The Pennsylvania State University

The Graduate School

College of Information Sciences and Technology

THE DEVELOPMENT OF AN EMERGENCY CRISIS MANAGEMENT SIMULATION TO ASSESS THE IMPACT A FUZZY COGNITIVE MAP DECISION-AID HAS ON TEAM COGNITION AND TEAM DECISION-MAKING

A Thesis in

College of Information Sciences and Technology

by

Rashaad Ennis Theophus Jones

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2006

The thesis of Rashaad Ennis Theophus Jones was reviewed and approved* by the following:

Michael McNeese
Associate Professor of Information Sciences and Technology
Thesis Advisor
Chair of Committee

David Hall
Associate Dean for Research

Guoray Cai
Assistant Professor of Information Sciences and Technology

Susan Mohammed
Associate Professor of Psychology

Joseph Lambert
Senior Associate Dean
Head of the Department of Graduate Program

*Signatures are on file in the Graduate School

ABSTRACT

The research presented addresses two problems evident in the Human Computer Interaction (HCI) and Cognitive Systems Engineering (CSE) literature. First, research has shown that emergency crisis management (ECM) work involves teams and "teams of teams". Yet, there are few models of naturalistic decision systems (i.e., crisis-management) that allow multiple teams to work collectively. Research efforts in modeling the ECM work domain typically involve single team-based simulations. However, few simulations have been designed to study multiple teams operating simultaneously.

Secondly, previous efforts in supporting teamwork have involved the development of decision-aids; however, a prevalent problem with team decision aids is the requirement of rigid and stable information needed to make decisions. The ECM work domain is highly dynamic with incomplete and ambiguous information. Consequently, decision-aids within the ECM domain oftentimes make inaccurate and erroneous recommendations. An alternative design that makes use of contextualized information could be more effective.

The aforementioned problems are addressed by 1.) the development of a simulation studying teams operating in an ECM work domain, and 2.) the conducting of an experiment to investigate the effect a decision-aid has on teamwork. The first objective of the research involved designing a simulation that effectively mimics ECM operations. Hence, the research constructed NeoCITIES as a crisis management scaled world to study team decision-making for a variety of examinations related to HCI and CSE suppositions.

The second objective of the research included the development of a decision-aid that utilizes contextualized information. Specifically, a Fuzzy Cognitive Map (FCM) that handles dynamic and incomplete information was developed to function as a decision-aid supporting teams operating within the NeoCITIES simulation.

Lastly, the research includes an examination of the efficacy of a Fuzzy Cognitive Map (FCM) decision aid on teamwork within NeoCITIES. The study measures whether the presence of a FCM improves decision-making involving resource allocation amongst teams. Experimental sessions using undergraduate students enrolled at Pennsylvania State University provided data to measure team performance and team cognition.

The contributions of the research are threefold. First, NeoCITIES is a technological artifact developed to function as a test-bed to study teams. Secondly, the research assesses the functional use of fuzzy cognitive mapping to serve as a decision-aid for teams. Lastly, in accordance with the interdisciplinary nature of the Information Sciences and Technology program, the present study serves as a model research agenda by developing technologies, examining theories, and incorporating user-centric methods from several disciplines: cognitive systems engineering, human computer interaction, computer science, psychology, and artificial intelligence.

THIS IS DEDICATED TO THE MEMORY OF QUINCY BROWN (1978-2006)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I would like to thank everyone that I have had an opportunity to meet and interact with. Regardless of good or bad, I have been touched by all in a constructive way; I have been motivated by every negative experience and inspired by every good experience that I have had encountered.

I am highly appreciative of my mother and father, Rosa and Willie Jones for instilling the importance of education and emphasizing that I should always give my best in whatever I do. They taught me how to be a successful black man in a “white man’s world” and more importantly, they stressed the importance of having God in my life. I would not be here without their invaluable lessons and principles. I also must express gratitude to all of my family members (aunties, uncles, cousins, and grandparents) for reaffirming those principles!

I would like to thank my twin brother, Rashaan. His no-nonsense approach has kept me grounded during the tough times. He has always been truthful to me regardless of the situation. His words of encouragement and words of correction are priceless!

Many, many thanks and much love to my folks from Morgan State. The “Fam” was there for me during the good times and the bad and they supported me without ever judging me. Specifically, I need to thank my best friend and my right-hand man, Michael Anderson. He has ALWAYS been in my corner and I am indebted for his friendship!

Thanks to my advisor, Michael McNeese for allowing me to turn his idea of a simulation that investigates team decision-making (i.e., NeoCities) into a reality. I am highly appreciative of his support. Additionally, I absolutely love talking about God during my “meetings” with him. Outside of my family, he is the first instance that I have had of a professional that puts his Christian beliefs above his career. I really needed to see that!

I am very grateful for my committee. I appreciate the time and effort they put in contributing to my research. Their insight, support, and advice were very invaluable.

I would like to thank the colleagues in the IST program, specifically Erik Connors. We were brothers in arm and I wish him the best in every aspect of his life.

I would like to thank the IST faculty for their advice and support. Specifically, I am highly appreciative of Dean Hall for going beyond the “call of duty” and opening doors for me that was otherwise closed.

I must give a very special expression of gratitude to the Way family. I consider them my State College-family. They accepted me when it seemed like others would not. I can say for a fact that there is no conceivable way that I would have finished my degree without them. I am so blessed that they “adopted” me into their family and treated me like I was one of their own. Specifically, I must express my sincerest gratitude for my State College-mom, Cindy. She helped me realize my full potential; she helped me understand that my flaws are what make me unique and special. I understand now that a flawed diamond is considered more valuable than a flawless diamond because it is the flaws that make the diamond shine a special way and it is those flaws that make it different from any other diamond.

Lastly, and probably most importantly, I must thank my Creator and Heavenly Father. God placed the aforementioned people in my life because He knew that I would need their guidance, prayers, friendship, and love. Getting a PhD has been a childhood dream and I strongly believe that He provided the right opportunities and people to ensure that I would accomplish my dream.

Chapter 1

Introduction

1.1 Problem Background

Emergency crisis management (ECM) covers a multitude of situations that can arise from a wide range of circumstances including natural disasters (e.g., tornados, wildfires, landslides, thunderstorms, blizzards, and volcanic eruptions) to human-caused mishaps such as mine subsidence, toxic spills, nuclear meltdowns, and even terrorism. Each type of emergency requires decision-makers to take immediate steps to aid response and recovery. Similarly, emergencies call for preventative measures, such as planning and preparedness, as well as sustained efforts to reduce long-term risk to hazards or mitigation. Additionally, decision makers must make appropriate decisions because errors can "snowball" into larger disasters and even become catastrophic (e.g., the disaster at Three Mile Island; see, Woods, Johannesen, Cook, & Sarter, 1994). Furthermore, ECM takes place in a wide variety of disciplines (e.g., emergency medicine, battle management, hurricane centers, piloting, terrorist activities in cities, and bomb threats in schools), emerges in context, and takes different forms across time and space.

Gredler (1994) conceptualizes the ECM domain as having three major characteristics: 1) the high-priority goals of a decision-making unit are threatened, 2) the time available to decision-makers for reacting to situations before they transform is restricted, and 3) high situational awareness from the decision-makers is required. Crisis situations typically exclude extended strategy planning as a means of resolution. Instead, decision-makers are required to accurately assess the situation as quickly as possible while they consider ways to apply available resources. Klein defines this cognitive strategy as recognition-primed decision-making (Klein, 1993). Similarly, other cognitive strategies have shown to be effective (e.g., situated cognition; Hutchins, 1995; Vera & Simon, 1993).

Furthermore, the allocation of resources is also accompanied by periodic reassessments to determine whether changes in response are needed (Gredler, 1994).

Cannon-Bowers and Salas (2001) contend that in a crisis management domain, the nature of work is so complex that it would be impossible for any single team member to hold all the knowledge required to succeed. Rouse, Cannon-Bowers, & Salas (1992) contend that this kind of environment can be turbulent, with time-varying goals and requirements. Further, such environments often require the coordinated efforts of a team operating within an organizational or hierarchical context. According to Rasmussen and Balston, effective organizations must be flexible and adaptive because it is doubtful that coordination "by plan" can satisfy all of the contingencies likely to arise. Moreover, the ECM domain is highly emergent and complex, and requires people to work in teams interacting through various levels of collaborative and communicative technologies. As such, ECM offers a real world domain that serves as an excellent area for examining teamwork and team cognition. Several research efforts have identified the major behaviors of teams working within this type of domain: (1) teams operate effectively when individual members share the workload, (2) team members continually monitor the work behavior of other members, (3) team members develop and contribute expertise on subtasks (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Cooke, Salas, Cannon-Bowers, & Stout, 2000).

Additionally, it is reasonable to classify the ECM work domain as a complex system because Rouse, Cannon-Bowers, & Salas (1992) characterize a complex system as having the following elements: 1.) the environment is highly dynamic 2.) goals vary in time and not infrequently conflict; there are many "acceptable" decisions and no "best" decision. 3.) information is typically incomplete, uncertain, and ambiguous. 4.) multi-person teams are involved 5.) member of teams have differing roles and responsibilities 6.) decision-making is embedded in an organizational context. Complex systems shape behavior in context in ways that may be maladaptive or induce errors. Rouse and colleagues (1992) assert that the operators working within a complex system are effective when team of individuals are able to coordinate action; integrate information and resources, and adapt to changing task demands. They conclude that team dynamics and performance are central in such situations and in order to achieve success in such environments people must continually monitor situations and adjust behavior according to changing task demands (Rouse et. al, 1992).

Alternatively, the ECM work domain can be termed as what Woods, Johannesen, Cook, and Sarter (1994) refer to as a *cognitive system*. Cognitive systems may be identified as large and complex systems that create demands on human activities. Additionally, since individuals are working in a distributed or co-located manner, the ECM domain can be characterized as a *distributed cognitive system* (Hutchins, 1995; Salomon, 1993).

1.2 Problem Description

1.2.1 Distributed Teams

Since the events of 9-11, and continuing with the formation of the department of homeland security, there has been an increased attention on developing technology to support teamwork within emergency crisis management centers (especially in the case to counter terrorist attacks). Recent research has shown that emergency crisis management work involves teams and “teams of teams” interacting through various levels of technological sophistication under certain levels of stress and emotion (McNeese, Brewer, Jones, & Connors, 2006). Teams working together (and a team of teams working together) introduce additional complexities because there exist layers of coordination and collaboration that are necessary to achieve a goal (McNeese & Hall, 2003). For example, activities performed at these layers can be dubious because team members may not share the same assessments of the situation or the teams may be overwhelmed with information, relevant or otherwise. McNeese describes this phenomenon as *cogminutia fragmentosa*; the term describes when a breakdown of the operator's cognition occurs (McNeese, 2002). Once in this state, the user is unable to adapt properly to new information, which results in errors, failures, and potential disasters.

In order to study issues related to teams and teamwork, a definition of teams is provided. Salas, Dickinson, Converse, and Tannenbaum (1992) define a team as a "distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common and valued goal/object/mission, who have each been assigned specific roles or function to perform, and who have a limited life span of membership". Similarly, Mathieu et. al. (2000) define a team 1.) as a distinguishable set of two or more people, each with interdependent and specific roles, who interact and adapt within a dynamic environment towards a common and valued goal and 2.) has a limited life-span of membership. The research conceptualizes the term team as *a distinguishable set of two or more individuals, each member possessing specialized <u>responsibilities and roles</u> working together in a dynamic environment towards a common goal.*

Furthermore, the focus of the research concerns teams that solely work in a distributed manner, which includes team members working co-located or in a completely distributed mode. In order to collaborate, the only communication mediums that these types of teams can use are computer mediated communication (CMC) devices (e.g., telephone, instant messenger, email, video-telephone conferencing, etc). Distributed teams produce more complex issues that impact work in unique ways. For example, information (or shared meanings) can be lost in non-face-to-face communication. A person working with his/her team members in a collaborative fashion must overcome the absence of their team member presence in order to work effectively. Recently, researchers have begun focusing upon studying the performance of distributed teams (e.g., Banks & Millward, 2000; Cooke et al., 2000; Hollenbeck, DeRue, & Guzzo, 2004; McNeese, 2001a, 2001b; McNeese et al., 2005a, 2005b) and the variables that affect their performance, such as team mental models (e.g., Klimoski & Mohammed, 1994; Mathieu, et. al., 2000; Mohammed & Dumville, 2001). Consequently, the definition of teams can be amended to only include a set of two or more individuals, each member possessing specialized knowledge and skills, working together in a dynamic environment towards a common goal *working within a distributed (or co-located) mode.*

Researchers contend that understanding teams is not an easy task. For instance, Cooke and associates (2000, 2004) assert that a team is a special class of groups that presents unique challenges for researchers above and beyond those experienced in group cognition study. Since team members are assigned distinct, interdependent roles, issues arise regarding the concept of shared knowledge. Similarly, Levesque, Wilson, and Wholey (2001) note that "when members have different views of the group process, team member skills, or individual contributions the group experience can be frustrating, time consuming, and ultimately, ineffective," (p. 135). Correspondingly, Salas and Fiore (2004) observe that teamwork "involves the adaptation of coordination strategies through closed-loop communication and a sense of collective orientation so that they can reach [their shared and valued] goals," (p. 4). Interestingly, issues have come to light that point to a degraded ability of distributed team members to structure and transmit information between each other, understand and accurately comprehend those messages, which leads to an inability of distributed teams to respond in a timely and appropriate manner (Daft & Lengel, 1984).

1.2.2 Supporting Teams via Decision-Aids

McNeese (2003) states that the solutions to many of today's complex problems involving teams revolve around the concept of team cognition. Team cognition can be defined as a consequence of team member interaction via communication, coordination, and other process behaviors, all of which cause the transformation of individuals' knowledge into team knowledge that ultimately guides a team's action (Cooke, Salas, Kiekel, & Bell, 2004). Mathieu et al. (2000) stresses that researchers and practitioners should conduct thorough team task analyses to identify the most critical knowledge requirements for a given situation and abstract which parts of knowledge must be shared. Cannon-Bowers and Salas (2001) hypothesize that understanding the concept of shared cognition will further our understanding of what is essential for a team to be effective.

Previous research in cognitive systems engineering (McNeese, 2000, 2002; McNeese & Brown, 1986; McNeese, Rentsch, & Perusich, 2000; McNeese, Zaff, & Brown, 1992; Wellens & McNeese, 1987; Whitaker, Selveraj, Brown, & McNeese, 1995; Wilson, McNeese, & Brown, 1987) has explored a range of team cognition elements and how assorted tools may be adapted to support collaborative activities. Recent efforts combined computer-supported cooperative work (CSCW) and CSE (e.g., Brewer, 2002; Cai, MacEachren, & Bolelli, 2004; Fuhrmann, MacEachren, Dou, Wang, & Cox, 2005) and have just begun pursuing solutions for the emergency crisis management domain, with a strong emphasis on geographical information systems. Similar work has extended CSCW to application domains such as emergency communication/service centers (e.g., A. C. Jones, 2004; Pettersson, Randall, & Helgeson, 2002) and hospitals (e.g., Reddy & Dourish, 2002).

Additionally, numerous decision aids developed to support teamwork (e.g., Rouse, 1988; Sarter & Woods, 1995) have been explored from a wide range of research communities. Those include artificial intelligence (Bates, 92; Magnenat & Thalmann, 1991; Perlin & Goldberg, 1995), decision support systems (Bisantz & Vicente, 1994; Hajdukiewicz & Vicente, 2002; Rasmussen, Pejtersen, & Goodstein, 1994; Vicente, 1999; Vicente & Burns, 1995), psychology (Tolman, 1948; Gibson, 1958; Buzsaki, Horvath, Urioste, Hetke, & Wise, 1992), and HCI (Cooke, 1994; Hoffman, Shadbolt, Burton, & Klein, 1995; Olson & Reuter, 1987; Woods, 1993; Ehrhart, 1994; Andriole & Adelman, 1995), just to name a few.

However, McNeese and associates (2000) asserts that decision aids are limited because they only provide rigid support. The design of these decision aids comes from a techno-centric approach and is oftentimes not informed by contextual and environmental information (i.e., knowledge about teams, user needs, cognitive processes, etc).

Consequently, this results in technologies that are non-adaptive, fail to enact human intentions, and produce “automation surprises” (McNeese, Rentsch & Perusich, 2000; Woods, Johannesen, Cooke, & Sarter, 1994). Interestingly, these problems were noted over 30 years ago when decision aids design was in its infancy (Wohl, 1981).

The view that this research takes is a human-centered approach in designing an intelligent decision aid. The aid makes use of contextualized information, resulting in a new innovative technology that effectively supports team cognition and improves team performance. The research examines the impact such an intelligent decision aid has on team performance. The aid is designed to specifically support teamwork within an emergency crisis management setting.

While there exist a plethora of designs that could be employed for an intelligent decision aid, this research is interested in utilizing a fuzzy cognitive map (FCM). A fuzzy cognitive map (FCM) is a model of the cause and effect relationships that define a complex system (Kosko, 1987). Perusich and McNeese (1998) indicate that the reasoning characteristics of a FCM make it a significant decision support aid for team decision makers. Additionally, the utilization of fuzzy cognitive mapping affords one form of qualitative modeling that is useful for representing contextualized information (McNeese, Perusich, & Rentsch, 2000). Papageorgiou & Groumpos (2004) consider the main advantage of FCMs is their flexibility to system design, model and control. The research investigates the role a FCM intelligent aid has on supporting teamwork. Specifically, the extent of the research is to measure the impact of an intelligent aid designed to facilitate team cognition and assist decision-making and resource allocation in distributed teams working within an emergency crisis management environment.

1.3 Problem Scope

The research limits itself to 3 areas related to distributed teams working in the ECM domain: (1) *distributed teams*, the constraints and factors that affect team performance and team cognition (2) *collaborative technologies*, the socio-techno environment designed to facilitate communication and collaboration amongst distributed teams; and (3) *cognitive modeling tools*, the software modules that support decision-making and team cognition. There are many themes that could be addressed within each area; in regards to distributed teams, the present study is concerned with *team decision-making,* the essentials for a team to reach a decision and the ways they formulate a decision; *distributed cognition,* the cognitive factors and requirements that affect team cognition; *team situational awareness,* the understanding of the team's role and responsibilities; and *team mental model,* a way to represent team knowledge. Additionally, collaborative technologies are divided into *socio-technical systems, scaled worlds, and synthetic task environments*. Lastly, the present study selected a Fuzzy Cognitive Map as its cognitive modeling technique. Hence, Chapter 2 provides a historical review of research efforts in distributed teams and collaborative technologies, and its concluding section provides rationale for the selection of the Fuzzy Cognitive Map.

Chapter 2

Literature Review

This chapter is organized as follows. The first section provides a historical overview of the several themes that involve distributed teams. While various factors exist that could be discussed (i.e., leadership role, activity theory, etc), specifically, the research limits itself to the following areas: team decision-making, distributed cognition, team situational awareness, and team mental models. The second section, collaborative technologies focuses upon the systems developed that have been used to examine teams in an attempt to support teamwork, facilitate collaboration, and enhance team performance. The last section provides an overview of efforts that have been to develop software to augment decision-making and it supplies the reader with justification for using a Fuzzy Cognitive Map for the present study.

2.1 Distributed Teams

2.1.1Team Decision-making

Many important tasks in industrial and military environments rely upon coordinated team efforts and effective distributed decision making (Wellens & McNeese, 1987). Cohen and Thompson (2001) suggest that the proverbial, “two heads are better than one” may provide an advantage for team decision making over individual decision making. Team decision-making involves complementary information and perspectives that help to provide a more complete picture of a given situation, which also leads to alternative solutions. It should be noted, however, that sharing perspectives and alternative solutions may only be considered an advantage if the team practices a suitable method for choosing among alternative solutions (Kerr, MacCoun, & Kramer, 1996).

Individuals often hold unique, expert situational perspectives and knowledge gathered from unique job experiences and tasks (Hedlund, Ilgen, & Hollenbeck, 1998; Larsen, Christensen, Franz, & Abbott, 1998; Lehner, Seyed-Solorforough, O'Connor, Sak, & Mullin, 1997). Oftentimes this individualized knowledge is unshared; which is a phenomena termed as “hidden profiles” in the literature (Stasser, 1992). Team decision making requires the disclosure of unshared information to other team members (Stasser, 1992; Stasser & Titus, 1985). If unshared information is not pooled and disclosed to all team members, the accuracy of a team decision could be drastically decreased (Larsen et al., 1998).

2.1.2 Distributed Cognition

The research conceptualizes cognition from a social-technological viewpoint rather than a social-organization perspective. In other words, this research views distributed cognition by emphasizing the cognitive process teams undergo over the social factors that are present in distributed teams. The research readily recognizes that there are alternative positions to view distributed cognition. For example, the Computer Support Cooperative Work[1] (CSCW) community views distributed cognition by highlighting the social-organization structure of teams over the cognitive processes that teams perform. Instead, the research bounded itself within a cognitive systems engineering and naturalistic decision-making perspective where the emphasis is on technology's impact on the social psychological of teams and the social construction of knowledge.

Cognition in teams has been approached from a variety of perspectives, several of those efforts include team situation awareness (Endsley, 1993, 1995a, 1995b; Wellens, 1993), shared cognition (Cannon-Bowers & Salas, 2001), distributed cognition (Rentsch & Hall, 1994), group sensemaking (Nosek & McNeese, 1997) and transactive memory (Moreland & Myaskovsky, 2000). According to Hollan, Hutchins, and Kirsch (2000), cognition and sources of information are not limited to an individual mind nor are they limited to a collection or group of minds. In accordance with previous research advances in understanding team cognition, Jefferson, McNeese, Theodorou, Ferzandi, and Ge (2004) assert that cognition is situated when it is focused on specific events that emerge through the context of experience within a given environment. Similarly, Lave & Wenger (1991) report "cognition observed in everyday practice is distributed". Hence, the research conceptualizes team cognition in a similar manner described by Eggleston (2002) and Hutchins (1995), the first being *situated* because it relies on knowledge that is generated from actors interacting with other team members and the environment (Bereiter, 1997; Brown, Collins, & Duguid, 1989; Salomon, 1993; Suchman, 1987) and that it is *distributed* because it involves how people constantly recognize situations, acquire and assess knowledge, and use that knowledge to make decisions via asynchronous interactions (Jefferson et al, 2004). The following graphic is provided to summarize the various perspectives of cognition.

---

[1] Computer Supported Cooperative Work (CSCW) can be defined as the development of computer systems to mediate group work and the design of shared interfaces to facilitate collaboration amongst group members to reach a common goal (Eseryel, Ganesan, & Edmonds, 2002).

Figure 2-1

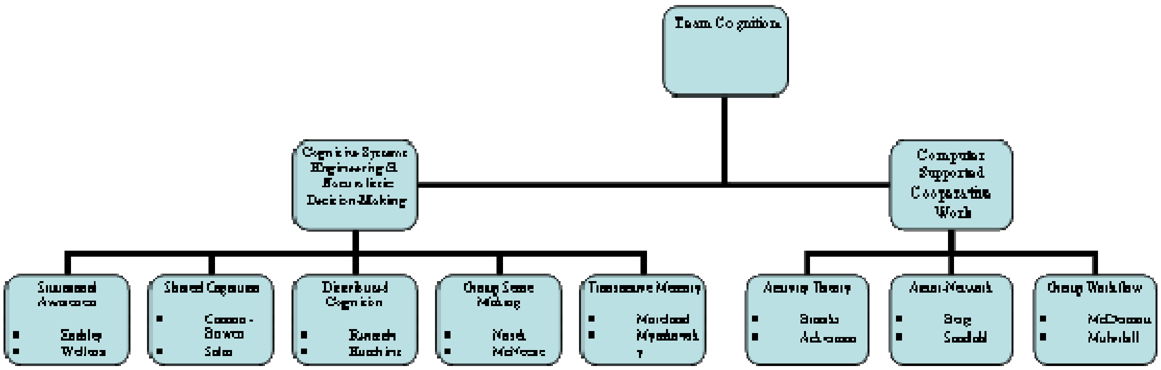


Fig. 2-1: Organizational chart illustrating the bodies of literature that examine cognition

2.1.3 Team Situational Awareness

Team situational awareness has been explored some in the past (Wellens, 1993), but recently has gained the interest for cognitive systems engineering, social psychology, and organizational/team function (Endsley & Jones, 1997; Rentsch et al, 1998). Team SA is defined as “the degree to which every team member posses the SA required for his or her responsibilities” (Endsley, 1995). Bolstad and Endsley (1999) argue that good team SA is needed for effective team performance. Similarly, Wellens (1993) defined group SA as “the sharing of a common perspective between two or more individuals regarding current environmental events, their meaning, and projected future status.” (p. 6). Likewise, shared SA is defined as “the degree to which team members posses a shared understanding of the situation with regard to their shared SA requirements,” (Endsley & Jones, 1997) and it is an important aspect of team SA (Bolstad & Endsley, 1999). The following graphic is provided to illustrate how the terms (shared, group, and team) characterize situational awareness in similar ways.

Figure 2-2

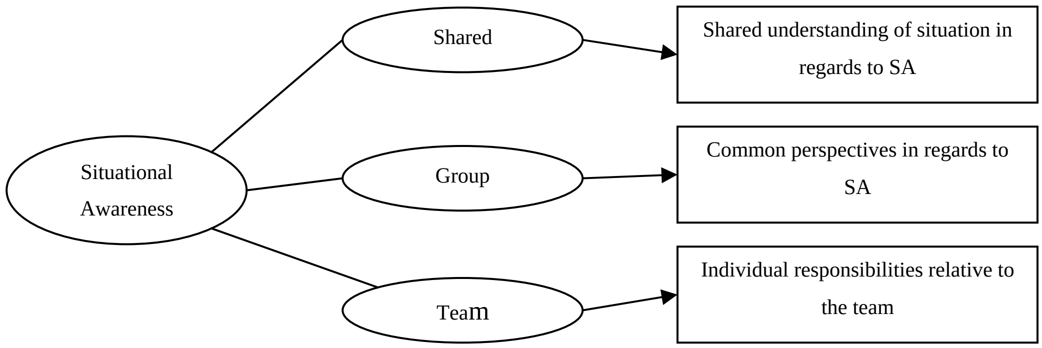


Figure 2.2: 3 different ways to characterize SA

Fig. 2-2: Characterization of SA

Bolstad & Endsley provide a framework (1999) that characterizes shared SA as involving how team members share requirements (which information needs to be shared), the devices available for sharing SA (i.e., direct communication, shared displays, collaborative environments, etc), mechanisms for sharing SA (the degree to which team member possess team mental models to interpret information), and shared SA processes for questioning assumptions, manage conflicting information, coordinating tasks, and enacting contingency planning within a socio-technical system. Correspondingly, other efforts have been utilized to examine shared SA within an ecological framework (McNeese, 1996; McNeese, Rentsch, Perusich, 1999; 2000) that utilizes a socio-technical system.

The position taken in this dissertation research is that shared SA is essential to effective team performance because it necessitates that each member has an understanding of their role and responsibilities and how their skills and knowledge are applied to the overall team effort.

2.2 Team Mental Models

One of the major constructs employed in contemporary team cognition research is that of the mental model. A seminal paper that discusses mental models is Rouse and Morris (1986). They define a mental model as, “a mechanism whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states and predictions of future system states.” Rouse and associates conceptualize a team mental model as the mechanisms that underlie the formation of expectations and explanations, within are encoded knowledge needed to perform the task (1992). Several researchers have hypothesized that team knowledge can be represented in the form of team mental models (Langan-Fox, Code, & Langfield-Smith, 2000; Hanisch, Kramer, & Hulin, 1991; Johnson-Laird, 1983; Wilson & Rutherford, 1989). Orasanu & Salas (1993) have explored the extent to which individuals share and construct similar mental models when interacting as a team. Langan-Fox et al. (2000) hypothesize that more effective team performance requires the existence of a shared or team mental model. They assert that a team mental model must encompass the perception, encoding, storing, and retrieval mechanisms of each individual. Similarly, others emphasized the notion of "shared" or "team" mental model, as the development of a common representation of a team’s situation that leads to mutual expectation, and thus improved coordination among team members (Cannon-Bowers & Salas, 1990). They proposed that team members form mental models of the external environment, the team environment, teammates, and the task, and that these are organized in a hierarchical manner. Similarly, Rouse and colleagues (1992) argue that team mental models are most useful for highly structured tasks where team behavior must conform to task requirements. They contend that by explicitly identifying mental models, an understanding of team performance will be gained.

The concept of shared mental models has different names in the research community. The list includes: group situation awareness (Wellens, 1993), team schema similarity (Rentsch & Hall, 1994), intersubjectivity (Levine, Resnick, & Higgins, 1993), collective mind (Weick & Roberts, 1993) and transactive memory (Wegner, 1987, for a review see Klimoski & Mohammed, 1994). Correspondingly, each concept theorizes that group members typically have some sort of organized knowledge structures relating to various aspects of the group's situation, such as their task, their interaction process, their environment and their fellow group members. The development of shared understanding (often referred to as *common ground*) helps group members to predict future actions and work together in a coordinated way (Cannon-Bowers, Salas, & Converse, 1993; Kraiger & Wenzel, 1997; Moreland, Argote, & Krishnan, 1996).

It is important to note that the conceptualization of team mental models should not be confused with shared SA. There is a high degree of interrelationships between them as they deal with similar concepts and phenomena pertaining to team cognitive activities, but the distinction is made in their representation of team knowledge. Knowledge about the task and the environment are chief elements for both, however shared SA generally involves the projection of a future status while team mental models typically includes expectations of team members role. The following graphic is provided to illustrate the similarities and differences in shared SA and team mental models.

Figure 2-3

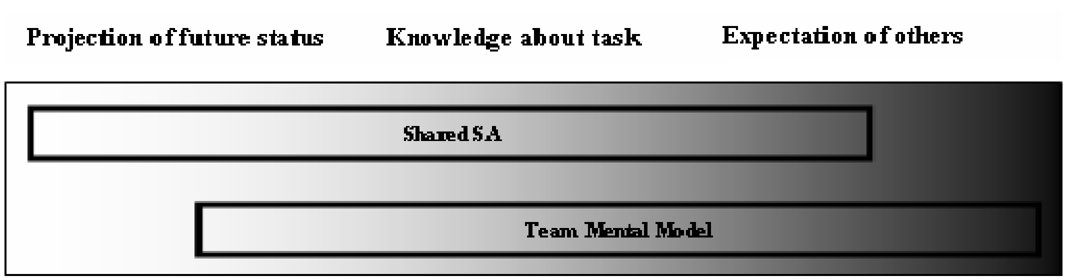


Fig. 2-3: Shared SA vs team mental models

According to Cannon-Bowers and associates (1993), shared mental models help explain how teams are able to cope with difficult and changing task conditions. Additionally, some researchers have suggested that the ability to adapt is essential in teams (Cannon-

Bowers, Tannenbaum, Salas, Volpe, 1995; McIntyre & Salas, 1995). The shared mental model theory (see Mathieu et al., 2000) provides an explanation of the mechanisms of adaptability, in other words, how teams quickly and efficiently adjust their strategies. The theory contends that in order to adapt, team members must predict what their teammates are going to do and what they are going to need in order to do it. Consequently, the shared mental model theory postulates that the function of shared mental models is to allow team members to draw on their well-structured knowledge as a basis for selecting actions that are consistent and coordinated with those of their teammates. Evidently, the conceptualization of team mental models is tied to the aforementioned categories of team knowledge--each element of team knowledge is a foundational component of the construct of team mental models.

Moreover, in a complex environment with events that emerge over time, team members may experience difficulties communicating and coordinating effectively. Stout, Cannon-Bowers, and Salas (1996) argue that a shared mental model becomes crucial when team members are unable to communicate and strategize freely, which could be due to excessive workload, time pressure, or other environmental issues. Thus, a shared mental model is posited as an effective mechanism because it allows members to predict the information and resource requirements of their teammates.

As such, the research examines the decisions of each member and the factors resulting in the decision that was made. Additionally, the research collected data that deals with team member’s expectations of others. These measures yield an assessment of the quantity of shared knowledge and the factors that resulted in any incongruities and inconsistencies. The following section summarizes techniques to measure team mental models and it describes the measurement that the present study utilized.

2.2.1 Historical Overview of Team Mental Model Techniques and Measurement

There exist a myriad of techniques that can be used to develop and measure a team mental model. Some techniques for measuring individual mental models extract the model from text provided by the subject; these techniques include content analysis, procedural mapping and cognitive mapping (Carley, 1997). Alternatively, in other techniques, subjects are asked to impose an organization on task-related stimuli provided by the researcher, such as multidimensional scaling (Kraiger & Wenzel, 1997), or to provide a numerical assessment of the strength of various relationships in the model (belief association matrix; Hinsz, 1995). Some techniques allow quantitative scores of congruence or discrepancy between individual mental models to be computed (Kraiger & Wenzel, 1997) which result in an overall score for the team. However, an alternative approach to studying shared mental models would be to examine behaviors that take place during team preparation and performance that reflect the encoding or the use of shared mental models. For example, Moreland and his colleagues have developed a program of research on transactive memory which uses this approach (Liang, Moreland, & Argote, 1995; Moreland, Argote, & Krishnan, 1996; Moreland et al., 1998; Rulke & Rau, 1997).

Interviews and surveys are excellent for gaining an understanding of task-related concepts (Cooke, Salas, Cannon-Bowers & Stout, 2000). Langan-Fox and Code (2000) cite several examples of cognitive interviewing techniques: open interviews, question-answered interviews, and inferential flow analysis. Several researchers have used this technique to elicit domain, or problem-solving concepts and the relationships between them (Cavaleri & Sterman, 1997; Lange & Burroughs-Lange, 1994; Redding & Cannon, 1992). A transcript of the interview is constructed and the researcher analyzes the text using propositional or discourse analysis and the result is a graph that represents domain concepts and conditional and causal associations among them (Langan-Fox et al., 2000). While cognitive interviews serve as an ideal method to extract conceptual information, several drawbacks are mentioned by Langan-Fox and Code (2000): it is significantly constrained by the interviewer's interpretation, which is typically predicated by their understanding; additionally, interview techniques are unable to capture tacit information; and lastly, interviews conducted after task performance are subjected to retrospective distortions.

Another method to form a team mental model that is referenced by Langan-Fox et al. (2000) is verbal protocol analysis. Ericsson and Simon (1984) have stated that this technique is primarily useful to obtain information about decision-making strategies and general reasoning processes. Rouse and Morris (1986) has described this as a "verbalization method" and it is an approach that provides appropriate descriptions of explicit tasks, such as solving a distance-time problem or following instructions to assemble a piece of furniture. Langan-Fox et al. (2000) report that this technique will yield sets of production rules, decision trees, heuristics, etc, that enables the researcher to identify the relationship between objects, or concepts within a domain. A limitation to this approach is that the mental model will be constrained by the subjects' inability to accurately verbalize their intentions and thoughts (Langan-Fox et al., 2000; Rouse & Morris, 1986); however, this technique is recommended when researching issues related to team decision-making because verbalization is a natural part of task performance (Langan-Fox et al., 2000). Similarly, Cooke et al. (2000) term this as a "process tracing technique" and it is very useful to elicit data in a dynamic setting because of its online nature. As mentioned earlier, there are many methods that provide measurements of team mental models and while these techniques are very interesting, a complete review of these techniques are outside the scope of this research. Accordingly, the following table is provided to give a review of the different techniques mentioned, it also provides recent examples of research efforts utilizing the respective techniques, and it reports the strengths/weakness of each approach.

Table 2-1

Table 2-1: Review of mental model measuring techniques

Technique	Synopsis	Examples	Strengths/Weakness
Cognitive Interviewing Techniques	Concepts relevant to the domain are elicited through open form, question-answer form, or inferential flow analysis.	Lange & Burroughs-Lane (1994) Redding & Cannon (1992)	*Strengths.* Straightforward and systematic. Yields a complete representation of domain. *Weakness.* Captures only information that can be expressed verbally. Relies heavily on interviewer's interpretations.
Content Analysis	A systematic method for analyzing written or verbal statements.	Langan-Fox & Tan (1997)	*Strengths*. Systematic; not constrained by availability of respondents. *Weakness.* Incomplete representation.
Participant Observation of a Task	Researcher observes participant during task and draws inferences about critical domain concepts and the relationship between them.	Chen (1996) Vandenplas-Holpher (1996)	*Strengths.* Highlights causes of user error. *Weakness.* Dependent upon the researcher to identify important concepts within the domain.
Visual Card Sort Technique	Participant lists all the concepts that are relevant to the domain of interest, which are written on notecards. The participant then cluster cards according to the strength of relationship.	Von Hecker (1997) Carpendal, McBride, & Chapman (1996)	*Strengths.* Unintrusive, easy to administer, and flexible. *Weakness.* Most likely represents information stored in short-term memory.
Repertory Grid Technique	Participant is provided with three concepts at a time and is asked to describe how two are alike and different from the third.	Valinejad (1997) Langan-Fox & Tan (1997)	*Strengths.* Grounded in theory *(*good and reliable). *Weakness.* Can involve subjective interpretation. Time intensive.
Multi-Dimensional Scaling	A matrix of pairwise similarity ratings of all the concepts are analyzed and placed in the space of dimensionality specified by the researcher.	Streveler (1994) Rusbult, Onizuka & Lipkus (1993) Roth, Gabel, Brown, & Rice (1992)	*Strengths.* Yields a pictorial representation of how the teams are clustered. *Weakness.* Since there are a number of variations of scaling to choose from, the most appropriate technique is not always easy to identify.

2.2.2 Team member schema similarity

The present study measures team mental models through the use of a team member schema similarity construct (Rentsch & Hall, 1994). Team member schema similarity, or TMSS, refers to the degree to which team members' team-related schemas, or their understanding of team-relevant information is similar (McNeese, Rentsch & Perusich, 2000). TMSS has two measures: *schema agreement (also known as congruence),* which is the overlap of schema content and structure that exists among teams and *schema accuracy,* which refers to the degree to which team members' descriptions of each other's schemas reflect the other team members' actual schema (McNeese, Rentsch & Perusich, 2000). Additionally, Rentsch and colleagues (1994, 1998, 1999, and 2000) contend that high teamwork TMSS enhances team performance. High TMSS will allow team members to interact efficiently and effectively with one another, thus affording them the ability to anticipate, facilitate, and compensate for each other actions (McNeese, Perusich, & Rentsch, 1999; 2000). Research efforts within military domains have gathered empirical evidence to support the direct relationship between TMSS and team performance (Rentsch, Burnett, McNeese, & Pape, 1999; Rentsch, McNeese, Pape, Burnett, Menard, & Anesgart, 1998).

2.2.3 Summary

The research addresses the call for research involving team mental models found in the literature. For example, Smith-Jentsch, Campell, Milanovich, and Reyonlds (2001) report research involving *team mental model measurement* - as related to effective team performance - is in its infancy. Similarly, Langan-Fox et al. (2000) report that there are few frameworks and taxonomies of team mental models. Additionally, they pronounce that there is a need to clarify the nature of the relationship between team mental models and performance. Likewise, Mathieu et al. (2000) note that while researchers have proposed many variables that influence team performance, there has been little empirical examination to substantiate those claims. Furthermore, Kraiger and Wenzel (1997) suggest that decision-making and communication are areas that can be greatly impacted by shared mental models. Hence, the crisis management domain is an excellent area for observing the practical constraints and benefits of a team mental model because it involves highly emerging events and ill-structured information requiring people, each with specialized roles and knowledge, to interact as a team, acquiring, assessing, and sharing information to make effective decisions. In turn, the research examines the decisions of each member and team member's expectations of others. These measures yield an assessment on the quantity of shared knowledge and the factors that resulted in any incongruities and inconsistencies.

2.3 Collaborative Technological Environments

2.3.1 Overview

While it is often desirable to study teams in their operational environments, both practical and methodological difficulties often preclude this approach (Wellens & McNeese, 1987). Thus, collaborative technological environments are important because it provides a way to study teams by simulating their operational environment, while maintaining the control associated with traditional laboratory research. A brief description of three studies is mentioned to serve as examples to orient the reader on collaborative technological environments used within the distributed decision-making paradigm.

2.3.2 TRAP

The Team Resource Allocation Problem (T.R.A.P.) was developed as a tool to study dynamic group decision making (Brown & Leupp, 1985). T.R.A.P. effectively simulates a command, control, and communication environment (C3) and it involves the experimental control to measure decision-making. The T.R.A.P. experiment involved three-person teams working together to accumulate information by processing "tasks" that were displayed on a color video screen. Conclusions are summarized as effective team performance is dependent upon the development of a shared group awareness of task demands and decision strategies (Snyder, Wellens, Brown, & McNeese, 1989).

2.3.3 CITIES

The C3 Interactive Task for Identifying Emerging Situations (C.I.T.I.E.S) was developed to test a series of hypotheses involving Wohl's model of information fusion and decision-making (Wellens & Ergener, 1988; Wohl, 1987). The C.I.T.I.E.S tasks effectively mimic an ECM center and it involves remotely located teams to pool information derived from various sources to formulate decisions relating to complex resource allocation problems (Wellens & Ergener, 1988). Results suggest that a feeling of "teamness" is obtained when the bandwidth (or richness) of the communication channel is increased (Wellens & McNesse, 1987).

2.3.4 DDD

Possibly one of the more popular simulations is The Distributed Dynamic Decisionmaking (DDD) scaled world. It is a distributed multi-person simulation and software tool for understanding decision-making in a dynamic team environment (Kleinman & Serfaty, 1989; Kleinman, Pattipati, Luh, & Serfaty, 1992). DDD is a very flexible scaled world in that it provides a high level of experimental control to vary elements of the environment (MacMillan, Entin, Hess, & Paley, 2004). As such, the DDD task has been used as a test-bed for various command and control team performance tasks for the past 3 decades.

As one can see from the examples, a collaborative technological environment can be thought of as a kind of virtual environment that is used to examine teams by simulating an operational setting that accurately reflects the work domain of interest. Additionally, collaborative technological environment provides a degree of experimental control that is usually not found in a natural setting; this is especially true when teams are being studied. The following section conceptualizes a virtual environment and it provides an overview of the various types of virtual worlds.

2.3.5 Virtual Environment

Virtual environments are rapidly becoming the standard technological artifact when it involves examining human performance. Efforts to apply virtual reality technology can be seen in various fields: the medical profession (Stytz, Frieder, & Frieder, 1991); visualization of complex data sets (Defanti & Brown, 1991); stock market prediction and analysis (Coull & Rotham, 1993); military training tools (Dix, Finlay, Abowd, & Beale, 1993); and human factors (Scott, 1991). Stanney, Mourant, & Kennedy (1998) envision virtual environments as being systems that can enhance the communication between humans and computers; they posit that human factors research involving virtual environments will lead to advancement of technologies that are better suited to meet the needs of its users. An advantage to using virtual environments in human factors related research is that the knowledge of how to perform a task is embedded in the contextual environment (Greeno, Smith, & Moore, 1993) and human users can use the environmental contextual cues to support and mediate their task actions (Druckman & Bjork, 1994). Additionally, virtual environment systems capitalize on fundamental and distinctively human sensory, perceptual, information-processing, and cognitive capability (Stanney, Mourant, & Kennedy, 1998). There are several types of virtual environments (or virtual worlds) that have been employed for research, namely: *simulations*, *scaled worlds*, and *synthetic task environments*. While they are similar in appearance and use, there are practical advantages and disadvantages. The following is an overview of each and a concluding section comparing the trade-offs.

2.3.5.1 Simulation

Simulations address questions regarding performance in specific situations (Humphrey, Hollenbeck, Ilgen, & Moon, 2004). The chief advantage provided by simulations is their ability to contain a high degree of realism. It is for this reason that simulators are used; their capability to accurately represent an abstraction of the real world makes it very feasible for training applications. Car and flight simulators are common examples of simulators that are used for training. In regards to research, Raser (1969) outlined 4 criteria for determining the validity of simulations as research tools: (1) psychological reality, (2) structural validity, (3) process validity, and (4) predictive validity. However, while Humphrey and associates assert that simulators are viable research tools because of their high fidelity, they also contend that simulators are usually limited by low number of subjects and difficulty in controlling multiple environmental effects (2004).

2.3.5.2 Scaled Worlds

Scaled worlds are another type of virtual environment. Scaled worlds are useful in simulating real world complex problems that are derived from or highly related to situated activities. For example, McNeese and associates (1996; 1999; 2004; & 2005) have cited several examples of using scaled worlds to examine socio-organizational constraints, modeling teamwork processes, and measuring team schema development. A chief advantage of scaled worlds is their ability to "scale-up" to meet the demands of the work domain. In other words, scaled worlds can run numerous participants simultaneously. As such, scaled worlds have been cited in the literature as being a viable method to study and model teams (Marks, Mathieu, & Zaccaro, 2004). Lastly, scaled worlds focuses on the mutual interplay of understanding, modeling, and measuring teamwork within complex systems (McNeese, Perusich, & Rentsch, 2000).

2.3.5.3 Synthetic Task Environments

Synthetic tasks environments (STE) are a type of scaled world that house research tasks constructed by systematic abstraction from a corresponding real-world task (Cooke & Shope, 2004). STEs are designed to mimic complex multi-role team tasks but are scaled to be distributed using more standard technology (Elliot, Dalrymple, Regian, & Schiflett, 2001). STE systems can create complex environment for networked multi-operator training and performance research and this environment has been cited in the literature as a research platform that bridges the gap between controlled studies using artificial laboratory tasks and uncontrolled field studies on real tasks or using high-fidelity simulators (Cooke& Shope, 2004). Lastly, as operational analogues, they can be distinguished from other PC-based “gaming” environments by three core features: (a) operational relevance, (b) experimental control, and (c) extensive tracking and assessment of operator activities as well as outcomes (Elliot, Dalrymle, Schiflett, Miller, 2004).

2.3.5.4 Summary

This section provides a summary by comparing the different types of virtual worlds that are mentioned. Simulations generally provide high fidelity virtual worlds, however there is a hindering tradeoff. Simulations often replicate the environment at the expense of the simulation’s flexibility as a research tool. As such, with simulations, researchers are oftentimes limited in the degree to which they can alter or control the simulation and the measures that they can derive from it. Recently, researchers have cautioned against the use of simulations unguided by training principles or an understanding of the actual task requirements and have extolled the virtue of low-fidelity simulations that take such factors into account (Cooke & Shope, 2004).

Unlike simulations, STE systems are expected to enable more controlled investigations of complex, multi-operational performance (Driskell & Salas, 1992). They are developed through functional and cognitive task analyses of an operational performance domain, and thus are analogues of particular aspects of the performance domain (Schiflett & Elliot, 2000). As such, STEs can be used as a research tool provides university-based researchers with a tool that increases the operational relevance of performance research, while allowing strong experimenter control (Elliot, Dalrymle, Schiflett, & Miller, 2004). The effectiveness of an STE depends on appropriate scaling of the operational performance domain to assure core functions and task characteristics are maintained at proper levels of complexity (Elliot, Schiflett, Hollenbeck, & Dalrymple, 2001). The following graphic is provided to illustrate the differences between simulations and synthetic task environments.

Figure 2-4

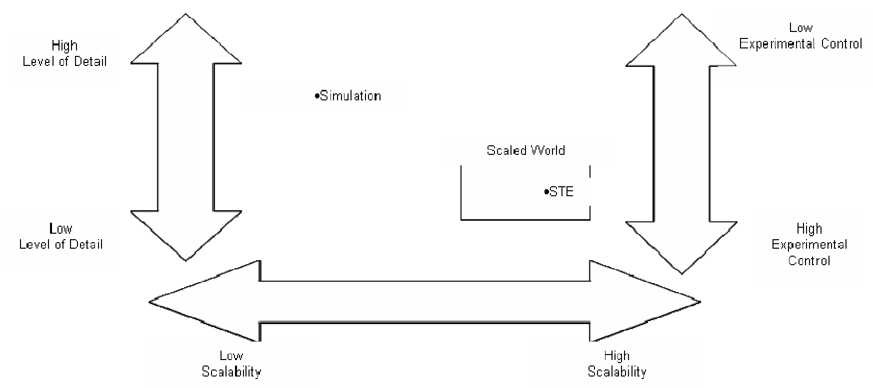


Fig. 2-4: Simulation vs. Scaled Worlds vs. STEs

2.4 Naturalistic Decision-Making

Naturalistic decision making (NDM) is described by Orasanu and Martin (1998) as "decision making by individuals with some level of domain expertise in real world contexts (e.g., aviation, nuclear power, offshore oil process control, fire fighting, command and control)". NDM occurs typically when limited time, dynamically changing conditions, goal conflicts, and information sources of varying reliability are present. In these cases, decision makers may operate in team and organizational contexts, and have available tools or other information resources available to aid their decision making (Orasanu & Connolly, 1993). NDM typically involves recognizing that a problem exists and sizing up the situation to define the nature of the problem and relevant factors (Orasanu & Martin, 1998). A candidate solution is retrieved, evaluated and applied if it meets a criterion of adequacy (Klein, 1997). Zsambok, Beach, and Klein, (1992) characterizes the following elements as being vital to NDM:

- situation assessment in addition to option selection
- single option construction and modification (versus generating many options for comparison purposes)
- single option evaluation (versus comparing multiple options to themselves or to a standard)
- changing conditions and ambiguous information versus stable conditions and information within the decision event
- shifting goals versus stable goals within the decision event
- time constraints in deciding what to do
- previous experience by the decisionmaker in the decision event

By recalling the characterization of an ECM by Gredler (1994) and Rouse, Cannon-Bowers, & Salas (1992), it is reasonable to suggest that NDM is inherent in an ECM work domain. In fact, this research posits that the ECM work domain is an ideal setting to study NDM. The following graphic is provided to clearly convey the similarities:

Table 2-2

Table 2-2: ECM vs NDM

	ECM (As defined by Gredler, 1994 and Rouse, Cannon-Bowers, & Salas 1992)	**NDM** (As defined by Zsambok, Beach, and Klein, 1992 and Orasanu and Martin 1998)
Goals	Vary in time	Shifting
Work environment	Dynamic	Changing conditions
Time	Restricted	Constrained
Information	Incomplete	Ambiguous
SA	High	High
Work structure	Teams and organizations	Teams and organizations

The following section is provided to describe various models of decision-making that follow NDM principles. These models of naturalistic decision-making attempt to address how people can make decisions in situations where the conditions are changing over time, where information is ambiguous, and where the plausibility of potential goals and courses of action is shifting over time. The following section does not attempt to cover all models of NDM; it merely lists a few that have been cited in the literature to give the reader an indication of the characteristics of NDM models. Additionally, the following section serves to set a foundation on NDM modeling principles, and then it introduces a new NDM modeling technique, namely a Fuzzy Cognitive Map.

2.4.1 Models of Naturalistic Decision-making: Recognition-Primed Decision-making

Klein (1989; Klein, Calderwood, & Clinton-Cirocco, 1986) has developed a model of Recognition-Primed Decision-making (RPD) that describes how experienced people commonly make decisions in their operational settings. The RPD model is perhaps the most popular NDM model found in the literature. For ease of description, the model differentiates between highly familiar and moderately familiar situations. However, it does not presume that this dichotomy exists as a psychological reality--it assumes that familiarity with situations is a continuum (Zsambok, Beach, & Klein 1992). The model describes how decision-makers are able to draw upon their experience to assess situations and to arrive at a course of action. Based on observations from five field studies in different domains such as firefighting and tank platoon maneuvers, Klein et al. (1986) found that commanders were often able to quickly size up the situation, arrive at a course of action to deal with it, and modify the course of action as necessary to accommodate changes in the situation (Zsambok, Beach, & Klein 1992).

Klein et al. (1986) found in situations that are highly familiar, that most often, experienced decision-makers recognize it being a situation that they have seen previously. For example, a fire chief arrives on the scene of a house fire. He perceives a number of cues from the environment: smoke coming from under the eaves of a pitched roof, a red flame shooting out the attic window, and a yellowish flame forming at an adjacent second-story window. The model postulates that these perceptions cue the chief's memory for other similar situations he has seen or for a prototypical instance which is the amalgamation of many such situations he has seen (Zsambok, Beach, & Klein 1992). In other words, remembered situations are presumed to contain information about additional *critical cues* to look for; about feasible *goals*; about *typical actions* and about *expectancies* (Zsambok, Beach, & Klein 1992).

2.4.2 Models of Naturalistic Decision-making: Noble's Cognitive Model for Situation Assessment

Noble (1989) developed a cognitive model to describe the situation assessment portion of a decision event. Through a series of experiments, Noble's model was able to capture the expertise which Navy operators use when they resolve report-to-track problems--when they localize and identify ships from a sequence of situation reports. (Noble, Boehm-Davis, & Grosz, 1986). In Noble's model, each type of previously experienced problem (or decision event) is treated as if it is stored in memory as a separate reference problem. Noble states that memory storage may not actually correspond to a structure consisting of reference problems and their solutions, but he argues that other plausible structures, such as those containing prototypes, would be functionally similar (Noble, 1989). Reference problems contain information about context, goals, solution methods, and other information useful for adapting these solution methods to future problems. He identifies this view as similar to the RPD model, in which it is proposed that goals, expectancies, cues, and actions are contained in memorial representations of experienced situations (Zsambok, Beach, & Klein 1992).

Compared to Klein's RPD, Noble's model is both deeper and narrower than the RPD model--deeper in that it specifies greater detail about how simple RPD (feature matching) takes place, and narrower in that it does not attempt to describe complex processes like modification or evaluation of a course of action (Zsambok, Beach, & Klein 1992).

2.4.3 Models of Naturalistic Decision-making: Image Theory

Image Theory (Beach, 1990) attempts to explain a wide variety of decision behaviors. Similar to RPD, the Image Theory presupposes that a decision-maker uses features of the context (the stimulus situation) to probe memory. If the stimulus features of the current context are virtually the same as memorial features, the current context is said to be recognized. However, if stimulus features only resemble them, then the memorial information and all that is associated with it constitute an ad hoc definition of the current context, and it is said to be identified.

2.4.4 Models of Naturalistic Decision-making: A Skill/Rule/Knowledge-Based Model of Cognitive Control

Rasmussen developed a NDM model that involves all portions of peoples' tasks, including sensory-motor behaviors that operators perform unconsciously and effortlessly, as well as those that require conscious deliberation (Zsambok, Beach, & Klein 1992). Consequently, his model is slightly different, but yet still compatible from the authors already mentioned. Rasmussen's model is derived from analyzing accidents and simulated decision-making of experienced operators of complex automated systems like power plants (1983; 1986). Based on analyses of these accidents and simulations, Rasmussen developed a model of cognitive control that includes three control levels: skill-based, rule-based, and knowledge-based. Familiarity with a situation (and level of expertise) will determine the level of cognitive control that is exercised and, thereby, the nature of the information used to control the activity and the interpretation of observed information (Rasmussen, 1988). This model reflects his view of "practical decision-making" in which situation diagnosis and action are intimately connected (Zsambok, Beach, & Klein 1992).

Skill-based control represents the highly skilled sensory-motor performance controlled by automated patterns of movements. This type of control is exercised by masters, or people with a high level of expertise. Skill-based control is like a program that runs without conscious attention, freeing the person to think of other things. The next level is rule-based control, characterized by consciously controlled actions. A rule is a procedure or subroutine stored in memory that prescribes actions for a particular situation. For example, while engaged in automatic skill-based control, a person may experience ambiguity or deviation in the environment compared to their internal world model of what the environment ought to be, given this familiar situation (akin to RPD). The next level of control is knowledge-based, which also involves conscious decision-making. This level of control is necessary if people's goals change during a decision event, or if they enter an unfamiliar situation where know-how and rules for control are not available. Here, the control must move to a higher conceptual level in which performance is knowledge based and goal controlled. The goal must be explicitly formulated (or reformulated, in the case of changing goals), based on an analysis of the environment and the overall aims of the person. According to this model, in unfamiliar situations, the person develops different plans to reach the goal.

Conceptually, there are stark similarities between Klein’s RPD model and Rasmussen’s model. The main difference is that Rasmussen highlights situations in which people create multiple plans and select the best one, whereas the RPD model depicts single option evaluation.

2.4.5 Models of Naturalistic Decision-making: Summary

While there are many differences in each of the NDM models discussed, it is necessary to abstract characteristics that they share. An abstraction of similar characteristics represents the essential activities that can be performed in a naturalistic decision-making process. Thus, these similar characteristics can be considered "core elements" to naturalistic decision-making. Consequently, the core elements can be used as the foundation for new NDM models. It is important to note that the research does not suggest that <u>every</u> NDM model must contain all (or some) of these elements. The abstraction is not a prescription, but more of a proposition, which involves commonalties found within several NDM models. One aim of this research is to use the following proposition to inform the design of a new NDM decision-aid model.

First, each model involves a decision-making process that can be divided into 5 sub-processes. The first sub-process is a *recognition task.* This process involves receiving new information, which can come in various forms (i.e., a report from an agency or an emergency like a house is set on fire). The second sub-process is an *assessment of the situation.* An assessment of the situation also includes an understanding of the environment and a projection of the future status (SA). It also includes an *identification* and/or *diagnosis* of the current situation and how it fits within the context (i.e., the fire appears to be a level 3 fire with a potential for high causalities). The third sub-process is an *analysis* routine that involves a reasoning mechanism that compares the current situation to what is already known. It also contains a formulation of a set of goals that are needed to address the current situation. The fourth sub-process transforms goals into a plan by decomposing the goals into tasks. Lastly, the fifthprocess is an action space that performs the tasks outlined in the plan.

Formally, the foundational elements are: 1.) Recognition 2.) Diagnosis/Identification 3.) Analyze 4. Plan/Evaluate and 5.) Act. The following graphic is used to convey the process.

Figure 2-5

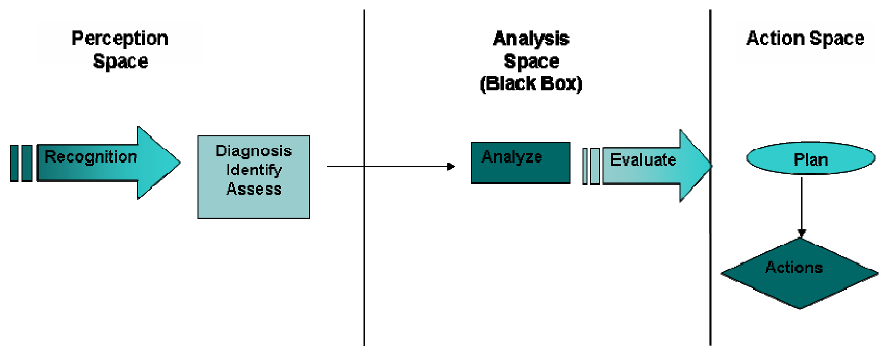


Fig. 2-5: Foundational elements of NDM models

An objective of the research is to develop a decision-aid for the ECM work domain. It has been dutifully noted that decision-makers working within the ECM field typically incorporate a "naturalistic decision-making process". Consequently, this research attempts to continue the efforts by researchers in NDM by developing a new model that follows the propositioned core elements of NDM. The following section outlines the design of the model.

2.5 Fuzzy Cognitive Maps

Fuzzy cognitive maps (FCMs) provide an efficient soft computing tool that supports adaptive behavior in complex and dynamic worlds (Siraj, Bridges, & Vaughn, 2001; Stylios & Groumpos, 2000). In recent years, FCMs have been used in domains (e.g., medical; Taber, 1991; distributed networks; Ndouse & Okuda, 1996; and health care;

Smith & Eloff, 2000). More akin to the present research, FCMs have also been used as knowledge mapping instruments[2] (Özesmi, 1999; Özesmi & Özesmi, 2003; 2004; Hobbs, Ludsin, Knight, Ryan, Biberhofer, & Ciborowski, 2002; Banini & Bearman, 1998; and Gettings, Bultman, & Fisher, 2004) and decision support systems (Siraj et. al., 2001; Geogropoulous, Malandraki, & Stylos, 2003; Xirogiannis, Stefanou, & Glykas, 2004).

FCM is a model of the cause and effect relationships that define a complex system. Unique to FCMs is their ability to incorporate attributes as qualitative states, rather than numerical characteristics. A FCM is a transfornational grammar used to model complex systems with emergent and non-linear qualities. In a sense, the FCM provides an adaptive structure that affords qualitative reasoning as assessed from the current levels or states of a complex system. Perusich indicates that the reasoning characteristics of a FCM make it a significant decision support aid for team decision makers (Perusich & McNeese, 1998).

FCM uses predefined knowledge, or constructs of the causality of concepts (represented as nodes), to define a system. It uses a graphical structure with variable concepts connected via cause/effect relationships. The strength of the causal connection is represented by a numerical quantity defined on the interval [-1, 1], with -1 representing an inverse causality and 1 meaning direct causality (Kosko, 1987). Additionally, fractional values are used for the causal connection when combinations of multiple nodes lead to an effect (e.g., a many-to-one relationship).

Conceptually, a FCM can be thought as a combination of fuzzy logic and concept mapping. Fuzzy logic is derived from fuzzy set theory dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic. It can be thought as the application side of fuzzy set theory dealing with well thought out real world expert values for a complex problem (Klir, 1997). It has been a popular

---

[2] A knowledge mapping instrument can be defined as a system that makes the way of thinking of an expert about a given system insightful (Balder, 2004).

misconception to confuse fuzzy logic with probability. The distinction between fuzzy logic and probability involves how each refers to uncertainty. Fuzzy logic is specifically designed to deal with **imprecision** of facts (fuzzy logic statements), while probability deals with **chances** of that happening *(but still considering the result to be precise)*.

Explicated further, degrees of truth (fuzzy logic statements) should not be confused with probabilities. Conceptually, fuzzy truth represents membership in vaguely defined sets, not likelihood of some event or condition. To illustrate the difference, consider the following scenario: Bob is in a house with two adjacent rooms: the kitchen and the dining room. In probabilistic terms, Bob is either in the kitchen or not in the kitchen (for this specific case, there's a 50% chance that he's in the kitchen. However, fuzzy logic statements will not refer to his *chances* of being in the kitchen; instead it will convey the *degree* of Bob being in the kitchen. This becomes especially important when Bob is in the doorway; then he may be considered "partially in the kitchen".

2.6 Summary

It is important to note a FCM decision-aid is a rather new technique. Alternative techniques that have been widely used to model decision-making are determinable and indeterminable systems (i.e., Bayesian, stochastic, or neural networks). As such, this section provides further rationale for an FCM by discussing the limitations of alternative techniques and highlighting the advantages offered by a FCM.

Deterministic systems always have one possible value, in most cases either a 0 or a 1. They are designed to adhere to behave similar to basic scientific laws; the output is always true and stable. Deterministic systems can be considered predictable systems. One main disadvantage to these types of systems deals with their inability to handle uncertainty. For example, Bob is either in the kitchen or not in the kitchen, but there isn't an indication as to the *certainty* of that statement. Probabilistic systems were designed to handle situations where uncertainty exists. Consequently, the output from these systems generally ranges between 0 and 1. They are considered unstable, indeterminate systems with the probability being an indication of certainty, and moves towards one value (0 or 1) with increasing information. These systems can come in the form of Bayesian, stochastic, or neural networks, etc. While these systems are capable to handle more situations, it has a problem managing during “in-between” situations (i.e., the case when Bob is in the doorway).

Fuzzy logic incorporates probabilistic terms by taking the view that an option either happens or does not, so each event has a fuzzy set membership of either one or zero. However, fuzziness extends the concepts of probability to encompass partial events, each of which can now occur with a strength between 0 and 1, for example we now allow for the probability of a half (eaten) apple (apple present = 0.5). Fuzziness is thus a measure of completeness and has no relation to uncertainty. Consequently a snapshot of a system (collapse of uncertainty) is no longer a vector of just binary 'truth' values, all either present or absent, but a vector of real values.

The following scenario is given to further illustrate the differences between each technique. Consider a terrorist is inside a building that only has two rooms, Room A and Room B. A solider, equipped with a decision-aid is tasked to enter the building and kill the terrorist. The decision-aid is needed because there are no lights inside the building (the solider is going in blind). The decision-aid is used to 1.) gather environmental information and 2.) assist in pinpointing the location the terrorist in Room A. However, the information may be incomplete (for example, the terrorist could have moved). The solider has to quietly infiltrate the building and has one chance to kill the terrorist.

In determinable systems, the terrorist is either in Room A (indicated by a value of 1) or not in Room A (indicated by a value of 0). If the information is incomplete, as if the terrorist has moved to the other room, then the soldier could enter the wrong room. In indeterminable systems, the decision-aid could report that from the information provided, there is an 80% likelihood that the terrorist is in Room A. However, that number could also be used if the terrorist is in the doorway. Technically, the terrorist is still in Room A, but not in a position where the solider could successfully kill the terrorist. Alternatively, by using a FCM, the decision-aid could report that 50% of the terrorist is in room A.

2.7 Research Question

The main research question is as follows:

> *How does the presence of a FCM impact team cognition and team performance in a distributed team simulation?*

The FCM identifies high priority emergencies and it serves to test the efficacy of the FCM by exposing it to issues that are prevalent in teams. In other words, the FCM becomes a valuable tool by enhancing the team overall cognition and situational awareness.

FCM's have been cited in the literature as being useful in managing existing information. This research posits that a more interesting inquiry is to investigate FCM's ability to make projections based upon existing information. This type of modality further enhances team situational awareness as it provides information regarding future elements (Endsley, 1988). One research question of interest is how team's perform during times of potential crises (or potential terrorist attacks), and how an FCM impacts team performance during these periods. In other words, does a FCM effectively assist teams in decision-making and resource allocations by identifying potential threats? This is tested by measuring the response times and resource allocations decisions of teams when a potential terrorist attack is looming. For example, the study examines if teams are more vigilant when they believe that the current state is critical with terrorist attacks being ominous.

2.8 Research Objective

The primary interest of research is to (1) develop the technological artifact that represents and an ECM work domain, (2) develop a technological entity that serves to support team decision-making, and (3) integrate the two parts into a synthetic task environment to make empirical studies involving team performance and team cognition. Specifically, the integration leads to an investigation the impact that a FCM has on team performance and team cognition in an ECM simulation. The research involves the following questions:

- *What effect does a fuzzy cognitive map have on team performance in a resource allocation-based simulation?*
- *What effect does a fuzzy cognitive map have on the formation of team mental models during a resource allocation-based simulation?*

2.9 Hypothesis

The hypotheses are as follows:

1) *H_1: The presence of a fuzzy cognitive map directly improves team performance.*
2) *H_2: The presence of a fuzzy cognitive map directly facilitates the development of strong team mental models.*
3) *H3: The presence of a fuzzy cognitive map directly speeds up the time to make a decision.*
4) *H4: The presence of a fuzzy cognitive map directly decreases the number of errors.*
5) *H5: The presence of a fuzzy cognitive map directly reduces the number of failed events.*

Except for H2, the hypotheses are analyzed using an ANOVA for each dependent variable. Team cognition is evaluated using a chi-squared measure. The assessment and evaluation of team performance is determined for every dependent measure collected in the NeoCITIES simulation environment.

Chapter 3

Methodology

3.1 Philosophy

Before describing the experimental process, it is important to note the methodological approach and the rationale for its selection. The present study adopted a "human-centered" focus. This research philosophy was chosen due to a need to better understand the social-psychological outcomes and complex decision-making processes generated by humans. The core research principle involves investigating cognitive and contextual foundations underlying knowledge acquisition and use of distributed teams in complex problem-solving domains.

3.2 System-Centric vs User-Centric

3.2.1 Problems with system-centric.

Historically, system-centric (also known as technology-centric) has dominated technology development and practice. This typically is the case where marketers, developers, and designers foist technology upon users "just because they can make it." Unfortunately, technology designed without understanding user's constraints, knowledge, and processes usually ends up creating more problems, confusions, mistakes, errors, and even catastrophic failures for users. Multiple cases exist where the user is expected to conform and learn the complexities of the design despite a device's clumsy automation or automation surprises (Woods et al., 1994). The goal is to develop technology so that it fits the user's job constraints, knowledge, and tasks. In other words, technology should cater towards the user and not vice-versa. Thus, a user-centric perspective is espoused to develop effective solutions for the users. A user-centered approach can also come in many different forms (e.g., human-computer interaction, industrial-organizational psychology, and cognitive systems engineering).

3.2.2 Rationale for User-Centric.

There are various researchers that cite designs stemming from a technology-centric approach result in many so-called "human errors" (Klein, 1989, Norman, 1988, Papanek, 1985 & Hennessey, 1977). Designs developed without user intervention have a higher chance of failing because they neglect to consider the needs, knowledge, skills, social context, organizational culture, and abilities of the intended users. Within the human factors community there has long been an acute interest in basing system design upon an understanding of the user's needs and capabilities (McNeese, Zaff, Brown, Citera, & Wellens, 1992). As Kantowitz and Sorkin (1983) suggest, human factors can be summed up by the principle, "honor thy user."

The user-centered perspective that is utilized is derived from cognitive systems engineering (Hollnagel & Woods, 1983; Rasmussen, Petjersen, & Goodstein, 1994) and is termed the Living Lab Approach. The focus is upon social-technical systems designs that emphasize the interactive impacts of cognition and context. The Living Lab approach (Living Lab) is used because it is a suitable approach to examine decision-making in a distributed setting.

3.3 Living Lab Approach

There exists a wide variety of CSE frameworks (Dowell & Long, 1998; Eggleston 2002; McNeese 2002; Vicente 1999. The Living Lab is a user-centered perspective that focuses on the interactive impacts of cognition, which is situated and distributed, and context, which involves the worker interacting with the work environment.

There are four nodes of the Living Lab framework used to assess workers in context: (1) *knowledge elicitation*, (2) *scaled worlds*, (3) *reconfigurable prototypes*, and (4) *ethnographic study*. The evolution of theories and models is represented across the upper portion of the framework while progress in *use* and *practice* is embodied by the lower. Notably, practice and use are tightly interwoven by the four nodes of the Living Lab – the output from one becomes the input for another. The process is cyclic, but flexible enough to move forwards, backwards, and crosswise between nodes. (See figure).

Figure 3-1

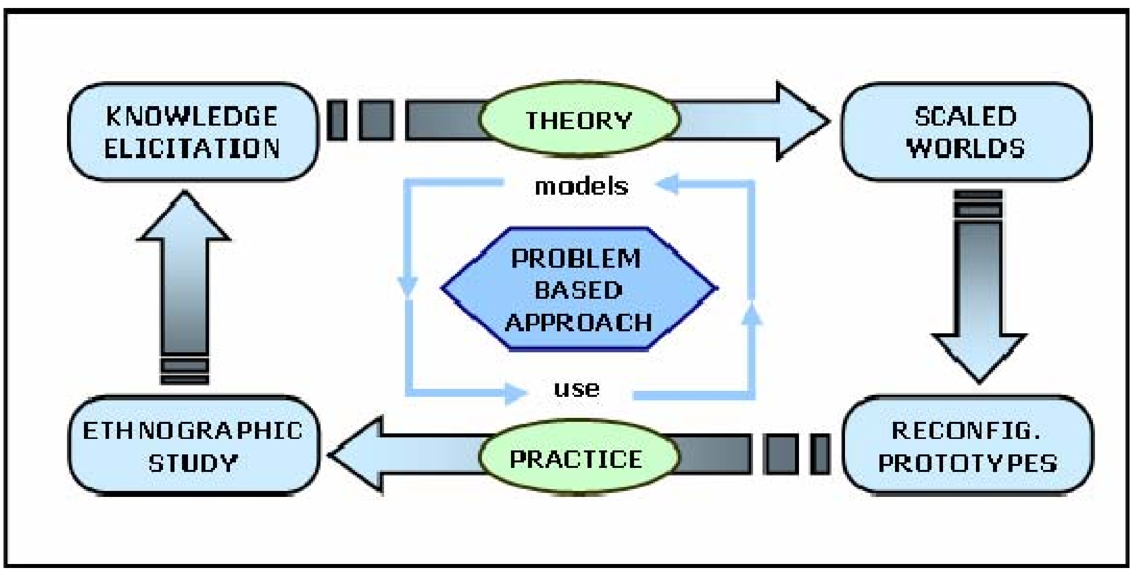


Fig. 3-1: The Living Laboratory Approach (adapted from McNeese, 1996)

The Living Lab is a cognitive systems engineering approach that addresses distributed cognition settings wherein information is distributed broadly across time, place, and people. The framework presupposes that these settings require the effective design of (1) supporting technologies (e.g., information and communication technologies), (2) knowledge management strategies, and (3) safe work practices (for more on the Living Lab, see McNeese, Connors, Jefferson, Bains, Jones, Terrell, Brewer, Craven, Van Vactor, Rotthoff, & Hall, 2004).

The Living Lab incorporates a problem-based approach (McNeese, Perusich & Rentsch, 2000; McNeese, 1996) that necessitates the researcher to examine the overall problem to find which node to start from. For this particular effort, the research mainly fits within the *scaled worlds* node, but it also stretches into the *knowledge elicitation* and *reconfigurable prototypes nodes* and incorporates the *theory and models* processes. In development is a scaled-world simulation, called NeoCities, that uses theory that was derived from information processing theories (Wellens & Ergener, 1988) and findings from previous research that utilized knowledge elicitation to examine the role of cognitive processes in facilitating knowledge acquisition and transfer (McNeese, 2000). Additionally, the output of the simulation involves eliciting knowledge to develop a team mental model. The team mental model could be used to inform the design of a socio-technical system that uses intelligent and adaptive features to support group decision-making. Subsequently, within the *scaled worlds* node, elicitation techniques is needed to represent and analyze a team mental model.

3.4 Participants

The original pool of participants derived from 2 undergraduate Business classes. The classes were chosen because 1.) the researcher has a professional relationship with the instructor of the classes and he [the instructor] thought it would be a worthwhile opportunity for his students to participate in a meaningful experiment and 2.) the classes held over 400 students, which were significantly more than the study needed for participants. As such, participation was solicited from these classes by giving a short talk during a lecture period and by posting advertisement fliers in their online angel class website. However, despite the opportunity of being paid $10 for their involvement, the majority of the students decided not to participate. Consequently, the researcher opened participation to all undergraduate students by posting flyers, emailing professors, and sending advertisements to student listserves. Even with this effort, the number of people signing up was not sufficient to hold the desired number of full sessions (i.e., 6 individuals running as 3 three teams). Since running a session consisting of 6 students became very difficult, an ad-hoc decision was quickly made to run sessions consisting of 2 participants running as one team, or 4 participants running as 2 teams. As a result, 34 sessions consisting of approximately 800 students participated in the experiment with the majority of sessions being 1 or 2 teams instead of the full 3 team operation[3].

Once students arrived at the lab, they were given an informed consent document that required a signature before continuing. Then, they were randomly assigned to a team (Police, Hazmat, or Fire) and a role (Information Manager or Resource Manager). After assignments were made, the students sat at their respective work stations where training commenced.

3.5 Data Requirements

The present study focuses on team performance and team cognition at the intra-team level. The unit of analysis is the dyad for each team. As such, a full session yields 3 data points (one from each team). Data is collected and compared at the team level (dyad). Since all teams are exposed to exactly the same manipulations, the team names are not recorded. Additionally, valid comparisons can be made because all teams undergo the similar experiences.

---

[3] The impact of running sessions consisting of 1 or 2 teams (compared to a full session) is discussed in the Discussions chapter.

3.6 Task

The NeoCities task is a simulation that was chiefly designed and developed by the researcher. The simulation took 11 months to program with approximately 10000 instructions (or actual lines of code). The design was abstracted from the popular video game SimCity. In fact, NeoCities could be described as an online team-based version of SimCity. The task requires at least 3 machines, 1 server and 2 clients (1 team setup); a full session requires 7 machines (1 server and 6 clients). The clients are the I-Mgrs and R-Mgrs for a specific team. NeoCities can run on any Windows XP computer that is equipped with Java 1.4.2. There are over 60 core java files and approximately 10 utility files. The development process involved the following tasks:

- Design and development of interface.
- Design of scenario.
- Integration of scenario module.
- Formulation and implementation of scoring module.
- Design and implementation of communication architecture (i.e., network).
- Implementation of a metric collection module.
- Coding of batch files to run NeoCities.

The present study involved participants completing a terrorist scenario within the NeoCITIES synthetic task environment. The episode was scripted specifically for the present study with a terrorist storyline comprised of various events, some terrorist-related, others routine that required participants to acquire information, share knowledge, and make resource allocation decisions.

The NeoCITIES task provides a variety of taskwork- and teamwork-related content over the course of the simulation. These elements have been identified by Cooke and colleagues (2000) as essential components in the formation of team mental models. The simulation is an ideal framework because it represents an excellent task for measuring team performance, specifically regarding decisions and quality of responses, due to its resource allocation structure. Other studies utilizing synthetic task environments have validated use of such tasks and led to the successful measurement of team performance (see Cooke, Kiekel, & Helm, 2001). In addition, much like the DDD simulation environment used by McNeese and colleagues (2000) to reveal team member schema similarity effects, the NeoCITIES scaled world affords high team member interdependence, distributed cognition, and an increased length of interaction time. Since a major effort of this research was to develop a simulated task, the following section is provided to give a comprehensive description, covering the theoretical and technical aspects of the simulation.

3.7 Artifact

NeoCITIES is an ECM simulation that was primarily designed from knowledge elicitation and observational studies involving dispatchers, analysts, operators, and decision-makers. Additionally, the NeoCITIES task is viewed as a both an update and extension of the original CITIES task (Wellens & Ergener, 1988). The primary objective of NeoCITIES (Jones, Jefferson, Connors, McNeese, & Perusich, 2005; McNeese, Bains, Brewer, Brown, Connors, Jefferson, Jones, & Terrell, 2005; Jones, McNeese, Connors, Jefferson, & Hall, 2004) is to test team collaborative decision-making processes, knowledge acquisition and knowledge management within a command, control and computer-mediated communication (C4) environment. As an adaptable problem interface, NeoCITIES is designed to allow researchers to closely examine team behaviors, identify patterns of response to time-stressed situations, and monitor the performance outcomes of semi-autonomous, spatially distributed, decision-making teams. MacMillian and associates have commented on DDD's flexibility making it unique amongst other war-games (2004). Similarly, NeoCITIES is a highly dynamic and robust task because modules (i.e., type of communication bridge, task-type, team structure, decision-aid, map style, etc) could be "dialed-in" at run-time.

Users participating in NeoCITIES are presented with a wide-variety of emerging, dynamic resource allocation events designed to be addressed by either single or multiple teams. To solve these events, teams are required to meet the needs of their given constituents, while working around various problem space constraints related to the underlying emergency crisis management scenario.

At its core, NeoCITIES is a team resource allocation problem designed to mimic the emergent situations that comprise real-life emergencies and measure decision-related outputs in a virtual environment. The simulation emulates the complex functions involved in the resource management of a city's emergency services. Crisis management is conducted through the joint interaction of three distinct teams: a *Police team*, a *Fire/EMS team*, and a *Hazardous Materials team*.

Each crisis-management team in NeoCITIES is comprised of dyad with each member assuming a specific role. The *information manager* is in charge of handling all incoming events. The information manager for each team is the first to receive notification about the events; additionally, they are responsible for processing information about incoming events, which includes notifying their teammate of those events requiring action, and oftentimes, they are required to communicate across all 3 teams. The *resource manager* is tasked with managing resources throughout the simulation. Their responsibilities include: allocation of resources to events, monitoring progress of on-site resources, and communicating on-site information back to their information manager. Each resource manager is given a finite amount of 3 specialized resource types. Each resource type has unique capabilities that can handle a specific set of emergencies. Since the resource manager is expected to have knowledge about how to handle emergencies, they are given information that pertains to their resources (explains the specialization of each resource), protocols (details quantity of resources for a given emergency), and other important facts that they may find useful (i.e., the resource manager for the police team is given crime statistics that illustrates the percentages of crimes committed within the area over the past year).

The communication architecture in NeoCITIES provides chatting functionality for each team member. Since the teams are assumed to operate in distributed mode, face-to-face talk is not an option; as such communication can only be performed within the NeoCities interface. The communication architecture is implemented as instant chatting modules for each team member. The features and functionality of these modules are similar to those found in mainstream instant chat technologies (i.e, yahoo, msn, or aol). The communication channel is designed to allow the information managers for each team to freely discuss and share information. As such, each information manager can chat to each other and chat to their respected resource manager (police information manager is allowed to chat to the police resource manager, but is not permitted to chat with the fire resource manager). Since the resource manager primary responsibilities involve managing resources, their communication is limited so that they can only chat with their information manager.

Information in NeoCITIES can come in top-to-bottom or bottom-to-top. The information manager is always the first to receive information pertaining to a new event.
As such, an example of information being dispensed in top-to-bottom fashion involves the information manager receiving notification about an event, which then can be sent to the resource manager. The resource manager can then send resources out in the field. This emulates someone calling 9-1-1 and the dispatcher notifying the proper emergency response units, which would then arrive to the site of the emergency (given by the caller). Similarly, the resource manger can receive feedback pertaining to a handled event, and in some cases, could lead to information pertaining to an event that will occur later in time (information being distributed following the bottom-to-top approach). In this case, information will need to be propagated upwards to the information manager, especially if the future event involves another team. This emulates the discovery of information (i.e., tips or rumors from reliable sources) that could be used to prevent or thwart a future emergency.

As a way of modeling teamwork, decision-making and communication in circumstances of crisis management, the three teams must address events that may involve potential terrorist activities. These events are conceptually similar to those undertaken during routine and non-routine law enforcement. In the case of non-routine events, their structure is analogous to the structure of events that emerged out of analysis of the September 11 attacks in that they are comprised of vague, unexpected, complex, highly stressful incidents, and typically occur under severe time-pressure. The teams are tasked with responding to events by applying different numbers of certain resources, combine different types of related resources, or perhaps even share resources across teams for these events to be resolved. Events range from isolated and mundane occurrences, to larger, more widely inclusive events that have the potential to escalate as a function of inadequate resource allocation.

In addition to assigning resources to events, participants are also tasked with scanning and identifying the underlying scenarios that manifest themselves in the simulation as terrorist-related events. These events escalate rapidly in intensity and require resources from each of the three teams to be brought under control. Once participants have identified these underlying scenarios, they can use this information to inform their upcoming decision-making and resource allocation, thus compensating for the presence of these background factors while anticipating and allotting resources to address these likely events.

NeoCITIES borrows the same event growth formula that was used in CITIES (Wellens and Ergener, 1988) to drive the escalation of event magnitude. Each event begins, or "fires", with an initial magnitude that indicates the severity of the event. The magnitude increases in value as the simulation updates the event growth formula as a function of time and participant response (i.e., the number of correct resources allocated to the event). Each event remains active until reaching one of two possible end-states. The success, or positive, end-state occurs when the magnitude of the event growth formula is non-positive. In this case, the participants responded to the event with the correct resource type in sufficient quantity. If an event magnitude were to grow past a set endpoint (e.g., greater than one plus the initial magnitude), the event would be said to have reached the failure, or negative, state. Once either end-state is attained, the event is deactivated and participants are no longer able to allocate resources to it.

The CIITES simulation involved only single team events and the players were not responsible to choose resource types. The players in CITIES were only charged with deciding the appropriate quantity of resources for allocation and the system automatically sent the appropriate resources to the event. As such, NeoCITIES have expanded the formula to include events that require multiple teams and decisions involving resource types. The revised formula includes a team performance variable that becomes important in events requiring multiple teams. The team performance variable is designed to reward a team when the other team is operating successfully towards an event and penalize a team for another team's failure. For example, in the case of an event that requires a police and fire team, if the police team responds successfully, but the fire team has not, then the fire team score will incorporate the police team's success and the police team's score will factor into the fire's team inability to make the correct decisions. This variable links teams as one super ordinate team, so when one team fails, the other team will partially fail and when one team succeeds, the other team partially succeeds. The other revision involves detecting the decisions made by the resource manager. The magnitude of the event can only decrease by the resource manager allocating the correct resource type to a particular event.

NeoCITIES events may either be single team constructs, or they can be more complex, requiring two (or even all three) teams to pool the necessary resources to resolve the emerging problem. Initially, only a few events are triggered, as a way of easing participants into the simulation. As time increases, the number of events presented increase, while the informational content for many events will become more complex. Participants will need to process these events quickly to determine their importance, or the events in question will start to escalate and run out of control.

3.8 NeoCITIES Parameters

The present study utilized a terrorist scenario that was comprised of terrorist-related events, routine events, and informational events. All events were single team in nature (although this was not divulged to the participants). Routine events only required 1 resource type, while an event deemed as "terrorist" required 2 resources of the same type. Informational events required no resource allocation action; instead it provided the teams with additional SA of current environment. (The informational events can be regarded as an update on the current state of the world). A total of 52 events were developed for the episode with each team having 16 events (12 routine events, 4 terrorist events, and 4 shared informational events).

In developing a scenario, it is important to provide a realistic and interesting story that motivates player's participation. Additionally, it is essential that the scenario is not overly complicated because it would confuse the players; nor should it be simplistically easy because that would result in a lack of interest. Lastly, the story should be engaging, but epigrammatic to decrease negative consequences attributed to attention loss or boredom (see Hitchcock, Dember, Warm, Moroney, & See, 1999). As such, events were fired every 30 seconds and a score was computed every 6 seconds. The simulation exited after 29 minutes. The teams were given 8 minutes of "null" time after the firing of the final event to allow additional time for the teams to correct any erroneous decisions.

3.9 Measures

The dependent variables of interest for the present study are as follows:

- Team performance – a quantitative score indicating the overall progress of the team during the simulation. Since team performance is assessed at the intra-team level, it includes only events that require resources from the team of interest. For example, team performance for the police team will only include events that require resources from the police team. Team performance is a comprehensive measure indicating the team's timeliness and accuracy in responding to events. Each event has a magnitude value that negatively affects the score. At every update, the magnitude will either increase (event has not been attended to correctly) or decrease until it is zero. At the beginning of the simulation, team performance equals 100% and at every update, it is decremented by the magnitude of all active events. The magnitude of the event is computed via the following variables: initial magnitude and time duration. The initial magnitude of the event (predetermined by the scenario developer) determines the initial severity of the event; it indicates how much the score will be decremented at each update. The time duration is the length of time it takes for the team(s) to correctly respond to the event; it indicates the growth value of the magnitude. The ideal behavior for teams is to respond to events expeditiously with correct decisions, especially for events that have high initial magnitudes.

- Time to correct decision (average) – this variable is used to convey how much time it takes (on average) for the teams to make the correct decision regarding an event. Teams can either ignore events (that do not pertain to them), or respond to appropriate events with the correct quantity of a particular resource type that is designated in handling the emergency.
- Decision-making errors – this occurs when teams are unable to respond correctly to an event. 3 types of decision-making errors can occur for every event: 1.) either the team ignores an event they should respond to 2.) the team responds to an event that they should ignore and 3.) the team responds to event with incorrect resources.
- Number of failed events – the number of events that reaches a "fail status". In the case of events that require multiple teams, the team(s) that fail to respond correctly will count towards this variable.

In addition to quantitative variables listed, a qualitative measure indicating team cognition is used. This variable is denoted as "team mental model development" and it conveys the strength of the team mental model. A team member similarity schema assesses the development of the team mental model. Additionally, the actions of the teams (i.e., their resource allocation decisions, information that is shared, etc) will be analyzed in accordance to their cognition.

The Teamwork Similarity Questionnaires consists of 15 items in length was used to collect measures of team member schema agreement (or congruence) and team member schema accuracy. These items are rated on a 7-point Likert scale from 1 (extremely unimportant) to 7 (extremely important). Each individual completed the measure with respect to his or her own views on teamwork. For this portion of the measure, participants receive instructions similar to the following (Rentsch et al., 1998):

Think about what teamwork means to you. Think about teamwork as it may occur on any team. Thinking about teamwork in this way, please read each of the following statements. Think about how important these behaviors or events would be in telling you about the nature or meaning of teamwork. Ask yourself "Does this event tell me anything about the meaning of teamwork?" When considering the importance, keep in mind your view or meaning of teamwork. (p. 21)

This measure is compared to a teammate's in order to obtain team member schema agreement. To assess team member schema accuracy, participants answer the same questionnaire, but with respect to how they expected their partner to respond. Instructions for this section is similar to the following (Rentsch et al., 1998):

Now please complete the same questionnaire, but this time, respond how you think that your partner would respond to the questionnaire. Please rate how you think your partner would say each of the following items is to his or her meaning of teamwork. When considering the importance, keep in mind your partner's view of teamwork. (p. 21)

Team mental model development is assessed via TMSS surveys. The TMSS surveys yield the strength of the team mental model for each team. The values resulting from the TMSS are: *weak, weak-to-medium, medium, medium-to-strong,* and *strong*. The degree of agreement and accuracy among dyad members is determined by the following:

Schema congruence and schema accuracy among the dyad members were determined using the absolute difference formula (Cronbach & Gleser, 1953, as cited in Pape, 1998). The absolute difference formula is represented as $| D | = \Sigma | X_i - Y_i |$ where X_i is the response on item i for one member, and Y_i is the response on item i for the second member.

Each team member has a sub-TMSS value that is combined with their respected teammate to obtain a major-TMSS value that is used to reflect the team mental model development of the respected team. The following graphic illustrates the process:

Figure 3-2

Step 1:
Obtain sub-TMSS value for each team member

Accuracy	Agreement	Sub-TMSS
Low	Low	Low
High	Low	Medium
Low	High	Medium
High	High	High

Step 2:
Combine sub-TMSS value for at the team level

Sub-TMSS (IMGR)	Sub-TMSS (RMGR)	TMM
Low	Low	Weak
Medium	Low	Weak-to-Medium
High	Low	Medium
Low	Medium	Weak-to-Medium
Medium	Medium	Medium
High	Medium	Medium-to-Strong
Low	High	Medium
Medium	High	Medium-to-Strong
High	High	Strong

Figure: Obtaining team mental model development

Fig. 3-2: Obtaining team mental model

3.10 Data Collection

All dependent variables were logged via the simulation. The simulation is designed to monitor and store quantitative measures in a log file. Post-processing steps were performed to organize the raw data in a useful and meaningful way for data analysis. To collect TMSS-related data, participants were given a survey at the conclusion of the simulation that required to be completed prior to completing the experimental session. Additional post-processing routines were developed to compute the TMM score.

3.11 Design

The experimental design consists of a 2x6 study. The independent variable is the presence (or absence) of the FCM. (The sessions that do not have the FCM can be thought of as the control group.) The dependent variables of interest are team performance, team mental model development, number, average time to correct decisions, decision-making errors, and the number of events that reach a fail status. The quantitative measures are computed and recorded during the simulation, while the team mental model is assessed via a team member schema similarity survey following the simulation. The following graphic illustrates the experimental design:

3.12 Procedure

The experimental session commences by handing out an IRB-mandated inform consent document, which was required to read and sign prior to continuing. Next, participants were randomly assigned to a role (i.e., information manager or resource manager) and held the same role throughout the entire session. Dyads were randomly assigned to one of the three NeoCITIES response team types (Police, Fire/EMS, or Hazards). The participants were then given a briefing session that introduced and described the experiment and the simulation. The briefing session was a Powerpoint show that served as a basic training guide to familiarize the participants with information regarding:

1. The main objectives of NeoCITIES;
2. The primary functions of their role as either an I-Mgr or R-Mgr;
3. The function of role-specific tabs (e.g., the Information tab for I-MGRs)
4. The primary functions of their partner's role (e.g., the functions of an R-Mgr for an I-Mgr);
5. A detailed diagram of the NeoCITIES interface;
6. A basic operating procedure for event handling in NeoCITIES;
7. The event icon library used by I-Mgrs to assign a preliminary assessment of an event; and

8. Descriptions of the three resource types available to each team, including detailed information about each resource's capabilities once on-scene[4].

The participants were introduced to a short training session[5], after reading the guide. The training session, similar to a walkthrough in a video game, served to acclimate the participants orientation of the interface and how to use the features/components provided by the interface, and how to utilize the intelligent decision aid functionality (if it is present). In other words, the walkthrough activity allowed the users to put to action the material they read during the briefing. The walkthrough activity comprised of 9 events with 3 events per team.

The participants partook in a 30 minute main activity that involved the terrorist scenario. The teams were told the following prior to beginning the main activity:

> "This episode involves a terrorist attack. For any events that you believe to be terrorist-related requires 2 resources. As before, any routine event only requires 1 resource, but sending only 1 resource to a terrorist-related event will not properly handle the event."

If the condition included the presence of the FCM, the teams were provided with the following information:

> "The R-Mgrs have a computer-aided assistant. The assistant is designed to assess the terrorist level of any event the R-Mgr receives. If it believes the event involves terrorism, it will highlight that event in black in the Event Tracker. The aid is not perfect because a part of its reasoning is dependent upon your prior

---

[4] Hard-copies of the descriptions of the resource types were given to each R-Mgr for use throughout the remainder of the experimental session.

[5] Teams were instructed that their only form of communication was via the chat panels in the interface. Verbal chat was prohibited and it was discouraged by having each student wear headphones during the simulation.

> actions. However, if your team operates perfectly, the aid would accurately highlight all terrorist events."

During this session, participants handled numerous events over the course of the simulation. This allows for a measurement of performance in the form of overall simulation time. At the completion of the session, the participants answered the team member schema similarity questionnaires and their participation was noted. Following the questionnaire, the participants were debriefed and released, concluding the experimental session. The following illustrates the steps of the experimental session and the time of each activity:

1. Briefing & Walkthrough -> 10 mins
2. Main task (either condition 1 or condition 2) -> 30 mins
3. TMSS Survey -> 10 mins
4. Debriefing -> 5 min. (Conclusion of session).

Total Time: Approximately 55 minutes.

3.13 Scenario Design

The main activity episode involves a terrorist attack on the city and it is designed to swarm the teams with routine events to keep the number of available resources low, launch terrorist-related events at strategic times, and fire a severe terrorist attack on the city towards the end of the scenario when resources are unavailable. Thus, the goal for the teams for this episode is to process the information from terrorist-related events correctly to predict the terrorist attack (goal for the information manager) and manage resources shrewdly so that there are sufficient resources available when the terrorist attack occurs (goal for the resource manager). This strategy may include ignoring routine events and waiting for events with higher severity. Once the events with higher severity are handled, resources could then be allocated to previously-ignored routine events.

The episode is designed to last 840 seconds with 18 events per team. The episode is comprised of 7 phases lasting 120 seconds each. There are 4 unique phases that contain events that are characteristically different from events fired in the other phases. For example, phase A contains single team events requiring 1 resource of a particular type with the last 3 events involving terrorist events (2 resources). Phase B is similar to Phase A, except it does not contain any terrorist events. Phase C contains only 3 events (one for each team) and they are all terrorist-related. Lastly, Phase D is considered a "null" phase as it contains no events. The following table highlights each phase:

The first 120 second phase involves a firing a single team event for each resource (only 1 resource is needed) for every team (phase A). The events are fired in an event firing block that lasts for 40 seconds with each block firing one event per team (the resource type required changes per block). The last block includes events that are terrorist related; it is at this time when teams are initially exposed to information pertaining to a potential terrorist attack. This information is supplied automatically to the FCM for the next phase.

The second and third phases are a repeat of phase A. It is used to keep the amount of available resources low for the subsequent phase. It is during this phase that the FCM could start highlighting events (since information obtained from the first phase included a possible terrorist attack, which activated a flag).

The episode transitions to a new phase for the 4th and 5th phases. Phase B is used as a buffer or in-between for Phase A and Phase C. Since Phase C only contains terrorist events, which are the most important events in the episode (because they are climatic events), it is essential that the teams have sufficient resources to handle these events. 2 phase Bs are used to allot sufficient time for the resources that were sent in Phase A to become available. However, if the players incorrectly classify any events in Phase B as terrorist, it could jeopardize their ability to handle the events presented in Phase C (this was true in several of the experimental sessions).

Phase C is the 6th and most important phase of the episode. It contains only terrorist events (1 per team) with each event leading to disastrous results if failed, or successful capture of the terrorist organization and the saving of lives if completed correctly. The events in Phase C are fundamentally different from the terrorist events presented in Phase A in that the outcome for Phase C events have severe ramifications, while the outcome for Phase A events presents the user with additional information that can aid later in the scenario.

Lastly, there is a system-pause time of 5 secs after each event is fired. This system-imposed delay is used to offset some of the time it takes for the players to read the text pertaining to an event. The following is a graphic summarizing the timeline:

Figure 3-3

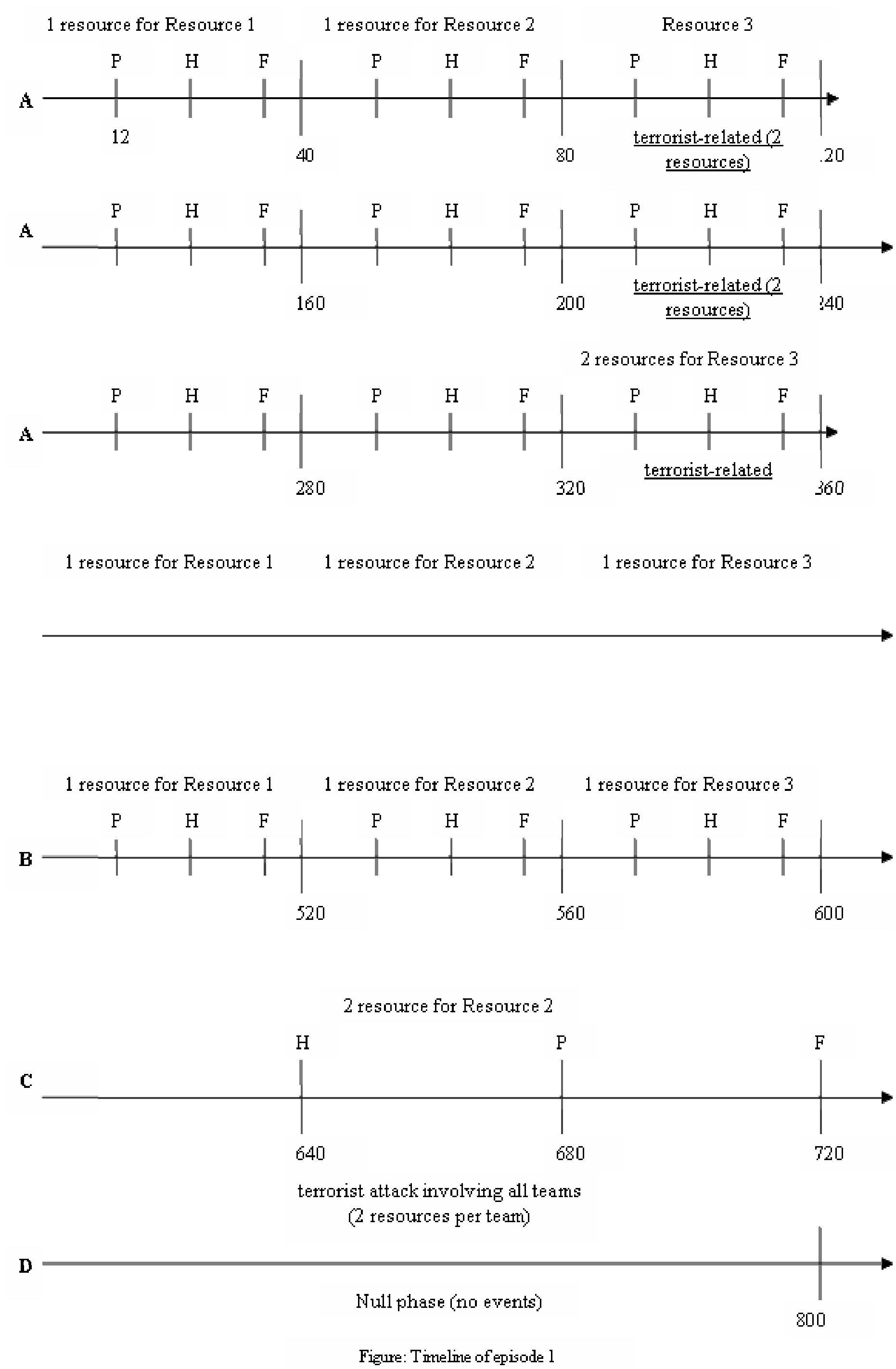


Fig. 3-3: Timeline of episode

Chapter 4

NeoCITIES

This section is intended to provide a detailed overview of NeoCITIES. The objective of this chapter is to (1) provide a comprehensive overview of NeoCITIES and (2) describe the technical aspect of NeoCITIES. It is important to note that the author programmed NeoCITIES and has intimate knowledge of the inner workings of the simulation and a complete understanding of the many intricacies of NeoCITIES. The author is the sole developer, or creator, of the NeoCITIES task. He developed NeoCITIES from the scratch and is responsible for its existence. The primary contribution of the research is NeoCITIES and the remainder of the chapter discusses the many components of the simulation. Unless noted, the variables used all reflect the instantiation of NeoCITIES that was used for the present study.

4.1 Team Structure[6]

NeoCITIES involves 3 interdependent teams that are equivalents of local police, fire, and hazardous materials units. Corresponding to the real world, each team is tasked with a limited amount of responsibilities. A tutorial session is provided that explains the governance of the team.

---

[6] It is noteworthy to mention that while the present study used examinations at the intra-team level, similar measures could have been employed to perform investigations at the inter-team level.

Collectively, the teams function as a super ordinate team or a crisis-response unit (CRU). The CRU is given the responsibility of addressing any and all emergencies. In order to do this correctly, team members must know and understand the capabilities of the resources available to them. Additionally, team members are expected to communicate and collaborate with other members of CRU in order to maintain high situational awareness and to have a comprehensive understanding of the current state of affairs in the virtual world. The following graphic illustrates the many components of collaboration of teams in NeoCITIES.

Figure 4-1

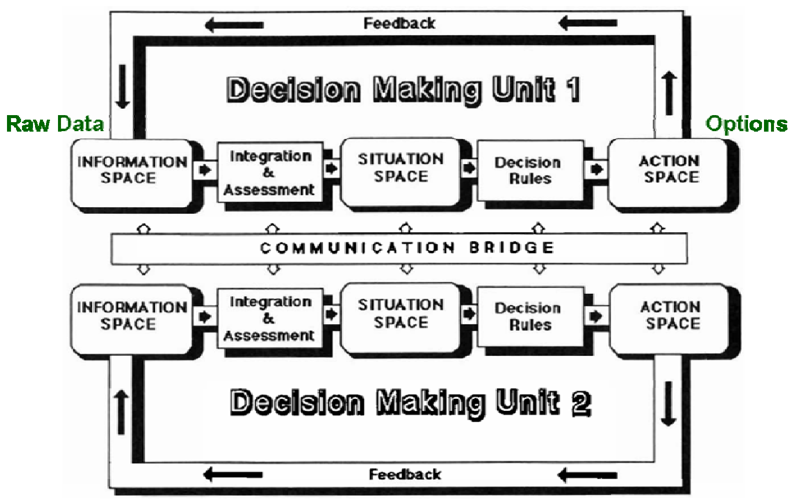


Fig. 4-1: Collaborative aspect of NeoCITIES

4.2 Resources

Each team has 3 types of specialized resources that have a finite amount throughout the simulation. The constraint on resources makes NeoCITIES a resource allocation task. Specifically, the present study allocated 3 resources per type for each team. All teams are given an investigative unit capable of conducting investigation which include, gathering information, interrogating witnesses, and collecting and analyzing evidence. The police team is given a set of 2 units: a Squad Car Resource Unit and a SWAT unit. These units are capable of dealing with events ranging from public disturbances, automobile accidents, hostage situations, domestic crimes, and other police-related emergencies. Squad cars are required for “basic” events, whereas SWAT units are involved with crowd-control, hostage situations, shoot-outs, and other serious crimes. The Fire team serves a dual purpose; (1) addressing fires and (2) providing emergency medical care by deploying EMS units. The Fire team has at its disposal a Fire Truck Resource Unit, which are required to handle emergencies that involve fires and an EMS Resource Unit that is capable to administer medical care for injured victims or preventative care. Lastly, the Hazmat team has two specialized resources that address events relating to chemicals (such as spills, bio-warfare, etc.), explosives, and other weapons. Namely, the Hazmat team has a Bomb Squad Resource Unit responsible for disarming, disposing of, detonating, and transporting explosives and a Chemical Truck Resource Unit responsible for clean-up, disposal, and transportation of all hazardous materials of a chemical, biological, or radiological nature. The following graphic is provided to summarize the resources made available to each team.

Figure 4-2

Police	**Fire**	**Hazmat**
Investigative Units	Investigative Units	Investigative Units
Squad Cars	Fire Trucks	Chemical Trucks
SWAT Units	EMS Units	Bomb Trucks

Fig. 4-2: Resource types available in NeoCITIES

Prior to running NeoCIITES, participants are asked to read and understand “resource specification sheets” that define the capabilities of each unit. The resource sheet is available in both softcopy and hardcopy. As a softcopy, the resource sheet is one component of the training module that appears as a slide in a PowerPoint file. An example is shown in the following graphic. (Please note that the appendix contains screenshots for each resource type). A hardcopy version is given to the resource managers during the experimental task. Resource specification sheets provide essential information necessary for the teams to match resource types and requirements for a particular emergency. The descriptions are written in a generic form to make resources applicable to a wide series of events. Particular care was taken with the wording of the descriptions to avoid conflicts of interpretation. No testing on the reliability of interpretations has been performed, but experiences during prototype testing, beta testing, and trial runs have shown that the descriptions can be easily understood by subjects. Additionally, events are written such that resources available to the teams can, in fact, address the event. In other words, an event is checked against the list of resource specification sheet to verify that subjects should know how to appropriately and correctly handle the event.

Figure 4-3

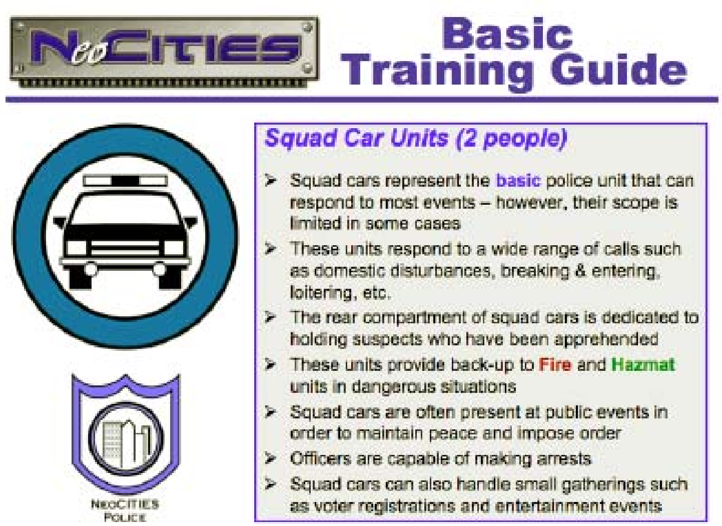


Fig. 4-3: Resource specification for Police team

4.3 Roles

Each team is comprised of: an information manager and a resource manager. Each has specific skills and capabilities. Information managers receive event information for the system and are responsible for dispensing and forwarding relevant information to other team members. They are also required to coordinate efforts with other teams and analyze information obtained throughout the simulation in search of emergent trends. The resource manager, as the name implies, is in charge of the allocation of team resources in response to events in the simulation. Resource managers are expected to review the details of an emergency and determine which resource(s) (if any) are needed to correctly address the particular event. Handling an event correctly may lead to the discovery of new information that pertains to future events. For example, interviewing a witness during event e(t) may lead to a description of the perpetuators (4 males with "x" characteristics). An event that occurs later, e(t+i) may involve 4 males with "x" characteristics. The additional information may help teams quickly determine the nature of the event and which resource types are needed.

Information managers are generally concerned with a broader perspective; they have knowledge that pertains to the entire CRU; while the resource managers have a focused concentration that primarily involves their specific resources and allocation decisions. In order to allow team members to function as a particular role, NeoCITIES includes a GUI for each role. Resource managers have an interface component that shows the number of available resources and the number of total resources. Information managers have an additional chat window to talk with information managers from other teams. The responsibilities for each role are summarized in the following table.

Table 4-1

Table 4-1: Roles in NeoCITIES

Role	Responsibilities
Information Manager	Receive event information
	Interpret information
	Pass assessment to RM
	Consult references
	Communicate with other teams
Resource Manger	Receive event information from IM
	Assess information
	Assign resources to events
	Respond to simulation feedback

4.4 Communication

Lastly, communication in NeoCITIES is setup to simulate teams operating in a distributed manner. Even though team members sat next to each other, they are only allowed to communicate through the interface. A communication interface component is provided in NeoCITIES that replicates popular instant messenger chat (Yahoo or AOL). Furthermore, communication in NeoCITIES is designed to create an additional layer of interdependence. Specifically, communication channels were constrained such that:

- The information managers could chat with all other information managers, but could only chat with the resource manager of their own team.
- The resource manager of a particular team can only communicate with the information manager of the same team

These constraints create a hierarchical communication network. First, the resource managers communicate in a "stove-pipe" structure; information they need must come from their team's information manager and if they have any information that another team needs must be directed to their team's information manager. Secondly, for the CTU to function correctly information must follow a particular channel; in certain cases, information can be sent horizontally amongst information managers (inter-team) or vertically between an information manager and a resource manager for a particular a team (intra-team). A benefit of this hierarchical communication network allows added control for experimenters in the distribution of information. The following graphic is provided to illustrate the hierarchical communication network.

Figure 4-4

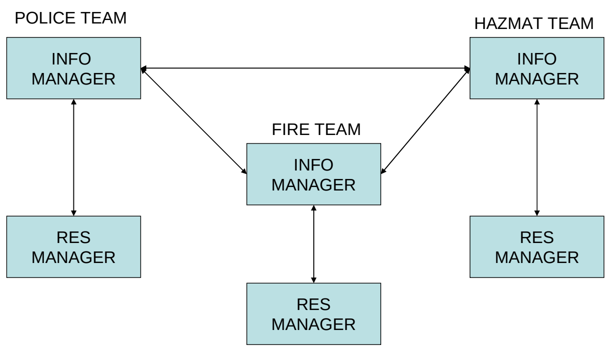


Fig. 4-4: Communication channels in NeoCITIES

4.5 System Architecture

NeoCITIES uses a server/client architecture. The server runs the game engine, which fires events (and advisory updates) and computes the score for each team. The scenario file resides on the server and at given times, specified when the server starts-up, sends an event (or advisory notice) to the teams. At a specified simulation time, the server computes the score, which is determined by assessing the penalty for every active event. On the client side is the GUI. The clients who run the GUI for the information manager receive the events (and advisory notes) from the server. The clients who run the GUI for the resource manager sends the resource allocation decisions to the server and it receives feedback on those decisions (via the server). The following graphics are provided to illustrate (1) the propagation of events in NeoCITIES and (2) the complete architecture that includes the components on each side.

Figure 4-5

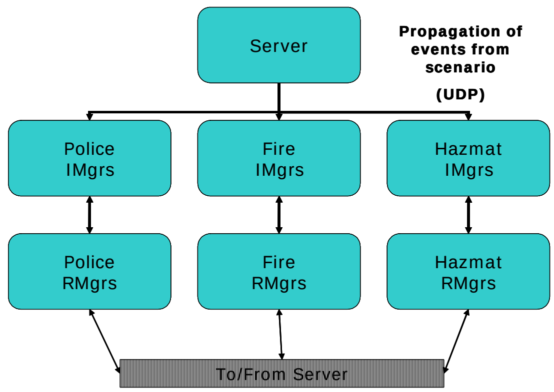


Fig. 4-5: Event propagation in NeoCITIES

Figure 4-6

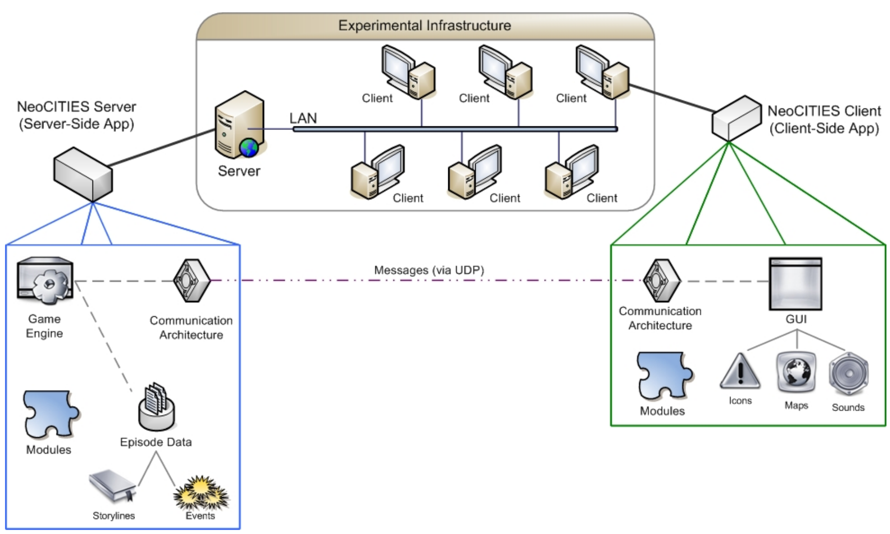


Fig. 4-6: Server/Client architecture

The architecture uses User Datagram Protocol (UDP). UDP was chosen due to its simplicity in implementation. Other protocols, such as TCP/IP requires extensive overhead, such as error checking and connection checks; however, UDP does not provide the reliability and ordering guarantees that TCP does. Datagrams sent by UDP may arrive out of order or go missing without notice. Without the overhead of checking if every packet actually arrived, UDP is faster and more efficient for many lightweight or time-sensitive purposes, such as instant messenger and online radio.

4.6 Score Model

Team performance is assessed by a computational scoring model. Specifically, the scoring model is used to measure the performance of teams based on the allocation of resources. This model is predicated upon the characteristics of emergencies that are found in the crisis-management work domain. Specifically, emergencies generally start with an initial severity (or magnitude) and if left unattended, escalate as time progresses. For example, a fire in a trash can burn an entire building if not addressed. Hence, the scoring model incorporates these characteristics via an *event growth formula*, as described by Wellens and Ergener (1988). It is a mechanism that determines event growth and decline." This measure of growth is labeled:

$$Mt = a(Mt\text{-}1) + b(Mt\text{-}1)2 - c(R)$$

where *Mt* represents the magnitude of an event at time *t*

M(t-1) is the magnitude at the previous time step,

*R is t*he number of resources, and

a = 0.995, *b* = 0.0075, and *c* = 0.04996

The magnitude of an event is calculated at every iteration that is specified at run-time. The formula is conceptualized such that event magnitude naturally increases as time progresses, unless appropriate resources are applied. The formula is implemented by the following explanation: Teams begin with 100% score and at every update penalties are assessed for every active event that requires a particular team. The penalty for each event is determined by (1) its initial magnitude and (2) the number of correct resources handling the emergency. Events that have a higher initial magnitude will escalate at a faster rate (i.e, a level 2 fire can become a level 3 fire quicker than a level 1 fire can become a level 3 fire). The quantity of the appropriate resources, namely *R*, will reduce the penalty (fire trucks diminish the severity of a fire). An event is deemed complete by the following:

1. its magnitude at time *t* is non-positive.
2. its magnitude at time *t* is greater than $1+M_0$

In the first case, the event has reached a *success* state. In the second case, the event has reached a *failure* state.

The only factor of interest in the model is the *R* variable because it is the only variable that teams have direct influence. In other words, the value of *R* is directly determined by the resource allocation decisions performed by the teams. In CITIES, *R* reflected the number of resources that were applied to an event. One major difference between NeoCITIES and its precursor is that in teams in CITIES were not required to make resource allocation decisions involving the *resource type*. Resource allocation decisions in CITIES only involved the quantity of those resources. Conversely, NeoCITIES require teams to make decisions involving (1) the quantity of resources and (2) the correct resources. Consequently, *R* is determined as such to account for the correct type AND correct quantity of resources.

4.7 Graphical User Interface

NeoCITIES features an application interface for each role. Separate interfaces are needed to support the functional purposes of each role. However, the layout of the interface is the same and both GUIs share many features. First, each interface has a top panel that includes the time of the virtual world, the team score, and the DHS-advisory threat level. Time is used for enhance a virtual presence by providing a temporal sense. Additionally, a timestamp is used for each event and teams are able to sort events by time in the message tracker. The team score is a feedback component that provides an indication on each team's performance. The DHS terrorist level provides an alert on the current state of the world. The present study used the threat level to inform the teams of potential terrorist attacks. For example, the threat level becomes more severe towards the beginning of the episode to reflect that terrorist attacks are looming and becomes less severe towards the end of the episode to reflect that there are no imminent threats to the city. The threat level update accompanies advisory notices and is encoded in the scenario file. The following graphic provides an illustration of the top panel.

Figure 4-7

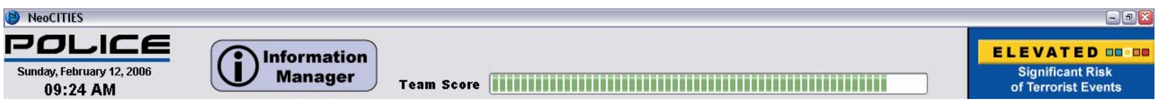


Fig. 4-7: Top Panel

The middle portion of each GUI features two styles of aerial maps. The first map is termed a *Mini-Map* and is used to provide a comprehensive view of the city. The *Main-Map* illustrates a subset of the *Mini-Map*. The *Main-Map* includes a navigational toolbar that allows zooming and panning compatibilities. Active events are displayed in the *Main-Map* as icons. The highlighted region of the *Mini-Map* indicates the area being viewed in the *Main-Map*. Each map is an image file and the locations are referenced by pixels. Additional GIS modules could be implemented to allow GIS functionality.

Figure 4-8

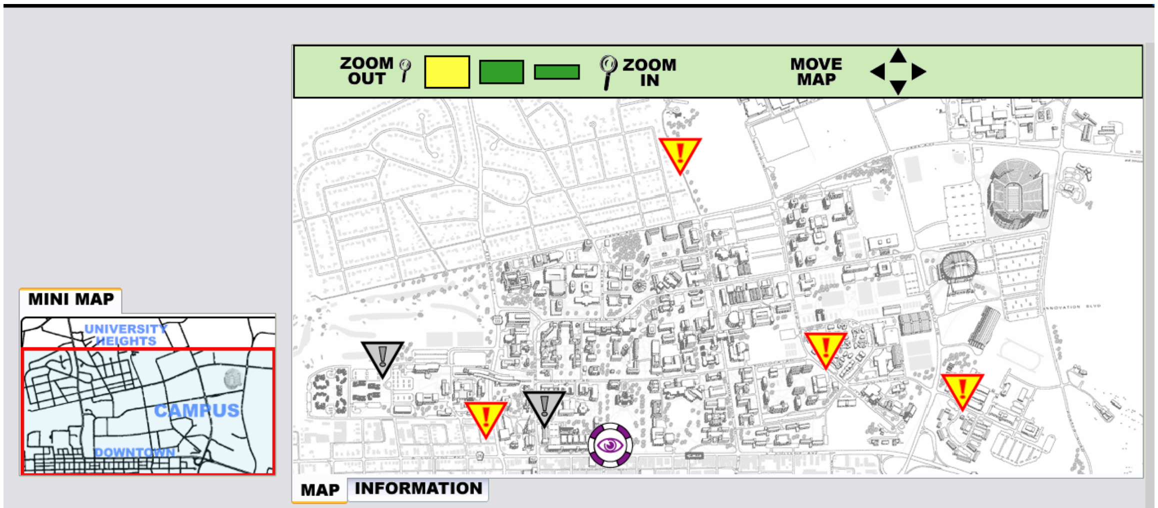


Fig. 4-8: Middle Panel

The bottom portion of the GUI includes an event tracker and the chat panel. A key element of the NeoCITIES interface is the event tracker. The event tracker is a table that provides a structured list of events as they fire in the simulation. The design of the event tracker was informed by ethnographic research of 911 call centers. Similar to the *Main-Map,* he event tracker provides an alternative presentation of the information in the simulation. It also summarizes information in textual form, allowing users to quickly assess the requirements placed upon their team (or other teams). The following graphic illustrates the bottom panel.

Figure 4-9

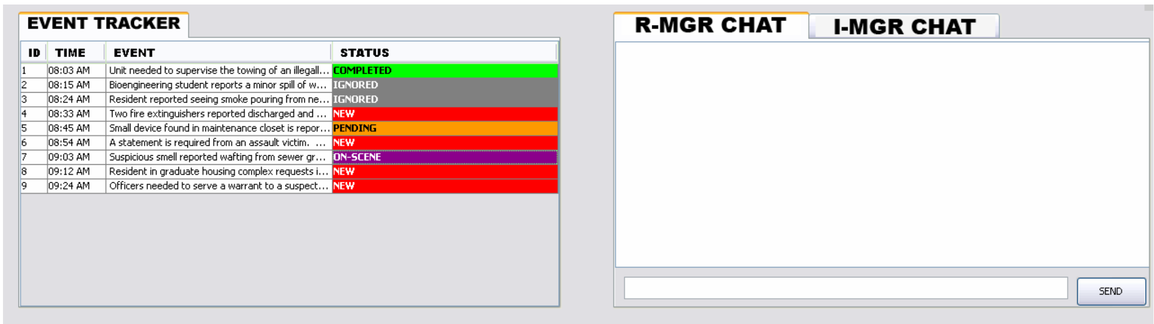


Fig. 4-9: Bottom Panel

4.8 Process

4.8.1 Event Firing Timeline

There are two time variables used in the simulation. Both are expressed in milliseconds. The first is the virtual clock, or *simulation time*. It is the metric of time in the virtual world. *Simulation time* represents the frequency of score computation. For example, for *simulation time* = t, at every t, the score for each team is computed. The second time variable represents the frequency of events that are fired. Formally, this variable is called *sleep time between events* and it is the number of milliseconds in real time that the game engine waits between firing a new event. For example, for *sleep time between events* = *s*, at every *s*, a new event is fired.

The two time variables are independent. The values of *s* and *t* are set at runtime and can be of any value. The present study used *simulation time* = 6000 milliseconds and *sleep time between events* = 30000 milliseconds. These values were determined to be sensible during extensive beta-testing of NeoCITIES. For example, it was observed that events fired at a faster pace than 30 seconds typically overwhelmed the teams because there was insufficient time to process information. Conversely, events that fired slower than 30 seconds typically led to boredom from the participants.

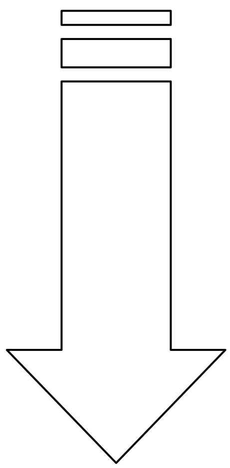
4.8.2 Responding to Events

When an event is fired by the simulation engine, an icon appears on the *Main-Map* and an entry is displayed in the *Event Tracker* of each information manager's interface. The user can then click on either the icon or the entry to view detailed information such as the location and description of the event. This information is displayed in a *Description Dialog* submenu. The information manager can then complete an event report using an *IM Event Description* submenu. The description of the event is automatically entered in the report and the information manager can include additional remarks. After completing the report, the information manager can submit the event report to their resource manager.

The resource manager receives the submitted report pertaining to a particular event from their information manager. The report is displayed in a *RM Event Description* submenu. From there, the resource manager can make resource allocation decisions and send a specified quantity for each desired type. The resources that are sent to the event become unavailable until the event reaches an end state, thus it is imperative that the correct resources are chosen. Once the resources arrive on-scene, the present study used an arbitrary 1 unit of *simulation time*, the resource manager is given feedback on their decision messages via an *On-Scene* submenu. Additionally, once resources arrive on-scene, the resource manager can send additional resources (more of the same type or another type).

Lastly, prior to submitting a report, the information manager can ignore an event. Additionally, prior to allocating resources to an event, the resource manager can ignore any events that were submitted by their information manager.

Figure 4-10



Submenu	**User**	**Function**
Description Dialog	Information Manager	View new events
IM Description Dialog	Information Manager	Submit report
RM Description Dialog	Resource Manager	View submitted report
On Scene Dialog	Resource Manager	View feedback of resource allocation decision

Fig. 4-10: Submenus provided in NeoCITIES

Figure 4-11

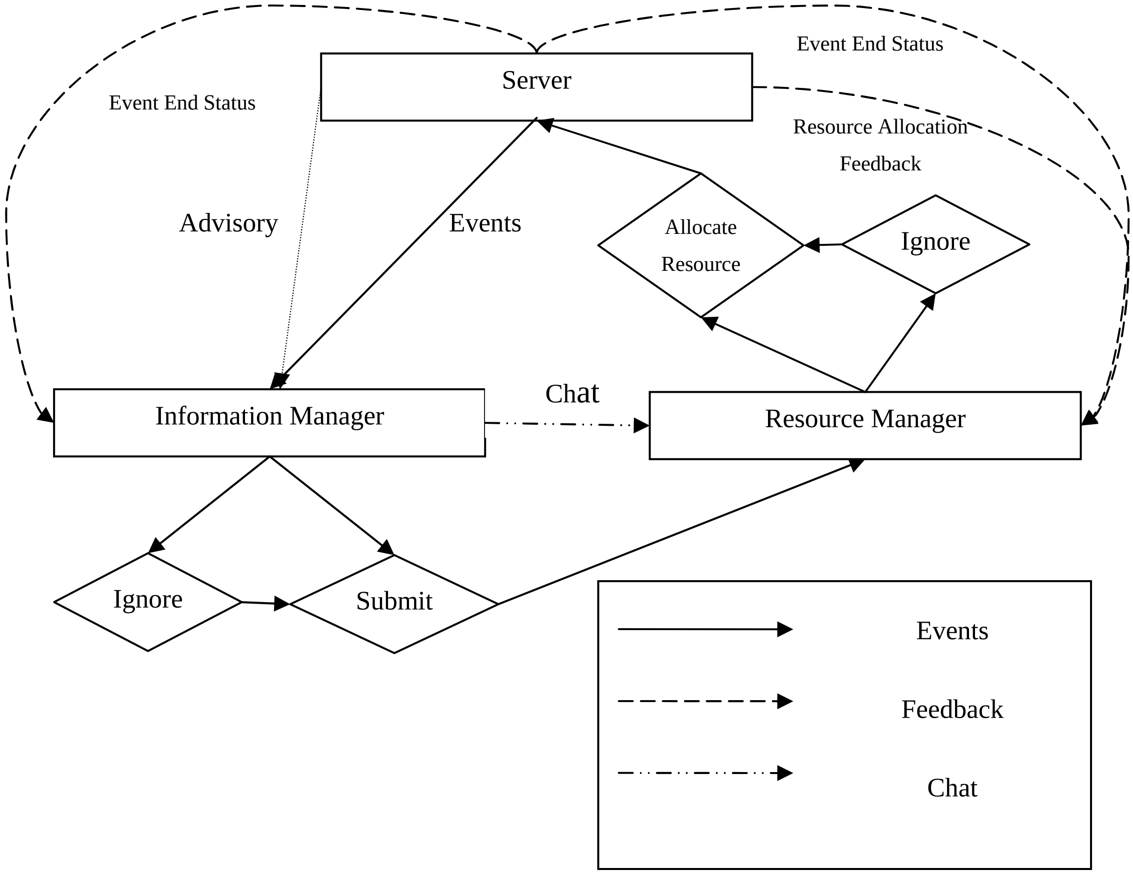


Fig. 4-11: Process Model

4.9 Scenarios

The present study utilized two types of events in the scenario. The first style is terrorist events. Terrorist events have a higher initial magnitude, which escalates faster and requires more resources to address. Additionally, terrorist events are related to other terrorist events and oftentimes have informational nuggets that, if processed correctly by the teams, could be used to identify future terrorist events. Upon general inspection, most terrorist events appear to be like any routine emergency, so it very important that teams share information obtained from correctly handling terrorist events. The second style of events is noise events. They function as distracter events and they are used to (1) keep teams occupied with events (2) require teams to prioritize resources and (3) make it difficult for teams to distinguish from terrorist events. Noise events require teams to communicate and make resource allocation decisions, so teams are engaged with the scenario by noise events. Secondly, noise events require teams to make careful decisions regarding resources because the incorrect prioritization of resources can lead to situations where teams do not have the requisite resources to properly address more severe terrorist events. Lastly, typically noise events appear to be very similar to terrorist events, so teams are not able to quickly distinguish between the two. Consequently, teams are expected to collaborate in order to determine the style of emergency. Each team received 12 noise events and 4 terrorist events.

4.9.1 Structure of Scenario File

The scenario file is an ASCII text file. It is read and loaded at runtime via an initialization method. The author developed a Reference Markup Language utility, or RML, which is a user-friendly XML tool. RML uses tags to denote references and ‘*’ as delimiters. A tag definition file is required that defines each tag and the number of parameters given for each tag. In the scenario file, each event is written on a separate line starting with the <event> tag (or <event_with_feedback> for events that have feedback). Additionally, the <advisory> tag is used to denote governmental advisory notices and DHS terror level updates. Following the tag are parameters, separated by the delimiter, that define the item. For example, the present study used the <event_with_feedback> tag which involved 19 parameters. The parameters are defined as:

<event_with_feedback>*ID*Description*Location*Lifespan*Investigative*Squad*SWAT*Investigative*FireTruck*EMS*Investigate*Bomb*Chem*Level*Escalate*X*Y*Feedback success*Feedback failure*FCM-VALUE

1. The "ID" parameter represents a unique number used to label an event.
2. The "Description" parameter is the event title that appears in the *Description Dialog* submenu.
3. The "Location" parameter denotes the site of the emergency. It is written as names of places, buildings, or street addresses.
4. "Lifespan" parameter is not used for this study.
5. The following parameters, "Investigative*Squad*SWAT*Investigative*FireTruck*EMS*Investigate*Bomb*Chem" denote the number of required units for each type for the given event.
6. "Escalate" parameter is not used for this study.
7. "X" and "Y" are the pixel locations for the event. The client uses this information to display the event icon over the site.
8. The feedback parameters, "Feedback success" and "Feedback failure" are used to provide additional information to the teams once an event has reached an end state. Typically, the feedback obtained from an event reaching a completed state contains helpful information that can be used later. Conversely, feedback obtained from an event that has reached a failure state typically are reminders that stress more vigilance is needed from the CRU. Either or both parameters could be set to null, which means no feedback will be given.
9. FCM-VALUE is used to update a flag in the FCM. The keywords "true" or "false" are used to denote the severity of the DHS threat level. "True" denotes that the DHS level is set to a more severe state and false denotes that the DHS level has not been set to a more severe state.

The present study used an underlying storyline that required the teams to protect a VIP. A visiting senator plans to officially declare his candidacy for the upcoming presidential election. A terrorist organization intends to kidnap the senator by creating a subterfuge that involves detonating a bomb at the site of the speech. During the senator's talk, the bomb is uncovered by police personnel and an evacuation is ordered. The diversion created by the pandemonium of the evacuation allows the terrorists to kidnap the senator and his family. A final shootout is staged involving the FBI and the terrorists. Once the senator is rescued, he needs immediate medical attention or he will die from his injuries.

Events occur during the scenario. The majority of the emergencies are routine events. Terrorist-related events are carefully and strategically placed among the routine events. A DHS alert is sent to the teams prior to the first terrorist event indicating that the VIP is in town and that members of a terrorist organization have been spotted in the city. The successful handling of events also leads to information pertaining to future terrorist-related events that occur later in the scenario.

The final three events are the most important events in the scenario. One event occurs for each team and the teams are required to address those events appropriately resulting in successfully thwarting the terrorists' plans. Specifically, the hazmat team must secure the evacuation of the building and disarm the bomb before detonation, the police are required to assist the FBI in the shootout, and the fire team must send EMS to give aid to the senator.

4.10 Feedback

4.10.1 Resource Allocation Decisions

Feedback is given to the teams by two mechanisms. The first is resource allocation messages, which participants receive after resources arrive on-scene. The resource manager is the only one to receive resource allocation feedback messages. There are three messages: *complete*, *incorrect*, or *incomplete*. A *complete* message occurs when the resource manager selects, at minimum, the correct number of the correct type. An *incorrect* message occurs when the resource manager sends resources to an event that does not require their resources. Finally, an *incomplete* message occurs when the resource manager sends the wrong type or an insufficient number of the correct type. The following figure illustrates each type of feedback.

Figure 4-12

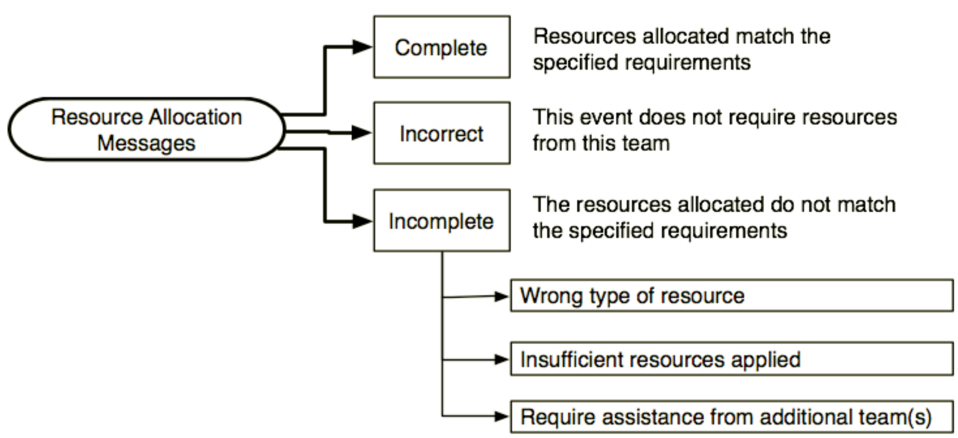


Fig. 4-12: Feedback for resource allocation decisions

4.10.2 Event Termination

Teams also receive feedback when an event terminates. An event has a *Complete* status if the correct amount of resources was applied in an timely fashion. Mathematically, a *Complete* status is reached if the event magnitude is non-positive. Conversely, an event can reach a *Failed* status that occurs if (1) the emergency is left unattended for an extended time period or (2) incorrect resources were applied to the event or (3) an insufficient amount of the correct resource type was allocated. In each case, the event magnitude was allowed to escalate and increase to an amount greater or equal than 1+initial magnitude.

Once an event is terminated, it is removed from the map. Information managers receive the end status of the event and it is displayed in the event tracker. Resource managers will only receive the end status of the events that were submitted to them from their information managers. Additionally, if feedback from the event is available, the resource manager for the team that was required to handle the event will receive a feedback message.

4.10.3 Progress Status

In addition to an end status, each event has a progress status that reflects the actions of the teams. The purpose of the progress status is to inform the team members of the current standing of each event. The progress status of the event reflects the actions of the teams. The simulation tracks the progress of each event and displays the status for each event in (1) the icon on the map and (2) the keyword in the message tracker. The progress status of the event changes throughout the lifespan of the event. The present study used four progress event statuses: *New, Ignore, Pending, and Resources Applied.*

Each event starts with a *New* status and changes based on the information manager's decisions. Unlike the event end status, which is universal, a progress status is relative to the role. For example, when an information manager submits an event report to the resource manager, an event status is displayed as *Pending* on the information manager's screen, but shown as *New* on the resource manager's screen. The following table illustrates the progress status of an event for both roles and how it changes based on the team member's decision.

Table 4-2

Table 4-2: Event status viewed by role

Sub-Menu	Status viewed by information manager	Status viewed by resource manager
Description Dialog	New	
Description Dialog	Ignore	
IM Description Dialog	New	
IM Description Dialog	Ignore	
RM Description Dialog	Pending	New
RM Description Dialog	Ignore	Ignore
On Scene	Pending	Resources Applied

4.11 Lab setup

The following graphics show the experimental space that was used for the present study. The experimental space is designed to simulate and support distributed teams. As such, 6 cubicles were constructed with a desktop machine in each space. The experimental space was divided into 2 sections, 1 row for each role. Participants were randomly assigned to a workstation. Subjects were instructed that verbal chat was not permitted and the only form of chat allowed was through the chat panels on the interface. As such, each person were required to wear headphones to discourage verbal chat. Additionally, wall partitions were setup to (1) provide physical separation and (2) discourage verbal chat.

Figure 4-13

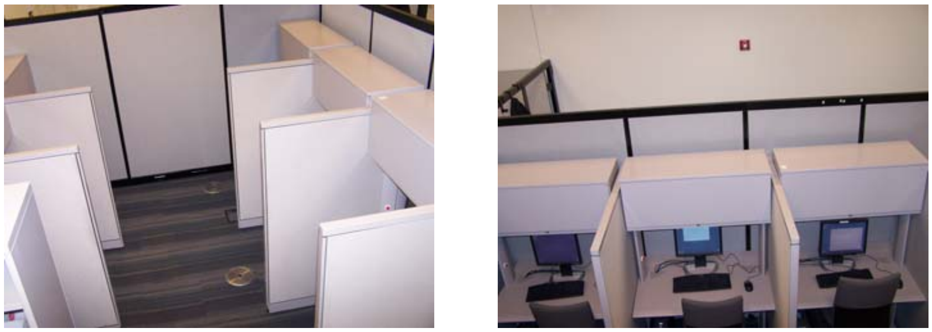


Fig. 4-13: NeoCITIES experimental space

4.12 Software

NeoCITIES is written in java. The development process took over 13 months that included a reiterative design phase and a beta-testing phase. The design phase was a cyclic phase that involved developing NeoCITIES and testing it in-house to ensure that the simulation met the research design requirements. The beta-testing phase entailed volunteers running the application to "break" the simulation. The objective of beta-testing was to run the application several times to uncover bugs, errors, and issues in the simulation.

In all, NeoCITIES includes over 70 class files with over 10000 instructions. The organization of the files is as follows, a com folder that contains the files for the GUI, game engine, and server/client interface. The com folder also contains universal files that contain constant values that are shared throughout the application. Additionally, an util folder is used to store utility files, such as the Scenario Reader module and the main methods for the server and client. Lastly, an img folder contains graphics of the buttons, panels, and maps on the GUI.

NeoCITIES runs on Windows XP. For optimal performance, the screen resolution should be set 1280x1024. At minimum, 3 computers are required to run NeoCITIES, one computer functions as the server and remaining serve as the clients. The instantiation of NeoCITIES requires a series of UDP ports to be opened to handle data transfer. Specifically, the present study used 8 ports (a listener port and a sender port) for the following items: chat, event, score, and fcm.

Batch files are used to invoke NeoCITIES. The batch file for the clients requires the following command line parameters:

java com.util.UDPClient -m F -i "img\\NC_Police.gif" -t Police -r I-MGR -c 6000 5 -p 12 4 0

-m char: module (F is for FCM)

-i filename: name of image used in title bar

-t string: teamname of client

-r string: role of client

-c # #: frequency (in milliseconds) to update clock and quantity of update (in minutes)

-p # # #: number of level 1 events, number of level 2 events, and number of level 3 events

The command line to invoke the server is as follows:

java com.util.UDPServer -s 30000 -t 6000 -k 58 -i 10000 -f scenario\terrorist_attack.txt

-s #: sleep time between events (in milliseconds)

-t #: simulation time (in milliseconds)

-k #: number of increments before NeoCITIES terminates

-i #: initial sleep time (in milliseconds) prior to first event

-f filename: name of scenario file

4.13 Summary

NeoCITIES is a viable research tool that performs scientific examinations on complex teams. NeoCITIES is engineered as a dynamic and robust test-bed for studying teams; specifically teams involved in resource allocation tasks. The software is available as an open-source project. Researchers are able to download and install NeoCITIES and employ it for other team-based experiments.

Chapter 5

Results

Analysis of Variance (ANOVA) methods are utilized to make simultaneous comparisons between two or more means; it is a statistical method that yields values that can be tested to determine whether a significant relationship exists between variables. From statistics, an ANOVA method is used to test a null hypothesis by determining if the means of two or more groups for some response variable are equal. If the group means do not differ significantly, then it is safe to conclude that the independent variable did not have an impact on the dependent variable. For this research, there is one independent variable (or factor), Fuzzy Cognitive Map (FCM), with two levels, presence of the FCM or absence of the FCM; thus, given one factor of interest, a one-way ANOVA is an appropriate method. Furthermore, there are several response variables of interest for this research and by using ANOVA methods; each response variable is evaluated using separate ANOVA analyses.

In order to conduct an ANOVA test, several assumptions about the data must be met. First, the data for each group should be approximately normally distributed. Secondly, the groups should have equal variances. This can be done by conducting a normality test on the data. Additionally, an examination of the histogram of the data is oftentimes used to assess normality. Moreover, equal variances can be checked by examining a graphical display of residuals or through Levene's test, a test for homogeneity of variances. While other tests that can be conducted to test homogeneity of variances, the Levene's test is fairly robust (i.e., not sensitive) to departures from normality.

In addition to the assumption of independence of teams (which is met by random selection of teams), proper usage of ANOVA procedures assumes equal variances for values at the two levels of the independent variable. Data that do not exhibit equal variances should be transformed. Levene's test provides an indication of whether or not the condition of equal variances is met. Levene's test evaluates a null hypothesis where the data exhibits homogeneity of variances, with the alternative hypothesis being that the data does not exhibit homogeneity of variances.

If the null hypothesis of homogeneity of variances is rejected, the conclusion does not preclude further analysis; Levene's test is a screening device and not a definitive test. As such, this research also examines the residuals to assess homogeneity of variances. The residuals should be approximately normally distributed, with a mean of zero. If the residuals are strongly skewed, then the appropriateness of using the ANOVA procedures is brought into question. The data may require transformation before using the ANOVA procedures, or alternative methods should be considered. It should be noted that ANOVA procedures are robust to departures from normality as long as the distributions are symmetric.

5.1 Normality Check

A normality check ensures that an ANOVA analysis can be performed. The normality tests[7] were done in Minitab using Kolmogorov-Smirnov Normality Test. The null hypothesis of a normality test indicates that there is no significant departure from normality**.** When the p-value is more than 0.05, the null hypothesis cannot be rejected and the assumption holds.

Table 5.1: Normality Test

Response Variable	p-value

[7] Prior to data analysis, a histogram plot and a residual plot were generated to examine the overall layout of the data. From these plots, it was apparent that 2 data points were extreme outliers (one both sides). As such, these data points were removed from the sample.

Team Score	0.15
Errors	0.15
Failed Events	0.15

In each case, the p-value is greater than 0.05, thus it is safe to conclude that the data is normally distributed for each response. The following depicts the results from the analysis.

5.2 Descriptive results

The following tables illustrate the mean and standard deviation for each response variable.

Table 5.2: Team Score

	Mean	Std Dev
F	49.625	14.994
N	55.6	9.914

Table 1.3: Number of Errors

	Mean	Std Dev
F	49.14	9.7199
N	35.75	2.062

Table 5.4: Number of Failed Events

	Mean	Std Dev
F	3.714	3.251
N	4.5	1.291

The p-values from the ANOVA analysis are provided in the following table.

Table 5.5: One-Way ANOVA Results

	p-value
Team Score	0.2248
Number of Errors	0.01295
Failed Events	0.33

A chi-squared procedure is used to measure team cognition. The team cognition variable produces results with five levels for both the teams that receive the FCM and the teams that do not receive the FCM. Additionally, since the team cognition represents ordered data (i.e., weak to strong in increasing levels), this variable is treated as an ordinal variable. The variable is not quantitative, yet there is some order to the categories of the response. The data was generated with a code for which a value of 1 represents weak team cognition, 2 represents weak-to-medium team cognition, 3 represents medium team cognition, 4 represents medium-to-strong team cognition, and 5 represents strong team cognition.

The appropriate test to conduct in a case in which the independent variable has two levels and the dependent variable is ordinal is a Mantel-Haenszel chi-squared test. The test is employed to investigate if there is an association between the treatment of the presence of a FCM (or not) and the team cognition level. In essence, the test examines if there is a trend for the values in one treatment group to be higher than (or lower than) the values for the other treatment group across the levels of the response variable, producing somewhat of a linear trend. The null hypothesis states that there is no association between the treatment and the response, and the alternative hypothesis is that there is a linear association between the treatment and the response. Thus, this study tests whether or not there is a linear association between the presence (or not) of a FCM and team cognition. The chi-squared results for the Team Cognition variable are illustrated in the table below.

Table 5.6: Chi-Squared for Team Cognition

Frequency Expected Row Pct.	Weak	Weak-to- Medium	Medium	Medium-To- Strong	Strong	Total

F	10 11.118 55.56	3 2.1176 16.67	1 1.0588 5.56	3 2.6471 16.67	1 1.0588 5.66	18
N	11 9.8824 68.75	1 1.8824 6.25	1 0.9412 6.25	2 2.3529 12.50	1 0.9412 6.25	16
Total	21	4	2	5	2	34

Mantel-Haenszel Chi-Square = 0.6875

5.3 Practical Limitations

A major constraint on the research is the number of students that participated in the study. Painstaking efforts were made to advertise the experiment and solicit participation. For example, the author spoke to 2 classes and disseminated emails to several other classes. Secondly, the author posted fliers in several public locations. Thirdly, the author posted ads on local websites and sent emails to student list-serves. In fact, more time was spent on soliciting participation than experimental hours. Despite the great efforts made to recruit students, a small few participated.

Chapter 6

Discussion

6.1 Significance

The interpretation of data results is as follows: using an $\alpha = 0.05$, the null hypothesis are rejected for the team score and failed_events responses. Thus, there is no statistical significance to show that the mean team performance for teams that receive the FCM is higher than the mean team performance for teams that do not receive the FCM. Similarly, there lacks statistical significance to show that the mean number of failed events for teams receiving the FCM is lower than the mean number of failed events for the teams that did not receive the FCM. However, in the case of the mean number of errors caused by each team, there is statistical evidence to support that the teams with FCMs had a different amount of errors compared to teams that did not receive the FCMs. By examining the means for each group, an interesting conclusion is revealed, the teams with the FCMs created more errors than the teams that did not receive the FCMs. For team cognition, the chi-square value is well above the alpha value, so the null hypothesis cannot be rejected, which means that there is not sufficient evidence to show that the FCMs had an effect.

Theoretically, the null hypothesis states that the means for each group are equal and the alternative maintains that the means for each group are not equal. Thus, for the following hypotheses, the null hypothesis cannot be rejected (the statements are shown not to be true):

- The FCMs will have an effect upon team performance.
- The FCMs will have an effect upon team cognition.
- The FCMs will have an effect upon the number of failed events.

The remaining hypothesis can be accepted:

- The FCMs will have an effect upon the number of errors. For this case, the effect is a negative effect.

The results suggest that the implementation of the FCM caused a distraction to the teams, which prevented effective teamwork. By only identifying events, it appears the R-MGRs have difficulty in 1.) prioritizing events and 2.) communicating the output of the FCMs to the I-MGRs. This is evident in the fact that the FCMs showed no effect upon team performance, team cognition, and the number of failed events, and, more importantly, the number of errors were higher for teams with the FCMs. This suggests that the FCMs may have confused the teams in their decision-making tasks. This interesting phenomenon can be explained in the following manner: it is reasonable to believe that the R-Mgrs' were aroused each time the FCMs highlighted an event as terrorist. This could have created a situation where the R-Mgrs felt overwhelmingly compelled to make an immediate decision. Cognitively, they could have directed all of their attention to the terrorist event and in essence creating a "tunnel vision" effect where their SA was substantially curtailed due to the identification of the terrorist event. In other words, the teams could have believed that a terrorist event is occurring and incorporated a "nothing else matters" position until that particular event is resolved. An anecdote that illustrates this occurs when 2 people are traveling in a car and the passenger, after noticing a police car, yells, "Police". The driver is affected in a couple ways resulting in becoming partially distracted to the task at hand: 1.) the driver may try to locate the police car and 2.) the driver may start to slow down. In the first case, if the driver tries to locate the police car, his/her attention is diverted to the road, which could cause a collision. In the second case, the driver may attempt to slow down, but may not be aware of the speed limit. In either case, in order to assess the presence of the police car and incorporate it within the current driving environment correctly, the driver may require more information than just knowing that a police car is present. First, it is a moot point if the driver isn't speeding or breaking the law (i.e., driving without a seatbelt, etc). Secondly, the police car may be unoccupied or positioned to monitor traffic in another direction. Thirdly, the police car could be driving, in a perpendicular fashion, away from the driver. The contention is that by simply identifying the presence of an instance (a terrorist event or a police car) does not ensure an appropriate response. In other words, just because one knows that something is present does not give one the ability, or the know-how to handle it correctly.

Lastly, it is noteworthy to mention that due to difficulty with getting students, most sessions were run as a 1 team operation. This was an adhoc decision and the impacts were not fully considered. For example, as outlined in the Methodology section, the team was given the same scenario; and even though they were instructed to ignore all events not pertaining to their team, some confusion typically arouse. Without the presence of the other teams, it was observed that certain participants were confounded by their responsibilities and sending resources to the wrong events (i.e., police car to a fire). The presence of the other teams would have generated discussion over the roles of particular events, thus possibly leading to less erroneous decisions.

6.2 Contributions

The primary contribution gained from the present study is NeoCITIES. The research developed a simulation that effectively models the ECM work domain. From a broad perspective, NeoCITIES is an adaptation and extension of its precursor, CITIES (Wellens, 1988). The NeoCITIES scaled world functions as an excellent example of a HCI artifact. A user-centric methodological design was adopted to create the virtual world making it an excellent testbed to run scientific examinations on complex teams. Lastly, NeoCITIES supports communication, collaboration, and coordination and it is operationally relevant for: Command and Control (C2), Command, Control, and Communication (C3), and Command, Control, Communication, and Computers, Intelligence, Surveillance, and Reconnaissance (C4SIR) applications.

From a more specific perspective, NeoCITIES explores socio-cognitive factors that regularly occur in homeland security and defense. Additionally, the deployment of

NeoCITIES enables researchers to launch examinations that can yield a better understanding of the social and psychological outcomes and complex decision-making processes generated by operators within the ECM work domain. The development of NeoCITIES involved a 13 month process with over 10,000 lines of code (instructions) residing in more than 70 files! The research effort included the development of a scoring module that encapsulates resource allocation decisions and a game engine that simulated a virtual world. Additionally, NeoCITIES has its own infrastructure that included two types of architecture, communication architecture and network architecture. Both architectures are needed for NeoCITIES to function as a distibuted team-based simulation. The communication architecture is similar to popular chat technologies (Aol or Yahoo). Likewise, the network architecture provides a server/client system (akin to online games), which links a server that hosts a scenario to client machines that represents roles (Information and Resource Manager) for each team. Furthermore, the development process included the construction of a user interface that allows teams to process information and make complex decisions. Lastly, NeoCITIES gives a high range of experimental control by providing a metric collection module that allows researchers to abstract desired elements of team performance and a scenario module that gives researchers the flexibility to launch various scenarios that are operationally relevant for their given examinations.

NeoCITIES is a research tool that has several key advantages over other simulations, such as TRAP, DDD, and CITIES. First, NeoCITIES is more akin to a real crisis-management task. The scenario presented in NeoCITIES involves a more realistic task and the activities performed by the teams are more similar to the actions of real crisis-management teams. Secondly, NeoCITIES is modular, so it can adapt to other operationally relevant tasks. For example, NeoCITIES could be used to support medical teams operating during an extreme outbreak. NeoCITIES, can adapt so that the scenario, map, team functions, team roles, etc all change to make the simulation operationally relevant for medical work domain. Thirdly, NeoCITIES is scalable, so that it can support various sessions, teams, and roles. Lastly, as a research tool, NeoCITES gives researchers a wide range of experimental control. Researchers are able to modify NeoCITIES in varies ways to make a valid tool for their examination.

Typically, the outcome of a research effort falls into one of the following 4 categories: theoretical experimental work, envisioned design work, proof-of-concept demonstrations, or technical artifact (i.e., deliverable or product). Traditionally, a research is conducted to investigate the validity of specific theories. Envisioned design work is where a technological design is developed from prior scientific investigations; thus the resulting software/interface/computer models follow an informed design approach. Proof-of-concept demonstrations routinely are used to establish the usability of models or new features added to a technical artifact (i.e., new features for a search tool). Lastly, a research can be done to develop a technical artifact, which could be ready for deployment at the conclusion of the research. This research is unique and significant in that it fits within all 4 components. It is representative of 1.) an empirical study to test theories of the efficacy of the FCMs in regards to teamwork, 2.) envisioned designs to illustrate how decisions-aids could be implemented to augment teamwork within the ECM work domain, 3.) a proof-of-concept to demonstrate how the NeoCities simulation is used to test teamwork, and 4.) NeoCities represents a full-scale simulation that could be used for other scientific inquiries. Currently, there are plans to use the simulation as a framework to house scenarios ranging from tsunami (crisis-management), intelligent agents (urban warfare), and notably, an empirical study testing adaptation strategies to design an intelligent group interface

Alternatively, framing the research within the Living Laboratory methodology conveys multiple contributions. Clearly, the research evaluated said theories on decision-aids, specifically, FCMs, and teamwork and models were created in the form of scoring model to assess team performance, a FCM model to represent a decision-aid, and a scenario-architecture model that makes the simulation robust to handle scenarios from various work domains and applications. The “knowledge elicitation” node is satisfied because the simulation was designed from KE sessions performed on intelligent analyst that ran through an early version of NeoCities and document analysis on the ECM work domain to appropriately formulate the design of the simulation. Clearly, the research contributes to the "Scaled World Development" node in that NeoCities is a robust synthetic task environment that is used for scientific examinations. Lastly, the development and integration of the FCMs incorporates the "Prototype Development" node. Furthermore, the research encompassed a problem-centered focus by examining problems within teamwork.

Moreover, the value of the study can be measured from a philosophical fashion. The research incorporates work from 3 disciplines: 1.) development of technology (computer science), 2.) integration of said technologies (engineering), and 3.) experimentation of technology on human users (psychology). As such, the research truly reflects the interdisciplinary nature of the IST program in that it incorporates principles from the said academic disciplines. Each area could have been its own dissertation study, so in essence, the study represents an integrated approach to develop intelligent technologies to assist teams overcome cognitive and communication-related barriers that affect collaborative problem-solving, knowledge management, and decision-making.

6.3 Lessons Learned

There are 3 main lessons gained from this research. The first two lessons place constraints on the experimental session in that it sets requirements on the number of teams and the participants. Specifically, the first constraint would only permit full sessions (i.e., 3 teams per session). The second constraint would require a minimum computer profiency for subjects. Essentially, these prerequisites limit who can participate and it sets the team structure to consist of 3 teams. The third lesson involves the implementation of the FCM. The present study entailed utilizing the agent to function in a passive manner—the agent did not have the capabilities to affect the state of the world. A replication of this study could utilize the agent so that its actions directly affect the virtual world. The following explicates each lesson; it describes the insight that was gained from the present study, how a replication of the present study could incorporate the said lesson, and the implications of including it in a future research.

First, a condition that specifices only 3 teams per session could provide additional opportunities for knowledge sharing and communication. This premise presupposes that the frequency of chat should increase for tasks involving multiple people. Each person can offer their own ideas and perspectives, which could lead to new insights[8]. During several 1 and 2 team sessions, team members often voiced to the proctor that they were confused about who was responsible for a particular event. While each participant was informed of the capabilities of each resource during the briefing activity; participants often seemed to forget the responsiblities of other teams. Thus, in cases where at least one team was absent for a session, participants were often in the dark regarding the responsibilities of the missing team. Obviously, this did not occur during any full session because participants could inquire about the functions of a particular team to the members belonging to the said team. Additonally, during full sessions, it was observed that the presence of the 3rd team provided opportunities for team members to discuss the capabilities of all resources and reach a conclusion on which team was responsible for handling a particular event. While regulating the number of teams for a session could have given the teams additional opportunites to share information, a major implication is that it required further coordination of the sessions with each subject's availability. In other words, sessions would only be open when 6 participants were avaiable. Consequently, this could have led to a smaller population size. A major concern for this research was obtaining a sufficient number of subjects to have statisical significane. As such, placing constraints on the team structure per each session could have lead to fewer sessions, thus resulting in insignificant results. Instead, the present study had open

[8] Alternatively, more people could also lead to more confusion and distraction. However, it is important to note that a general observation was made that when team members discuss relevant issues involving the task, teams often avoided distraction and did not appear to be confused.

sessions daily; students were welcomed to participate during the work day hours and sessions were run when there were an even number of students[9] available.

Secondly, requiring participants to have a minimun amount of computer profiency could have led to different results. It was observed during the briefing activity that several participants had difficulty with performing the basic functions of a computer. Generally, these participants also had a difficult time in performing their duties during the simulation. A replication of this study could include an assessment form to evaluate each individual's computer competency. Only the participants that exceeded a specified threshold would be allowed to participate. The present study did not require any computer proficiency because it would have considerably reduced the number of eligible participants. Additionally, it has been noted in research that there is a considerable difference in computer proficiency amongst gender (Jones, Terrell, & Connors, 2006). College males tend to have higher computer aptitude than their female counterparts. As such, the sample population for the experiment could have been a distorted representation of the general pool of participants. In other words, the sample population could have been skewed due to the high ratio of males-to-females. Consequently, this could have led to results with a gender bias factor.

Thirdly, the implementation of the FCM could have included the agent to serve as an active team member. For example, the FCM could have been designed to suggest resource type, team, and quantity of resources for a given event. Providing additional features to the FCM could have made it would be more effective in augmenting the R-Mgrs decisions. By using the anecdote of the car, the passenger would say "A squad car is present and the speed limit is 45 mph." The design of the FCM included the capability to allocate resources, could lead to a statistically significant effect upon the teams. A minimalist approach was used in designing the implementation of the FCMs.

[9] When 2 participants were available for a session, it was run as a session involving one team. When 4 participants were available, a session onsisiting of two teams was run. Lastly, full sessions occurred when 6 participants were available.

Consequently, the FCMs were only provided with the capability of highlighting events that could be terrorist-related. A serious implication exists when the agent is an active team member. Specifically, trust becomes a concern when the agents send suggestions to the team. Measuring trust is considerably outside the scope of the research as it involves other social factors.

6.4 Future Studies

There are several opportunities for additional research which would provide interesting insight in the role of FCM decision-aids involved in emergency crisis-management work. First, while the present study involved sessions that consisted of one, two, and three teams, it did not assess the differences amongst the sessions. In other words, the present study did not measure the relationship between the number of teams present amongst each dependent variable (team performance, number of errors, number of failed events, team cognition). A replication of the study could involve examining the role the number of teams has on decision-making and team cognition.

Lastly, further studies could include examining the effect computer proficiency has on computer based simulations. A replication of the study could measure the differences in team performance and team cognition in teams that are categorized as having high computer skills compared to teams with low computer skills.

Bibliography

Ackerman, M.S. (1995). Social activity indicators: Interface components for CSCW systems. *Proceedings of the 8th ACM Symposium on User Interface Software and Technology*

Andriole, S. J., & Adelman, L. (1995). *Cognitive systems engineering for user-computer interface design, prototyping, and evaluation.* Hillsdale, NJ: Lawrence Erlbaum.

Balder, D. (2004). Fuzzy cognitive maps their uses as knowledge mapping systems and decision support systems. *Retrieved from* student.science.uva.nl/~d**balder**/media/fcm.pdf

Banini, G.A. & Bearman, R.A. (1998), Application of fuzzy cognitive maps to factors affecting slurry rheology. *International Journal of Mineral Processing, 52* (4), 233-244

Bates, J. (1992). Virtual reality, art, and entertainment. *Presence,* 1(1), 133-138.

Bereiter, C. (1997). Situated cognition and how to overcome it. In D. Kirshner & J. A. Whitson (Eds.), *Situated Cognition* (pp. 281-300). Mahwah, NJ: Erlbaum.

Berg, Marc (1997b): *Rationalizing Medical Work. Decision Support Techniques and Medical Practices*. Cambridge: MIT Press.

Bisantz, A. M., & Vicente, K. J. (1994). Making the abstraction hierarchy concrete. *International Journal of Human-Computer Studies*, 40, 83-117.

Brannick, M.T., & Spector, P. (1990). Estimation problems in the block-diagonal model of the multrait-multimethod matrix. *Applied Psychological Measurement,* 14, 325-339.

Brooks, F. P., Jr. (1987). No Silver Bullet; Essence and Accidents of Software Engineering, *IEEE Computer*, 20(4),10-19.

Brown, J. S., Collins, A., & Duguid, P. (1989) Situated cognition and the culture of learning. *Educational Researcher*, *18*, 32-42.

Buzsaki, G., Horvath, Z., Urioste, R., Hetke, J., & Wise, K. (1992). High frequency network oscillations in the hippocampus. *Science*. 256, 1025-1027.

Bolstad, C. A., & Endsley, M. R. (1999). Shared mental models and shared displays: An empirical evaluation of team performance. *Proceedings of the 43rd Annual Meeting of Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomic Society Press.

Brewer, I. (2002). *Cognitive systems engineering and GIScience: Lessons learned from a work domain analysis.* Paper presented at the Second International Conference on Geographic Information Science, Boulder, CO.

Cai, G., MacEachren, A. M., & Bolelli, L. (2004). *GCCM: Map-mediated collaboration among emergency operations centers and mobile teams.* Paper presented at the Third International Conference on Geographic Information Science, Adelphi, MD.

Cavaleri, S., & Sterman, J. D. (1997). Towards evaluation of systems thinking interventions: A case study. *System Dynamics Review, 16,* 391-415.

Cannon-Bowers, J. A., & Salas, E. (1990, April). *Cognitive psychology and team training: Shared mental models in complex systems.* Paper presented at the Annual Meeting of the Society of Industrial and Organizational Psychology, Miami, FL.

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, *22*(2), 195-202.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision-making. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum.

Carley, K. M. (1997), Organizational adaptation and cognition. *Cognitive Science Proceedings*. Stanford, CA.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies. 41*. 801-849.

Cooke, N. J., & Shope, S. M. (2004). Designing a synthetic task environment. In S. G. Schiflett, L. R. Elliott, E. Salas & M. D. Coovert (Eds.), *Scaled world: Development, validation, and applications* (pp. 261-278). Burlington, VT: Ashgate.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring team knowledge. *Human Factors*, *42*(1), 151-173

Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 83-106). Washington, DC: American Psychological Association.

Coovert, M.D., Penner, L.A., & MacCallum, R. (1989). Covariance structure modeling in personality and social psychological research: An introduction. In C. Hendrick and M. Clark (Eds.), *Review of Personality and Social Psychology: Research Methods in Personality and Social Psychology*, *vol 2* (pp. 291-330). Newbury Park, CA: Sage.

Coull, T., & Rotham, P. (1993). Virtual reality for decision support systems. *AI Expert, 8*(8), 22-25.

Daft, R.L., & Lengel, R.H. (1984). Information richness: A new approach to managerial behavior and organization design. *Research in Organizational Behavior, 6,* 191-233.

Defanit, T. A., & Brown, M. D. (1991). Visualization in scientific computing. *Advances in Computers, 33,* 247-305.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (1993). *Human Computer Interaction.* New York: Prentice Hall.

Druckman, D., & Bjork, A. (Eds). (1994). *Learning, remembering, believing: Enhancing human performance.* Washington, DC: National Academy Press.

Eggleston, R. G. (2002). Cognitive systems engineering at 20-something: Where do we stand? In M. D. McNeese & M. Vidulich (Eds.), *Cognitive systems engineering in military aviation environments: Avoiding cogminutia fragmentosa* (pp. 77-116). Wright-Patterson Air Force Base, OH: HSIAC Press.

Ehrhart, L. S. (1994). Cognitive *systems engineering design handbook* Technical Report (Contract #F30602-92-c-0119). Fairfax, VA: C3I Center, George Mason University.

Elliott, L. R., Dalrymple, M., Regian, J. W., Schiflett, S. (2001). Scaling Scenarios for Synthetic Task Environments: Issues Related to Fidelity and Validity. Proceedings of the 2001 Meeting of the Human Factors and Ergonomics Society, September. Minneapolis, MN.

Endsley, M. R. (1988). Situation awareness global assessment techniques (SAGAT). *Proceedings of the IEEE Aerospace and Electronics Conference*, *3*, 789-795.

Endsley, M. R. (1993). Toward a theory of situation awareness requirements in air-to-air combat fighters. *The International Journal of Aviation Psychology, 3*, 157-168.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 65-84.

Endsley, M. R. & Jones, W. M. (1997). *Situation awareness, information dominance,* *and information warfare* (AL/CF-TR-1997-0156). Wright-Patterson AFB, OH: United States Air Force Armstrong Laboratory.

Ericson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Fuhrmann, S., MacEachren, A. M., Dou, J., Wang, K., & Cox, A. (2005). *Gesture and speech-based maps to support use of GIS for crisis management: A user study.* Paper presented at the Auto-Carto 2005: A Research Symposium, Las Vegas, NV.

Georgopoulos,, V.C., Malandraki, G.A., & Stylos, C.D. (2003). A fuzzy cognitive map apporoach to differential diagnosis of specific language impairment. *Artificial Intelligence in Medicine, 29*, 261-278.

Gettings, M.E., Bultman, M.W., & Fisher, F.S. (2004). A complex systems model approach to quantified minderal resource appraisal, *Environmental Management, 33* (1), 87-98.

Gibson, J. J. (1958). Visually controlled locomotion and visual arientation in animals. *British Journal of Psychology, 49,* 182-194.

Gredler, M. (1994). Crisis Management Simulations. Chapter 6 in Designing and Evaluating Games and Simulations: A Process Approach. Houston: Gulf Publishing Company.

Greeno, J., Smith, D., & Moore, J. (1993). Transfer of situated learning. In D. K. DEtterman and R. J. Sterberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction (pp. 99-167).* Norwood, NJ: Ablex.

Grosky, W. I. (1994). Multimedia information systems. *IEEE Multimedia*, *(1)*1, 12-24.

Hajdukiewicz, J. R., & Vicente, K. (2002). Designing for adaptation to novelty and change: functional information, emergent feature graphics, and higher-level control. Human Factors, 44(4), 592-611.

Hanisch, K. A., Kramer, A. F., & Hulin, C. L. (1991). Cognitive representations, control, and understanding of complex systems: A field study focusing on components of users' mental models and expert-novice differences. *Ergonomics*, 34, 1129-1145.

Hayduk, L. (1987). *Structural Equation Modeling with LISREL: Essential and Advances.* Baltimore, MD: Johns Hopkins Press.

Hinsz, V. B. (1995). Mental models of groups as social systems: Considerations of specification and assessment. *Small Group Research,* 26, 200-233.

Hobbs, B.F., Ludsin, S.A., Knight, R.L., Ryan, P.A., Biberhofer, J., & Ciborowski J.J.H.. (2002). Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems, *Ecological Applications 12*(5), 1548-1565.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes, 62*. 129-158.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Hollan, J. D., Hutchins, E. and Kirsh, D. (2000) Distributed Cognition: a new foundation for human-computer interaction research. *ACM Transactions on Human-Computer Interaction: Special Issue on Hu*Vol. 7, No. 2, pages 174-196

Jain, P. & Agogino, A. M. (1990). Stochastic sensitivity analysis using fuzzy influence diagrams. In R.D. Shachter, T.S. Levitt, L.N. Kanal, J.F. Lemmer (Eds). *Uncertainty in Artificial Intelligence* (pp. 79-92). North-Holland: Elsevier.

Jefferson, T., Jr., McNeese, M. D., Theodorou, E., Ferzandi, L., & Ge, X. (2004). Distributed cognition in spatially distributed decision-making teams. *Human Factors.*

Johnson-Laird, P. (1983). *Mental models.* Cambridge. MA: Harvard University Press.

Jones, A. C. (2004). *The information-technology-people abstraction hierarchy: A tool for complex information system design.* Unpublished Master of Science, The Pennsyvania State University, University Park, PA.

Jones, R.E.T., McNeese, M.D., Connors, E.S., Jefferson, T., & Hall, D.L. (2004). A distributed cognition simulation involving homeland security and defense: The development of NeoCITIES. *Proceedings of the Human Factors and Ergonomics Society 48th Annual meeting*, 631-634.

Jones, R. E. T, Jefferson, T., Jr., Connors, E. S., McNeese, M.D., Obieta, J.F., & Perusich, K. (2005). Exploring Fuzzy Cognitive Maps for use in a Crisis-Management Simulation. *Proceedings of the 14th Conference on Behavior Representation in Modeling and Simulation*, Universal City, CA.

Klein, G. (1993). A Recognition-Primed Decision (RPD) Model of Rapid Decision Making. In G. Klein, J. Orasanu, R. Calderwood & C. E. Zsambok (Eds.), Decision Making in Action: Models and Methods (pp. 138-147). Norwood, NJ: Ablex Publishing Corporation.

Kleinman, D. L. & Serfaty, D. (1989). Team performance assessment in distributed decision-making, in *Proceedings of the Symposium on Interactive Networked Simulation forTraining* (Orlando), 22-27.

Kleinman, D., Pattipati, K., Luh, P., & Serfaty, D. (1992). Mathematical models of team performance: A distributed decision-making approach. In R. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance* (177-218). Norwood, NJ: Ablex Publishing Corporation

Klimoski, R. J., & Mohamed, S. (1994). Team mental model: Construct or metaphor. *Journal of Management*, *20*.

Klir, G., St. Clair, U.H., & Yuan, B. (1997). *Fuzzy Set Theory Foundations and Applications*.

Kosko, B. (1987). Adaptive inference in fuzzy knowledge networks. In *IEEE Conference on Neural Networks* (Vol. 2, pp. 261-186). San Diego: SOS Printing.

Kraiger, K, & Wenzel, L. (1997). Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness. In M. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement* (pp. 63-84). Hillsdale, NJ: Erlbaum.

Langan-Fox, J., Code, S., & Langfield-Smith. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors, 42*(2), 242-271.

Lange, J. D., & Burroughs-Lange, S. G. (1994). Professional uncertainty and professional growth: A case study of experienced teachers. *Teaching and Teacher Education, 10,* 617-631.

Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Levesque, L. L., Wilson, J. M., & Wholey, D. R. (2001). Cognitive divergence and shared mental models in software development project teams. *Journal of Organizational Behavior, 22*(2), 135-144.

Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology, 44,* 585-612.

Liang, D. W., Moreland, R. L., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin, 21,* 384-393.

MacMillan, J., Entin, E.B., Hess, K.P., & Paley, M.J. (2004). Measuring performance in a scaled world: Lessons learned from the Distributed Dynamic Decision-making (DDD) synthetic team task. In S. Schiflett, L. Elliot, E. Salas, & M. Coovert (Eds.) Scaled worlds: Development, validation and applications. Aldershot, England: Ashgate Publishing.

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM*, *37(*7), 31-40.

Magnenat, N. & Thalmann, D. (1991). Complex models for animating synthetic actors. *IEEE Computer Graphics and Applications,* 11(5), 32-44.

Marks, M. A., Mathieu, J. E. & Zaccaro, S. J. (2004). Using scaled worlds to study multi-team systems. In S. G. Schiflett, L. R. Elliott, E. Salas & M. Coovert, Eds.). Scaled worlds: Development, Validation and Applications. Burlington, VT: Ashgate. (279-296).

Mateas, M. (1997). An oz-centric review of interactive drama and believable agents. (*Technical Report CMU-CS-97-156)*. Carnegie Mellon University.

Mathieu, J., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, *85*(2), 273-283.

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp.149-203). San Francisco: Jossey-Bass.

McNeese, M. D. (1996). An ecological perspective applied to multi-operator systems. In O.Brown & H. L. Hendrick (Eds.), *Human factors in organizational design and management-- VI* (pp. 365-370). The Netherlands: Elsevier.

McNeese, M. D. (2000). Socio-cognitive factors in the acquisition and transfer of knowledge. *Cognition, Technology & Work, 2*(3), 164-177.

McNeese, M. D. (2002). Discovering how cognitive systems should be engineered for aviation domains: A developmental look at work, research, and practice. In M. D. McNeese & M. Vidulich (Eds.), *Cognitive systems engineering in military aviation environments: Avoiding cogminutia fragmentosa* (pp. 77-116). Wright-Patterson Air Force Base, OH: HSIAC Press.

McNeese, M. D. (2003). Metaphors and paradigms of team cognition: A twenty year perspective. *Proceedings of the 47th Annual Meeting of Human Factors and Ergonomics Society* (pp. 518-522). Santa Monica, CA: Human Factors and Ergonomic Society Press.

McNeese, M. D., & Brown, C. E. (1986). *Large group displays and team performance: An evaluation and projection of guidlines, research, and technologies (AAMRL-TR-86-035)*. Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Laboratory.

McNeese, M. D., & Hall, D. (2003) "Computer Aided Cognition to Support Problem-Centered Decomposition of Complex Problems," *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society,* Santa Monica: Human Factors and Ergonomics Society Press, 2003, pp. 523-527.

McNeese, M. D., Zaff, B. S., & Brown, C. E. (1992). Computer-supported collaborative work: A new agenda for human factors engineering. In *Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON)* (Vol. 2, pp. 681-686). Dayton, OH: IEEE Aerospace and Electronics Systems Society.

McNeese, M. D., Bautsch H., & Narayanan S. (1999) A framework for cognitive field studies. *International Journal of Cognitive Ergonomics*, *3*(4), 307-331.

McNeese, M. D., Perusich, K., & Rentsch, J. R. (1999). What is command and control coming to? Examining socio-cognitive mediators that expand the common ground of teamwork. *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomic Society* (pp.209-212). Santa Monica, CA: Human Factors and Ergonomics Society.

McNeese, M. D., Perusich K., & Rentsch J. R. (2000) Advancing socio-technical systems design via the living laboratory. *Proceedings of the Industrial Ergonomics Association / Human Factors and Ergonomics Society*(IEA/HFES) *2000 Congress* (pp. 2-610 – 2-613), Santa Monica, CA: Human Factors and Ergonomics Society.

McNeese, M. D., Rentsch, J. R., & Perusich, K. (2000). Modeling, measuring, and mediating teamwork: the use of fuzzy cognitive maps and team member schema similarity to enhance BMC3I decision making. In *IEEE International Conference on Systems, Man, and Cybernetics* (Vol. 2, pp. 1081-1086): IEEE Press.=

McNeese, M. D., Brewer, I., Jones, R. E. T., & Connors, E. S. (2006). Supporting knowledge management in emergency crisis management domains: Envisioned designs for collaborative work. In R. L. Popp & J. Yen (Eds.), Emergent Information Technologies and Enabling Policies for Counter-Terrorism (pp. 255-281): Wiley-IEEE Press.

McNeese, M. D., Zaff, B. S., Citera, M., Brown, C. E., & Whitaker, R. (1995). AKADAM: Eliciting user knowledge to support participatory ergonomics. *The International Journal of Industrial Ergonomics, 15*(5), 345-363.

McNeese, M. D., Connors, E. S., Jones, R. E. T., Terrell, I. S., Jefferson, T., Jr., & Brewer, I.,. (2005). Encountering computer-supported cooperative work via the

Living Lab: Application to emergency crisis management. In *Proceedings of the 11th International Conference of Human-Computer Interaction*.

McNeese, Bains, Brewer, Brown, Connors, Jefferson, Jones, & Terrell (2005). The Neocities simulation: Understanding the design and experimental methodology used to develop a team emergency management simulation. *Proceedings of the Human Factors and Ergonomics Society 49th Annual meeting*. Orlando, FL: Human Factors and Ergonomics Society.

Moreland, R. L., & Myaskovsky, L. (2000). Exploring the performance benefits of group training: Transactive memory or improved communication? *Organizational Behavior and Human Decision Processes, 82*(1), 117-133.

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 57-84). Thousand Oaks, CA: Sage.

Moreland, R., Argote, L., & Krishnan, R. (1998). Training people to work in groups. In L. H. R. S. Tindale, J. Edwards, E. J. Posvac, F. B. Byant, Y. Sharez-Balcazar, E. Henderson-King, & R. Myers (Ed.), *Theory and Research on Small Groups* (pp. 37-60). New York: Plenum.

Ndousse, T. D. & Okuda, T. (1996). Computational intelligence for distributed fault management in networks using fuzzy cognitive maps. In *Proceedings of the IEEE international conference on communications converging technologies for tomorrow's application (1558-1562).* NY: IEEE.

Nosek, J., & McNeese, M. D. (1997). Augmenting group sensemaking in ill-defined, emerging situations. *Information, Technology and People, 10*(3), 241-252.

Olson, J. R., & Reuter, H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems. 4,* 152-168.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In G. A. Klien, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Mental models and methods* (pp. 327-345). Norwood, NJ: Ablex.

Özesmi, U. (1999). Ecosystems in the Mind: Fuzzy Cognitive Maps of the Kizilirmak Delta Wetlands in Turkey, http://env.erciyes.edu.tr/Kizilirmak/UODissertation/uozesmi5.pdf

Özesmi, U. & S. Özesmi. (2003). A participatory approach to ecosystem conservation: Fuzzy cognitive maps and stakeholder group analysis in Uluabat Lake, Turkey. *Environmental Management 31*(4), 518-531.

Özesmi, U. & S.L. Özesmi. (2004). Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. *Ecological Modeling 176*(1-2), 43-64.

Papageorgiou, E. I. & Groumpos, P. P. (2004). Two-stage learning algorithm for fuzzy cognitive maps. In *Proceedings IEEE International Conference on Intelligent Systems:* IEEE.

Parenthoën, M., Reignier P., Tisseau J. (2001). Put fuzzy cognitive maps to work in virtual worlds. In *Proceedings Fuzz-IEEE'01*.

Perlin, K. & Goldberg, A. (1995). Improve a system for scripting interactive actors in virtual worlds. *Computer Graphics.* 29(3), 3-11.

Perusich, K., & McNeese, M. D. (1997). Using fuzzy cognitive maps to define the search space in problem solving. In G. Salvendy, M. Smith, & R. Koubek (Eds.), *Design of Computing Systems: Cognitive Considerations.* (pp. 805-809). Amsterdam, The Netherlands: Elsevier Science Publ. B.V.

Perusich, K., & McNeese, M. D. (1998). *Understanding and modeling information dominance in battle management: Applications of fuzzy cognitive maps.* AFRL-TR-98-0040. Wright-Patterson Air Force Base, OH: Air Force Research Laboratory.

Perusich, K. A., & McNeese, M. D. (in press). Using fuzzy cognitive maps for knowledge management in a conflict environment. *IEEE Systems, Man and Cybernetics.*

Perusich, K., McNeese, M. D., & Rentsch, J. (1999). Qualitative modeling of complex systems for cognitive engineering. In *Proceedings of the SPIE* (Alex F. Sisti, Ed.), *Vol. 3996,* 240-249.

Pettersson, M., Randall, D., & Helgeson, B. (2002). Ambiguities, awareness and economy: A study of emergency service work. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (pp. 286-295). New York: ACM Press.

Raser, J.C. (1969). *Simulations and society: An exploration of scientific gaming.* Boston: Allyn & Bacon.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). Cognitive Systems Engineering. New York: J. Wiley.

Reddy, M., & Dourish, P. (2002). A finger on the pulse: Temporal rhythms and information seeking in medical work. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (pp. 344-353). New York: ACM Press.

Redding, R. E., & Cannon, J. R. (1992). Expertise in air traffic control (ATC): What is it, and how can we train for it? In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1326-1330). Santa Monica, CA: Human Factors and Ergonomics Society.

Rentsch, J. R., & Hall, R. J. (1994). Members of great teams think alike: A model of team effectiveness and schema similarity among team members. In M. M. Beyerlein & D. A. Johnson (Eds.), *Advances in interdisciplinary studies of work teams: Theories of self-managing work teams* (Vol. 1, pp. 223-261). Greenwich, CT: JAI Press, Inc.

Rentsch, J., Burnett, D. D., McNeese, M. D., & Pape, L. (1999). Team member interactions, personalities, schemas, and team performance: Where's the connection? *Paper presented at the 14th Annual Conference of the Society of Industrial and Organizational Psychology*, Atlanta, GA.

Rentsch, J. R., McNeese, M. D., Pape, L. J., Burnett, D. D., Menard, D. M., & Anesgart, M. N. (1998). *Testing the effects of team processes on team member schema similarity and team performance: Examination of the team member schema similarity model (AFRL-HE-WP-TR-1998-0070)* (Technical Report). Wright-Patterson Air Force Base: Human Effectiveness Directorate, Crew Systems Interface Division.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits on the search for mental models. *Psychological Bulletin*, *100*(3), 349-363.

Rulke, D. L., & Rau, D. (1997). Examining the encoding process of transactive memory in group training. *Academic Management Proceedings*, 349-353.

Salas, E., & Fiore, S. M. (2004). Why team cognition? An overview. In E. Salas & S. M. Fiore (Eds.), *Team Cognition: Understanding the factors that drive process and performance* (pp. 3-8). Washington, DC: American Psychological Association.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance (*pp.3-29). Norwood, NJ: Ablex.

Salomon, S. (1993). *Distributed cognitions: Psychological and educational considerations*. Cambridge, UK: Cambridge University Press.

Sandahl, T.I. (1999): *From Paper to Digital Documents: Challenging and Improving the SGML Approach*, PhDThesis, Dept. of Informatics, Univ. of Oslo

Sanderson, P. M., McNeese, M. D., & Zaff, B. S. (1994). Handling complex real world problems with two cognitive engineering tools: COGENT and MacSHAPA. *Behavior Research Methods, Instruments, & Computers*, *17* (2), 117-124.

Sarter, N. & Woods, D.D. (1995). How in the world did we get into that mode? Mode error and awareness in supervisory control. *Human Factors*, vol. 37, pp. 5-19.

Scott, W. B. (1991). Computer simulations place models of humans in realistic scenarios. *Aviation Week & Space Technology, 134*(25), 64-65.

Siraj, A. Bridges, S. M. Vaughn, R. B. (2001). Fuzzy cognitive maps for decision support in an intelligent intrusion detection system. *North American Fuzzy Information Processing Society.*

Smith, E & Eloff, J. (2000). Cognitive fuzzy modeling for enhanced risk assessment in health care institution. *IEEE Intelligent Systems and Their Applications (*69-75).

Smith-Jentsch, K. A., Campbell, G. E., Milanovich, D. A., & Reynolds, A. M. (2001). Measuring teamwork mental models to support training needs assessment, development, and evaluation: Two empirical studies. *Journal of Organizational Behavior*, *22*(2), 179-194.

Stanney, K. M., Mourant, R. R, & Kennedy, R. S (1998). Human factors issues in virtual environments: A Review of the literature. *Presence (7)*4, 327-351.

Stylios, C., & Groumpos, P. P. (2000). Fuzzy cognitive maps: A soft computing technique for intelligent control. In *Proceedings of the 2000 IEEE International Symposium on Intelligent Control (*pp. 97-102). Patras, Greece.

Stylios, C. D., Georgopoulos, V. C., & Groumpos, P. P. (1997). The use of fuzzy cognitive maps in modeling systems. *The 5th IEEE Mediterranean Conference on Control and Systems*. Paphos, Cyprus: IEEE.

Stytz, M. R., Frieder, G., & Frieder, O. (1991). Three-dimensional medical imaging: algorithms and computer systems. *ACM Computing Surveys, 23*(4), 421-496.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. New York: Cambridge University Press.

Taber, R. (1991). Knowledge processing with fuzzy cognitive maps. *Expert Systems with Applications.* 2, 83-87.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review,* 42(55) 189-208.

Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science, 17*, 7-48.

Vicente, K.J. (1999). *Cognitive work analysis.* Mahwah, NJ: Lawrence Erlbaum.

Vicente, K., & Burns, C. (1995). A field study of operator cognitive monitoring at Pickering Nuclear Generating Station - B (CEL 95-04). Toronto, Ontario, Canada: Cognitive Engineering Laboratory, University of Toronto.

Weick, K. E. & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly, 38:* 357-381.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.

Wellens, A. R. (1993). Group situational awareness and distributed decision making: From military to civilian applications. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 267-291). Hillsdale, NJ: Lawrence Erlbaum.

Wellens, A. R., & McNeese, M. D. (1987). A research agenda for the social psychology of intelligent machines. In *Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON)* (Vol. 3, pp. 944-949). Dayton, OH: IEEE Aerospace and Electronics Systems Society.

Wellens, A. R.. & Ergener, D. (1988). The C.I.T.I.E.S Game: A computer-based situation assessment task for studying distributed decision making. *Simulation & Games, 19*(3): 304-327.

Whitaker, R. D., Selveraj, J. A., Brown, C. E., & McNeese, M. D. (1995). *Collaborative design technology: Tools and techniques for improving collaborative design (AL/CF-TR-1995-0086)*: Wright-Patterson Air Force Base, Armstrong Laboratory, OH.

Wilson, J. R. & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors, 31,* 617-634.

Wilson, D., McNeese, M. D., & Brown, C. E. (1987). Team performance of a dynamic resource allocation task: Comparison of shared versus isolated work setting. In *Proceedings of the 31st Annual Meeting of the Human Factors Society* (Vol. 2, pp. 1345-1349). Santa Monica, CA: Human Factors Society.

Wohl, J. G. (1981). Force management decision requirements for Air Force Tactical Command and Control. *IEEE Transactions on Systems, Man, and Cybernetics*. *SMC-11*, 618-639.

Woods, D. D. (1993). Process-tracing methods for the study of cognition outside of the experimental psychology laboratory. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods (*pp. 228-251). Norwood, NJ: Ablex.

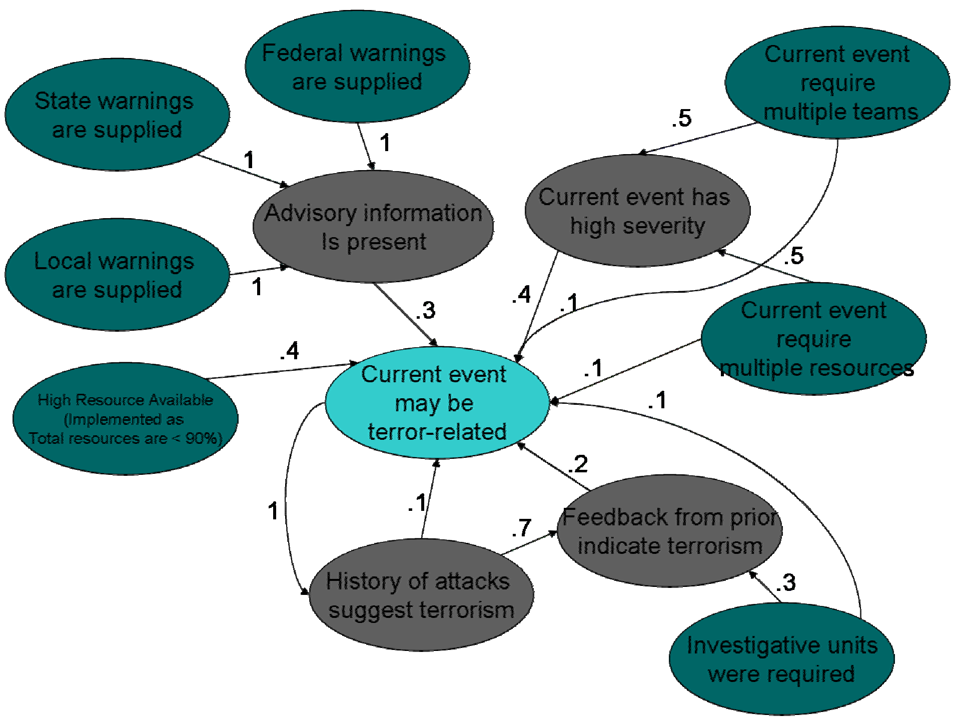
Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers, and hindsight*. Wright Patterson AFB, OH: CSERIAC.

Xirogiannis, G, Stefanou, J., & Glykas, M. (2004). A fuzzy cognitive map approach to support urban design. *Expert Systems with Applications, 26*, 257-268.

Zsambok, C. E., Beach, R. B., and Klein, G. (1992). *Literature Review of Analytical and Naturalistic Decision Making*. Naval Command, Control and Ocean Surveillance Center, San Diego.

Appendix A

Fuzzy Cognitive Map



Each team has a FCM agent that monitors several environmental concepts and each time the information manager submits an event report to the resource manager, the FCM assesses the likelihood of the said event to involve terrorism. If the FCM determines the event to involve terrorism, it highlights the event in black in the event tracker. There are 7 environmental states that the FCM uses as input nodes.

The FCM uses the following concepts:

- Low resources are available (LowRes): This concept is an input concept that becomes true when the team has more than 10% of their resources in the field attending to emergencies.
- Advisory information is present (Adv): This concept becomes true when one of the following is true:
    - Local warnings have been supplied to the teams
    - State warnings have been supplied to the teams
    - Federal warning have been supplied to the teams
- Current event has high severity (HiEventSev): This concept is set to be true if both of the following environmental concepts are true:
    - Current event requires multiple teams
    - Current event requires multiple resources types
- History of attacks suggests terrorism (Hist): This concept becomes true when previous events are assessed to involve terrorism.
- Feedback from prior events indicate terrorism (Feed): This concept is true if
    - History of attacks suggest terrorism
    - Investigative units are required to attend current event

The implementation of the FCM follows a series of logical statements. Perusich (1999) reports a reduction technique to simplify a FCM. As such, an event is determined to involve terrorism by the following:

- LowRes, Adv, and HiEventSev are all true
- Adv, HiEventSev, Feed, and Hist are all true
- HiEventSev, Feed, and His are all true and LowRes or Adv is true
- Feed, LowRes, Adv, and Hist are all true

Appendix B

Data

Presence	Team Performance	Number of Errors	Failed Events	Average Time To Decision
F	39	-	-	4.3125
N	49	-	-	4.145834
N	62	-	-	3
N	50	-	-	4.5
N	60	38	4	5.3125
N	48	48	5	2.0625
N	63	36	3	5.3125
N	28	39	10	4.333334
N	54	33	5	9.583333
N	55	37	4	5.5625
N	39	36	6	5.375
F	27	51	10	2.666667
F	61	45	0	8.8125
F	47	44	4	6.020834
F	31	57	9	-1
F	34	52	6	10.39583
F	35	68	5	-1
F	53	59	3	4.3125
F	59	43	4	0.979167
F	47	46	-	2.041667
F	67	42	1	4.5625
F	63	43	-	2.520833
F	39	45	6	4.3125
F	40	40	7	4.145834
F	66	44	1	3
F	55	10	5	4.5

Presence	Team Cogntion
N	4
N	1
N	1
F	H
N	1
N	H
N1	
N	1
N	1
N	1
N	H
N	H
N	1
N	3
N	1
N	4
N	1
F	1
F	2
F	1
F	1
F	1
F	1
F	1
F	1
F	2
F	4
F	2
F	1
F	4
F	1
F	1
F	3
F	4
N	2

VITA
Rashaad Ennis Theophus Jones

Education

Pennsylvania State University, University Park, PA
Major: Information Sciences & Technology
Degree: Ph.D., 4th year (ABD)
Focus: Human-Computer Interaction, Cognitive Systems Engineering, Decision-Aids, and Modeling and Simulation
Awards: Applied Research Laboratory Exploratory and Foundational Graduate Fellowship, Lockheed Martin Graduate Fellow.

Morgan State University, Baltimore, MD
Major: Electrical and Computer Engineering
Degree: BS (cum laude), 2002
Focus: Neural Network Design and Development, Computer Science, and Digital Signal Processing
Awards: Honors Scholarship

Research Experience

Research Assistant,
User Science Engineering Laboratory
Pennsylvania State University
University Park, PA

- Programmer
- Systems Designer for the NeoCITIES simulation
- Developed mathematical algorithm that computes team performance for the NeoCITIES simulation

Research Assistant
Synthetic Environment Applications Laboratory
Pennsylvania State University
University Park, PA

- Systems Designer
- Programmer

Work Experience

Software Developer
Agilent Technologies
Santa Rosa, CA

Programmer
Hughes Space and Communications
Los Angeles, CA

Publications

Referred Journal Articles

McNeese, M., Jones, R. E., Brewer, I., Connors, E., Bains, P., Terrell, I., & Jefferson T. (accepted-in prepartion). Understanding work in the intelligence community: Acquiring knowledge from expert image analysts. To appear in the *International Journal of Cognition, Technology, and Work*

Conference Papers

Jones, R. E. T., Jefferson, Jr., T., Connors, E. S., McNeese, M. D., and Obieta, J. F. (2005). Exploring fuzzy cognitive maps for use in a crisis-management simulation. *Conference on Behavior Representation in Modeling and Simulation.*

El-Nasr, M.S., Jones, R. E. T., & McNeese, M.D. (2004). A scalable and extensible interactive scenario architecture for distributed command and control simulations. *Proceedings of 2004 Command and Control Research and Technology Symposium*. Vienna, VA: Command and Control Research Program..

Jones, R.E.T., McNeese, M.D., Connors, E.S, Jefferson T, & Hall, D. L. (2004). A Distributed cognition simulation involving homeland security and defense: The Development of NeoCITIES. *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: Human Factors and Ergonomics Society.

Book Chapters

McNeese, M. D., Brewer, I., Connors, E., & Jones, R. (2006). Utilizing intelligent-group-interfaces to prevent terrorist attack: A distributed cognition perspective. In Yen, J. & Popp, R. (Eds.). *21st Century enabling technologies and policies for counter-terrorism.* Englewood Cliffs, NJ: Prentice-Hall.

Journal Articles Reviewed

IEEE Intelligent Systems.

ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE).