they are not performing well. Without specific diagnosis, team interventions cannot be developed, and the team is disbanded or continues performing as they have. If the team processes can be evaluated and quantified, it is possible to correlate them with outcome performance. The evaluation methodology proposed in this study could be used to diagnose team problems and build effective interventions to solve team problems.

Participants

The command and control team will be populated with participants from the Battle Simulation Center's contractor pool. These contractors train the active duty personnel on the use of the systems, and in most cases are retired military officers with extensive experience in command and control. While these individuals are considered experts in command and control, they do not often work collectively as a command and control team. Despite this, it is believed this group of individuals is representative of the active-duty personnel who generally operate in the TOC. Individuals participating as observers will be taken from the contractor pool and from the Army Research Lab: Ft. Hood Field Element personnel. Sixteen individuals will participate in the command team and as many as three will participate as observers.

The age of these participants will range from approximately 30 to 60 years of age. Since military commanders have historically been male, especially in the generations represented by the participants, it is expected that the majority of the participants will be male. However, everyone in the contracting staff

is being given the opportunity to participate in this exercise, regardless of gender. Since this exercise is being used by the Battle Simulation Center as a training exercise for the contracting staff, the participants will not be compensated, monetarily, for their time by this research endeavor. This exercise is considered a part of their employment with the Battle Simulation Center and the participants will be considered "on the clock" during the exercise. Demographic information will be collected from the command team at the end of the first day of the exercise.

A single individual will participate in the distributed observation portion of this study. This participant will be selected from the students who have participated in ISE 5616: Macroergonomics. This participant will be compensated at a rate of $7 per hour and will participate in active observation for the entire four days of the exercise. No qualifications are set for gender or age.

The participants will be informed of the purposes of this experiment prior to the first day of the experiment and will be required to complete an informed consent form before the experiment commences. Since this study will not be the only study performed using this field exercise and the time-sensitive nature of the experiments, the informed consent procedures will be completed prior to the first day of the exercise.

Participant Tasks

The battle simulation being observed in this study will be conducted for four consecutive days from June 10, 2003 to June 13, 2003. Each day's simulation will begin at 08:00 and proceed to approximately 18:00 CST. Lunch hours will be provided, but the schedule is yet to be determined by the Battle Simulation Center's staff.

Those participating in the command and control team will be asked to fulfill the duties of their positions, per the specifications in the training exercise as stipulated by the Battle Simulation Center. Those participating as observers will be asked to evaluate the command team using the developed tool. Observers will be trained in the usage of the tool, including when to make observations and recordings, and what to look for when observing the team. The observers are to continuously observe the team, while it is in operation. At the end of each hour of observation, the observers will be asked to complete the evaluation tool developed in this study, based on the previous hour's observation. The only difference between the local and remote observers is the observation medium. The remote observer will be asked to observe the team by streaming video and audio.

Risks and Benefits

Battle simulations are designed to emulate the stressful conditions experienced in actual battle command and control. This stress is accomplished by having the team play against as live opponent. The knowledge that the team is being evaluated and the competitive spirit found in most military commanders creates a stressful environment. As such, this exercise exposes the participants to a level of stress that might not be experienced in their normal routines. The participants also run a risk of injuring themselves by tripping over power or network cables used to drive the simulation. Ft. Hood's medical facilities will provide any needed medical attention. It should be noted that this exercise would proceed, even if this study were not being performed, since it is being used to train the contracting staff and as an exercise for another, Army Research Lab sanctioned, research study.

The participants will benefit from this study by this project's assessment of the command team's functioning, in the case of the command staff. The participants acting as observers will benefit by gaining

insight in observing how command and control teams function and how to evaluate them. Both of these items are particularly useful, because these participants are typically evaluating the command and control teams.

Confidentiality and Anonymity

No identification will be collected from you beyond your signature on this form and your demographic information collected in the first day's exercise. Due to the nature of the data collection, the team's commander (CO) will complete a specially marked form during each end of the day assessment. However, no names will be collected beyond the signature. You will remain anonymous in this study. Any references made concerning the participants in this study will be by the position occupied in the command group or by the label of "observer" only. No data will be collected linking individuals to the positions occupied.

Audio and video from this exercise will be transmitted, via the internet, to Virginia Tech. This is being done so that an observer can participate from there. There will be no recording made of this. However, it is possible, since it is being transmitted over the internet, that the signal can be intercepted by someone and recorded. Precautions are in place to minimize the risk of this happening, by choosing software that is more difficult to hack into.

Compensation

The exercise described in this study is considered part of the contracting staff's job, and they will be considered "on the clock" during this exercise. Therefore, no compensation will be given to the participants in this study, with the exception of the remote observer. The remote observer will be compensate at a rate of $7 per hour.

Informed Consent

Prior to the commencement of the exercise on the first day, all participants will be given the two informed consent forms for signing. The contents of this form will be read aloud to ensure that everyone agrees to comply with the statements included. Before signing, the participants will be asked if they have any questions regarding this study. All questions will be answered. One copy of the signed informed consent form will be collected from each participant prior to the start of the exercise.

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING (ISE)

Informed Consent for Participants of Investigative Projects

Title of Project: "Development of a Methodology for Assessing Military Team Processes"

Principal Investigators: Dr. Brian Kleiner, Associate Professor, ISE
Brent Fraser, Graduate Research Assistant, ISE

I. THE PURPOSE OF THIS RESEARCH

The purpose of this study is to develop and demonstrate a method for employing naturalistic observation techniques in evaluating team processes. This study will use a military Command and Control (C^2) team for testing and validation purposes. However, this methodology can be applied with any high-pressure, tactical decision-making team in the military and civilian realms. The methodology developed in this study will be used to facilitate future understanding of team processes. This study is a first step to developing meaningful, quantitative measures of team process effectiveness.

II. PROCEDURES

Participants in this exercise will be divided into two groups: players and observers. Players will participate in the battle simulation, as prescribed by the staff of the Battle Simulation Center. Observers will observe the players throughout the exercise and complete the assessment form each hour, based on the previous hour's observation. This assessment form gathers information on how well the players function as a team.

This exercise will run for four consecutive days (June 10 through June 13) from approximately 08:00 to 18:00 (CST), each day. At the end of each hour of the exercise, each observer will complete the assessment tool based on the observations made over that hour. At the end of each day's exercise, the team commander and the section commanders will be asked to complete the assessment form, based on their overall opinion of the team's functioning. Completed assessment forms will be submitted to the researcher at the end of each day.

This is an observation-based study only. The only difference between this exercise and any other battle simulation is that there will be a number of individuals observing how the team operates. You should feel free to ask any questions regarding the tool at any time.

III. BENEFITS AND RISKS OF THIS RESEARCH

This study is being performed in order to develop methods for assessing team functions. The results of this study, including the assessment of the team in this exercise, will be made available to you at the completion of this study. It is hoped the results of this study will aid you in your interaction with the command and control teams that use this simulation center by giving you a greater understanding of how teams function.

Since military command and control operates in a stressful environment, this exercise could expose you to a level of stress that you may not encounter in your everyday life. Elevated stress levels can potentially be hazardous to your health. While the researcher does not anticipate any unacceptably high levels of stress to occur, if you feel you are being stressed beyond your ability to cope, please inform the researcher of your discomfort. Every effort will be made to eliminate your discomfort during this exercise.

IV. EXTENT OF ANONYMITY AND CONFIDENTIALITY

No identification will be collected from you beyond your signature on this form and your demographic information collected in the first day's exercise. Due to the nature of the data collection, the team's commander (CO) will complete a specially marked form during each end of the day assessment. However, no names will be collected beyond the signature. You will remain anonymous in this study. Any references made concerning the participants in this study will be by the position occupied in the command group or by the label of "observer" only. No data will be collected linking individuals to the positions occupied.

Audio and video from this exercise will be transmitted, via the internet, to Virginia Tech. This is being done so that an observer can participate from there. There will be no recording made of this. However, it is possible, since it is being transmitted over the internet, that the signal can be intercepted by someone and recorded. Precautions are in place to minimize the risk of this happening, by choosing software that is more difficult to hack.

V. COMPENSATION

Your participation in this exercise is greatly appreciated. However, there will be no additional compensation for your participation in this exercise, because your participation is part of your obligation to the Army. The one exception is the participant in the remote observation, who will be compensated at a rate of $7 per hour. This individual is being compensated because this individual is not affiliated with the US Army.

VI. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason without penalty. However, this does not free you from any responsibility you may have with the Battle Simulation Center concerning this exercise, as this is out of the researcher's control.

VII. APPROVAL OF RESEARCH

This research project has been approved, as required, by the Institutional Review Board for projects involving human participants at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

VIII. PARTICIPANT'S RESPONSIBILITIES

I understand there are no foreseeable risks involved with participating in this exercise, and I know of no reason why I cannot participate in this study. I, also, agree to notify the investigator at any time about a desire to discontinue participation.

_____

Signature of the Participant

IX. PARTICIPANT'S PERMISSION

Before you sign the signature page of this form, please make sure that you understand, to your complete satisfaction, the nature of the study and your rights as a participant. If you have any questions, please ask the researcher at this time. Then, if you decide to participate, please sign your name above and on the following page. Please, repeat this procedure for your copy, as well.

Participant's Copy

Signature Page

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature _____

Printed Name _____

Date _____

The research team for this experiment includes Dr. Brian Kleiner, Associate Professor, ISE, and Brent Fraser, Graduate Research Assistant, ISE. Research team members may be contacted at the following addresses and phone numbers:

Dr. Brian Kleiner
Industrial and Systems Engineering (0118)
250 Durham Hall
Virginia Tech
Blacksburg, VA 24061
540-231-4926
bkleiner@vt.edu

Brent Fraser
1907 N Park St Apt 6
Killeen, TX 76541
254-628-8328
bfraser@vt.edu

Please, keep this completed copy for your records. Thank you for your participation!

Participant's Copy

Signature Page

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature _____

Printed Name _____

Date _____

The research team for this experiment includes Dr. Brian Kleiner, Associate Professor, ISE, and Brent Fraser, Graduate Research Assistant, ISE. Research team members may be contacted at the following addresses and phone numbers:

Dr. Brian Kleiner
Industrial and Systems Engineering (0118)
250 Durham Hall
Virginia Tech
Blacksburg, VA 24061
540-231-4926
bkleiner@vt.edu

Brent Fraser
1907 N Park St Apt 6
Killeen, TX 76541
254-628-8328
bfraser@vt.edu

Please, submit this completed copy of this form to the researcher. Thank you for your participation!

Vita

Brent D. Fraser

Brent Fraser is a graduate student at the Virginia Polytechnic Institute and State University in Blacksburg, Virginia where he is due to receive his Master's of Science degree in Industrial and Systems Engineering (Human Factors Option) at the end of the Fall 2003 semester. His graduate studies and thesis research have been partially sponsored by the Human Research and Engineering Directorate of the U.S. Army Research Laboratory where he conducted the research presented in this document. Previously he received his Bachelor of Science degree in Industrial Engineering at the University of Arkansas. In his senior year at UA his team won the national material handling design competition sponsored by the College-Industry Council of Material Handling Education (CIC-MHE). Also, his senior design team won the UA- Industrial Engineering outstanding design team for work with Tyson Foods. During and following his undergraduate studies Brent has worked on several projects in Industrial Engineering. This includes a study he performed for the U.S. Postal Service developing strategies for future automation and augmentation of mail handling systems in major processing and distribution centers in a 22 state area. He worked as a Logistics Engineer for J.B. Hunt Corp. where he performed data analysis, logistics support, and headed a project to analyze and optimize one of the corporate departments. Brent's current research interests include the study of organizational team performance employing naturalistic observation methods with the goals of generating quantitative indicators that can be used to better understand team performance. He is also interested in human-centered design processes, in general, and several areas outside the scope of human factors engineering. He is currently a student member of the Human Factors and Ergonomics Society and Alpha Pi Mu, the Industrial Engineering honor society.