Understanding Adaptive Behaviors in Complex Clinical Environments

by

Mithra Vankipuram

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2012 by the
Graduate Supervisory Committee:

Robert Greenes, Co-Chair
Vimla Patel, Co-Chair
Diana Petitti
Marshall Smith
Valentin Dinu

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

Critical care environments are complex in nature. Fluctuating team dynamics and the plethora of technology and equipment create unforeseen demands on clinicians. Such environments become chaotic very quickly due to the chronic exposure to unpredictable clusters of events. In order to cope with this complexity, clinicians tend to develop ad-hoc adaptations to function in an effective manner. It is these adaptations or "deviations" from expected behaviors that provide insight into the processes that shape the overall behavior of the complex system. The research described in this manuscript examines the cognitive basis of clinicians' adaptive mechanisms and presents a methodology for studying the same.

Examining interactions in complex systems is difficult due to the disassociation between the nature of the environment and the tools available to analyze underlying processes. In this work, the use of a mixed methodology framework to study trauma critical care, a complex environment, is presented. The hybrid framework supplements existing methods of data collection (qualitative observations) with quantitative methods (use of electronic tags) to capture activities in the complex system. Quantitative models of activities (using Hidden Markov Modeling) and theoretical models of deviations were developed to support this mixed methodology framework.

The quantitative activity models developed were tested with a set of fifteen simulated activities that represent workflow in trauma care. A mean recognition rate of 87.5% was obtained in automatically recognizing activities.

Theoretical models, on the other hand, were developed using field observations of 30 trauma cases. The analysis of the classification schema (with *substantial* inter-rater reliability) and 161 deviations identified shows that expertise and role played by the clinician in the trauma team influences the nature of deviations made ($p<0.01$).

The results shows that while expert clinicians deviate to innovate, deviations of novices often result in errors. Experts' flexibility and adaptiveness allow their deviations to generate innovative ideas, in particular when dynamic adjustments are required in complex situations. The findings suggest that while adherence to protocols and standards is important for novice practitioners to reduce medical errors and ensure patient safety, there is strong need for training novices in coping with complex situations as well.

To my parents

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisors, Dr. Vimla Patel and Dr. Robert Greenes, for their support and encouragement in the final stages of my PhD journey. The past four years have had some significant highs and lows during which I have found their guidance to be invaluable. Dr. Greenes has been a very ready and supportive guide. His trust in my abilities as a researcher has given me the confidence I need to finish my dissertation work. Dr. Patel has been a co-author on many of my publications. Her guidance through the formative years of my PhD has been wonderful. In addition to mentorship, she has given me numerous opportunities to excel as an individual and I thank for her for that.

In addition to my chairs, this dissertation would not have been possible without the added support of Dr. Diana Petitti and Dr. Valentin Dinu. They have kept me motivated throughout my four years in the Biomedical Informatics program. On many occasions, both Dr. Petitti and Dr. Dinu have taken the time to listen to the various trials and tribulations that I encountered in last two years. They have been fantastic role models for research and I appreciate their support, both as guides and committee members.

I would also like to acknowledge the contributions of Dr. Kanav Kahol and Dr. Mark Smith in my professional development, especially during the early years of my graduate studies. Dr. Kahol and Dr. Smith introduced me to the world of virtual-reality simulations that sparked my interest in biomedical informatics six years ago. In addition to the variety of projects they introduced me to they have taught me a lot about research methods and communicating with scientists from varied backgrounds, both of which are critical skills for a biomedical informatician.

I would also like to thank the IRB (Institutional Review Board) coordinators and all the trauma clinicians at Banner Good Samaritan who helped me complete my experiments. I would especially like to acknowledge the efforts of Dr. Vafa Ghaemmaghami and Dr. Denise Drumm-Gurnee. Their ready support this past year helped me cross the final hurdles in my dissertation work.

Finally, a majority of this work was supported by a grant from James S. McDonnell Foundation, "Cognitive Complexity and Error in Critical Care" to Vimla Patel. I thank the foundation for their support.

The PhD is an arduous process that challenges students mentally, physically and emotionally. Without the support of friends and family completing this journey would have taken much longer. I would like to thank Michael Astrauskas for his unending support and patience over the past five years. My colleagues, Aaron Ashby, Prabal Khanal and Akshay Vankipuram have in many different ways helped and guided me through this journey. Lastly, my parents have always believed in my abilities and supported me through thick and thin. I am appreciative of all of their support and encouragement.

Thank you.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

Page

INTRODUCTION

In an ideal scenario, hospital systems would deliver care in a timely manner to a large number of patients with a variety of diseases. There would be no hospital-acquired infections, staff-related oversights or prescription errors that result in complications. As patients, we would want to be treated in such an institution. Insurance companies, a principal (financial) driving force in the healthcare industry, would prefer that their customers visit hospitals where reduced complications result in shorter hospital stays and lower overall costs due to better outcomes. From the clinicians' point of view, working in a safe and efficient system increases their reputation and work morale. Such an institution would attract a large volume of patients. This will result in greater reimbursement, which would make a strong case for improving quality of care from a business perspective as well. Although not all the features described may be practicably achievable, quality of care is a fundamental concept that is critical to building a safe, cost-effective and sustainable healthcare system.

Unlike other domains such as aviation and nuclear power [1], medicine continues to rely on individual error-free performance as opposed to designing systems around principles of safety [2]. In order to build safer systems, understanding the cognitive mechanisms that drive errors and other adaptive deviations in complex systems is needed. The research described in this dissertation work elucidates the primary barriers for understanding complex healthcare systems and presents a methodology for studying the same. The goal of the research is to develop methods for understanding the nature of errors and other deviations that may occur in complex systems, so that the system can be redesigned around theoretically grounded principles of safe practice.

Current State of Quality of Care and Patient Safety

The Institute of Medicine (IOM) released a number of reports that have increased the public awareness about quality in healthcare and patient safety. The 2000 report "To Err Is Human" [3] drew attention to the vulnerability of the healthcare system to medical errors. This report estimated that in the United States (US) alone, 44,000 to 98,000 lives were lost annually due to preventable medical errors. These figures were based on injury rates estimated by two key studies that performed retrospective reviews of medical records [4]. The significance of this statistic lies in the fact that it is more likely to be an under-estimate. Chart review processes catch only errors reported in the hospital setting, which is only a small part of the care continuum [5]. Leape compared the reported figures to "three fully loaded jumbo jets crashing every-other day" [6]. In any field other than healthcare, such a high error rate would be unacceptable.

This report made a number of recommendations for reducing errors. These included setting national goals for patient safety, developing evidence-based knowledge, understanding the cause for errors and encouraging voluntary error reporting. A 2001 IOM report, "Crossing the Quality Chasm" [7], provided broad recommendations for the future of healthcare, stating that systems should aim to be "safe, effective, patient-centered, timely, efficient and equitable". Together, these two IOM reports have largely served to draw attention to the critical task of error prevention, enlist the support of stakeholders, and had impact on practices in all levels of care [8].

Following these reports, a variety of interventions have been implemented at various healthcare centers across the United States. These interventions include incorporation of computer-based provider order entry (CPOE) systems, protocol adoption and team training, to name a few [8]. There is evidence of small but significant improvement in patient safety at various institutions. Fewer patients die from medication errors [9, 10], and infection rates have been reduced due to the use of protocols and checklists for specific procedures [11, 12].

Despite evidence of some improvement, health systems nation-wide did not show an anticipated (and necessary) overall level of progress in improving patient safety (IOM recommended reducing errors by 50% within 5 years) [5, 8]. One of the reasons for the lack of sufficient improvement is that errors are often not caused by individual clinicians or practices, but are the result of some fundamental systemic problems. Leape and Berwick [8], in their assessment of barriers to quality improvement, suggested that system complexity compounded by professional fragmentation and a hierarchical authority structure, may dissuade the creation of a culture of individual accountability and coordinated teamwork, both attributes of a safe system. Therefore, in order to understand the root cause of errors, researchers would first need to investigate how clinicians behave and interact within the complex healthcare system.

Healthcare as a Complex Adaptive System

Recent research has approached the study of social systems, such as clinical environments, using scientific theory based on complex adaptive systems (depicted in Figure 1) [13]. Plesk and Greenhalgh define complex systems as *"a collection of individual agents with freedom to act in ways that are not always predictable, and whose actions are interconnected so that one agent's actions, change the context for other agents"* [14, 15]. Such systems typically involve a dynamic network of entities acting simultaneously, while continuously reacting to each other's actions [16, 17]. Complex systems are adaptive, unpredictable, and inherently non-linear [18]. Inconsistencies, tension, and anxiety are by-products of such environments [19, 20].

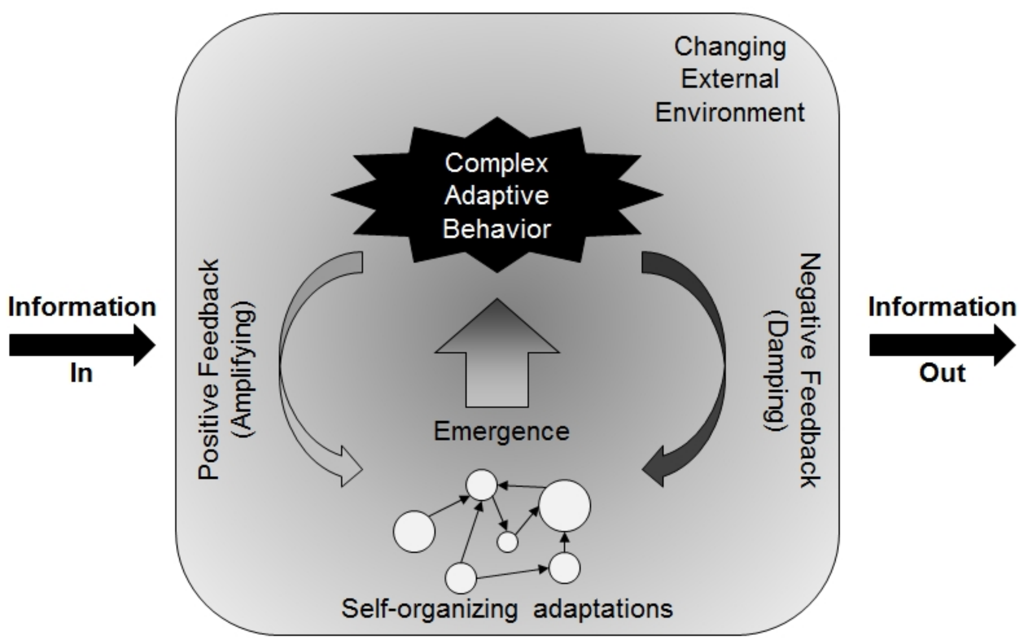

Figure 1. Overview of complex adaptive systems

Figure 1, an illustration adapted from “Complexity: Life at the Edge of Chaos” [21], depicts the key elements of a complex system. Typically, a large amount of information is utilized and generated by the system. In addition to the systems having an environment in which information and knowledge are dynamically changing, the overall behavior of such systems is also affected by the positive and negative feedback received through interactions among the individuals working in these systems. In order to cope with an unpredictable and dynamic environment, individuals tend to develop ad-hoc adaptations, which may eventually evolve into strategies. This “emergence” of stable strategies makes up the overall behavior of a complex adaptive system.

Clinical environments, such as emergency departments (ED), intensive care units (ICU) and trauma critical care are particularly complex and dynamic. Changes in staff (due to shift changes, rotations of staff, or departure/new hires) continually alter team dynamics and the plethora of technology and equipment create unforeseen demands on clinicians. These characteristics allow clinical environments to be categorized as complex adaptive systems.

Data Collection in Complex Systems

In addition to the challenges faced by clinicians, the very nature of complex environments makes studying interactions in these systems difficult as well. This is primarily due to a disassociation between the complex nature of the environment and the tools available to analyze cognitive and workflow processes.

The tools currently used for analyzing processes in these environments include qualitative methods such as ethnographic observation, shadowing of individual clinicians, surveys and questionnaires [22]. The data collected by these methods can be used to model segments of the clinical workflow centered on a particular individual and his or her activities [23]. Although the workflow documented in this manner captures many aspects of the overall system behavior, the presence of dense and interrelated interactions between various entities often makes operations in complex environments intractable. For example, observations are usually gathered from a single individual's point of view. A single observer may not be able to capture information on communication, movement and decision making, occurring at an instant of time. Theoretically, by increasing the number of observers it is possible to capture most of the information about the activities in the environment from several perspectives. However, based on informal interviews conducted with clinicians, more than two observers are considered disruptive to the clinical workflow. With such constraints imposed on data collection in complex environments, there is a need for an unobtrusive alternative that can augment existing methods of data collection, and help piece together a more complete description of system behaviors; both from individual and team perspectives.

Assessment of Behaviors in Complex Systems

In complex environments, adaptations ("deviations" from standards) and the resultant emergent behaviors provide insight into the processes that shape the system. In order to understand the root cause for errors in these systems, researchers would first need to examine the cognitive basis of these adaptive mechanisms. Protocols and guidelines have proven to be very useful in understanding complex tasks by dividing them into simpler observable units. Typically, protocols and guidelines suggest a sequence of atomic tasks and define a criterion for success. Checklists, a tool that has proven to be very effective in the management and control of processes in some complex environments (especially those structured by rigid protocols, as opposed to flexible guidelines) [24-26], are then utilized to assess clinician performance by examining the adherence to a protocol.

Much of the research assessing behaviors in complex systems follows this paradigm [27-29]. In these studies, deviations from protocols and guidelines are considered to be errors. The IOM, in fact, defines errors as *"…a deviation from that (protocol, procedure) which is generally held to be acceptable"* [30]. Although this definition of an error as a deviation is valid, the converse need not necessarily be true. In other words, while clearly all errors are deviations, not all deviations are errors. In fact, it is possible that a deviation from a protocol may be an innovation designed to maximize patient safety or an adaptation to enable the clinician to simply cope with the environment.

An example of complex social system that is similar to a clinical environment is aviation. Both pilots and clinicians operate in environments where teams interact with numerous technology and the risks originate from a number of sources in the environment. Errors, in these environments, occur due to a number of reasons; most of which are related to human error [31]. In contrast to medicine, however, errors in aviation often involve the loss of massive number of lives. A number of mechanisms have been adopted to minimize errors in aviation, focusing primarily on the task of error management in complex situations [32].

Crew resource management (CRM) [33], a major safety training in aviation, focuses on error training individuals in the countermeasures of human performance limiters (stress and fatigue). These counter measures include encouraging behaviors such as leadership, continuous monitoring, briefings, decision-making and dynamic modification of plans. In addition to CRM, simulation allows pilots to practice dealing with error management and receive feedback about they performance in dealing with complexity [31]. In addition to technical training, the domain of aviation has recognized the need to train both individuals and teams in dealing with complex error-prone situations, situations where plans may need to be altered dynamically to tailor the solution to the problem at hand.

An example of such an adaptive situation is the emergency landing of US Airways flight 1549 (on January 15, 2009) in the Hudson River is very well known. It involved a situation in which the airplane lost engine power shortly after takeoff. In this case, the flight captain used his own judgment and followed some protocols, while departing from others [34] and managed to land the heavy plane safely in the river. In emergency situations, the US Airways protocol calls for the first officer to take control of the flight, so that the captain can focus on making time-critical decisions. In this case, however, the captain quickly assessed the situation and deviated from the protocol. He took control of the plane instead and left his first officer to go through the checklist for restarting the engines. The decision was made because he felt that he was the more experienced pilot (and consequently had a better chances of landing the flight safely), while his first officer was more familiar with the specifics of the aircraft and would be able to go through the checklists more efficiently. The plane was in the river before the first officer completed the first page of the three-page checklist. This is an example where deviations from protocols (a dynamic alteration in action plan) resulted in a positive outcome.

A lesser-known example from aviation is that of Air France flight 447 that disappeared over the Atlantic on June 1, 2009. The analysis of the black box (published in December, 2011) revealed a disturbing finding [35]. The pilots encountered a storm and had to disengage from autopilot. This was not an unusual situation. The captain then left the helm to junior co-pilots for a routine break. Fifteen minutes later, the plane crashed killing the 228 people on board. The situation called for the junior pilots to coordinate their efforts in order to pass through the storm. However, the more inexperienced pilot of the two was overcome by the intensity of the situation and reverted to a protocol that was no longer applicable. By the time the captain returned to the cockpit, it was too late to prevent the crash.

"While (the first officer's) behavior is irrational, it is not inexplicable. Intense psychological stress tends to shut down the part of the brain responsible for innovative, creative thought. Instead, we tend to revert to the familiar and the well-rehearsed ...It's not surprising, then, that amid the frightening disorientation of the thunderstorm, (the first officer) reverted to flying the plane as if it had been

close to the ground (normal conditions), even though this response was totally ill-suited to the situation" [35].

This example highlights the fact that complexity, in some cases, cannot be controlled by protocols and standards. Individuals operating in such environments may be required to step outside the boundaries of "standard solutions" in order to solve time-critical problems. Based on safety mechanisms implemented in aviation it is evident that there is a need for research in medicine that examines the adaptive behavior of experts in order to improve the existing criteria for evaluation of performance in complex clinical environments.

Mixed Method Framework for Complex Systems

The research described in this manuscript examines the use of protocols and standards in a complex clinical environment with the end goal of understanding the cognitive mechanisms that initiate errors in these systems. The complex system under study is trauma care, a process occurring in a critical care environment. Trauma care is a highly dynamic process. Typically, teams of clinicians (with varying expertise, background and roles) treat a patient under critical conditions. The environment can become chaotic very quickly, due to the unpredictable nature of trauma cases and the unanticipated clustering of events. This makes trauma critical care a good representative environment for complex clinical systems.

Lapses in patient safety in trauma care and other complex environments have been linked to unexpected perturbations in clinical workflow [36, 37].

Effective workflow analysis is thus important to understanding the impact of these perturbations on patient outcome. The typical methods used for workflow analysis, such as ethnographic observations and interviewing, are limited in their ability to capture activities from different perspectives simultaneously. This limitation, coupled with the complexity and dynamic nature of clinical environments, makes understanding the nuances of clinical workflow difficult. In this work, a hybrid methodology is presented for capturing and analyzing workflow in complex environments.

Workflow analysis is an integral part of medical error research. A workflow is a description of a sequence of operations or activities performed by various entities or agents in the system [38]. It provides a description of the context and conditions in which errors occur. Careful analysis of workflow can be employed to model the distribution of cognitive work and the information flow in complex environments. For example, Malhotra et al. [23] utilized ethnographic observations and interview data to analyze the workflow in an intensive care unit. The workflow analysis helped the team to develop a cognitive model from which details of information flow could be extracted.

This type of analysis could lead to the discovery of latent systemic flaws that potentially result in adverse events. In addition, monitoring and assessment of workflow in complex clinical environments can provide clues regarding the efficacy of patient management. For these reasons, studying workflows in clinical environments is an important aspect of patient care and safety research.

Critical care environments, being complex adaptive systems, can become intractable when examined using qualitative methods of data collection. Observations, while rich in description, may not capture the cluster of events that occur simultaneously. There is a need for an unobtrusive alternative that can augment existing qualitative methods of data collection. This will help researchers to piece together a more complete workflow, both from individual and team perspectives.

In aviation (a complex social system similar to critical care), a key component of error analysis is the *black box*. The black box, as a tangible unit, refers to devices installed on aircrafts that track both communication within the cockpit of the aircraft, as well as performance parameters such as altitude, airspeed and heading. From a conceptual perspective, the black box is a continuous monitoring tool that does not interfere with the procedures of aviation and simply monitors parameters pertaining to the flight. If some tool akin to a black box were available for critical care units, analysis of adverse events would be far more accurate. The ability to automatically track all events that led to the adverse situation would be of great use in workflow modeling, error analysis and training of clinical professionals. In addition, an unobtrusive tool would enable monitoring of workflow without disrupting the activities of entities in the environment.

Methods used to analyze workflow in clinical environments can be one of two types – *qualitative* methods or *quantitative* methods. While qualitative methods involve subjective observations gathered by researchers, quantitative methods typically involve the usage of sensor technology or video recordings to capture data about workflow. The main differences between the data captured using quantitative methods and qualitative methods are as follows:

(i) Using quantitative methods, accurately time-stamped data can be obtained. Human-intensive methods can only produce time-stamped observations with near accuracy.

(ii) Qualitative methods of data collection produce relatively low volume, high quality data. On the other hand, quantitative methods produce a high volume of abstract data that in some way reflect underlying workflow.

(iii) Human-intensive qualitative methods are best suited for low-intensity situations, whereas automated quantitative methods are optimal for data gathering in high-intensity situations.

Qualitative Methods for Workflow Analysis

Malhotra and his colleagues analyzed the workflow in intensive care units (ICU) in order to understand the process of evolution of error in a critical care setting [23]. Ethnographic observations and interviews were utilized to gather data to model workflow centered on the entities and activities in the environment. The process of gathering observations involved following a key member of the critical care team and recording all of his or her interactions with both clinicians and equipment. These key players were then interviewed to corroborate the observations collected and to delineate their individual workflows. Using observations and interviews, a collective workflow was reconstructed by combining the individual workflows of each key player. The developed workflow summarizes how ICUs function and where errors are most likely to occur.

Laxmisan and her colleagues utilize ethnographic observations and interviews to analyze the workflow in an emergency department (ED) [39]. The workflow is analyzed to study the cognitive demands imposed by the workflow in the context of the work environment. Multi-tasking, interruptions, gaps in information flow and handovers during shift change were some of the aspects of the workflow that were studied in detail.

Quantitative Methods for Workflow Analysis

Quantitative methods provide means of gathering some information about the activities and whereabouts of entities in an environment. An entity could be a person (nurse, physician, patient, etc.) or a machine (such as ultrasound device). The tracked activities can then be pieced together (similar to integration of observations and interviews) to provide an aggregated overview of the workflow.

The sensors typically used for entity activity recognition include passive infrared sensors, radio frequency identification tags and pressure sensors. The sensors, depending on their type, are utilized to detect various activities in which the entity is involved. A number of systems have been developed for activity recognition and workflow monitoring using different types of sensors. These systems use the various types of sensors in some combination to model key activities of the entity being tracked. In general, these sensors are encased into a physical form representing a tag. These tags can sense different types of information like movement and location through the ensemble of sensors embedded in the physical form.

In the domain of healthcare, tags have been employed for tracking patients, equipment and staff to gather data that can be used to improve patient care and the efficiency of clinical workflow [40-46]. Fry and Lenert [46] developed a system for location tracking of patients, staff and equipment called MASCAL. The main aim of the system was to aid in streamlining patient care during mass casualty situations. RFID tags were used by the system to track the location of key players (clinicians and equipment) in patient care during emergencies. This information is integrated with personnel databases, medical information systems and other applications (such as those that enable registration and triage) in order to centralize the management of resources during critical situations. In addition, MASCAL included interfaces for centralized management of various entities in the system.

Chen et al. [45] studied the incorporation of RFID technology in a clinical setting in non-psychiatric hospitals in Taipei, Taiwan. Tags were used to identify patients and notify clinicians on the status of patients and patient related information (lab reports, radiology results etc.). Preliminary studies showed that using the RFID-enabled framework decreased the wait time for patients in intensive care units.

The other technique for activity monitoring is processing of video recordings. Hauptmann et al. [47] describe a system that recognized activities from videos captured using video processing techniques. The system was developed to recognize activities of daily living (ADL) for patients. Examples of ADL activities include visiting the washroom, eating, sleeping etc. Cameras placed at key locations within the environment provided video feeds. These video feeds were processed to identify the patients and draw conclusions on the possible activities in which the patients were involved.

Limitations of Qualitative Methods

Qualitative methods are human-intensive, i.e., they require significant amounts of human effort for data gathering and analysis. The dependency of qualitative methods on human effort has certain advantages and disadvantages. The main advantage is that human-intensive methods usually yield data that are of high quality. These data are detailed and descriptive, and potentially insightful inferences can be made using qualitative analyses of these descriptions. The disadvantage, however, is that the dependence on people for data gathering and analysis limits the capabilities of these methods to capture important details of the collective workflow in a critical care environment.

Observation gathering is a classical qualitative method for workflow analysis that suffers from its dependence on human effort. It is difficult for individuals to monitor and document all activities that occur at every instant in a dynamic and complex environment. Interviews, on the other hand, suffer from the poor recall of events on the part of clinicians being interviewed. Facts about events may be altered as the memory of the event changes temporally (post-hoc bias). Other real-time methods of data collection such as audio and video recording systems not only require consent from clinicians to be used to gather workflow data, but also require significant human efforts for processing data collected to retrieve meaningful information. Post-processing of real-time data involves manual analysis of audio and video data in order to detect various workflow events. The real-time data can then be manually annotated with the key events that have been detected. This process requires time, effort and researcher expertise in order to be completed successfully. Such limitations make these methods more suited for workflow analysis in simple, low-activity environments.

Limitations of Quantitative Methods

In most quantitative methods, sensors for monitoring activities and locations are placed at pre-defined locations. The rigid infrastructure often makes installation costs prohibitive. In addition, maintenance can be complicated if spatial configurations are altered. Another issue lies with the modeling approaches employed to track workflow. In all the current systems, the sensor systems are employed to determine the location of the entities from which activities are estimated. This system works well if the location identification is reasonably accurate. However, RFID systems can often be highly erroneous, resulting in close to 200% errors in location estimates [48]. Location in these systems is determined by geometric triangulation methods that have limited performance in environments with electromagnetic fields. Since clinical environments require large amounts of equipment, it is impossible to control for electromagnetic fields. To account for this high rate of error, activities that are covered by the current approaches limited to macro-level movement-based activities such as entering a room or going from one area of the hospital to another. Current systems are limited in documenting activities that occur in smaller area, because the sensors cannot discriminate location in these environments with acceptable accuracy.

Video-based tracking suffers from similar issues. The locations of cameras are fixed. Areas need to be analyzed to ensure that the cameras cover all parts that need to be monitored. In addition, real-time analysis of videos for entity recognition can suffer from typical video processing problems, such as occlusion of entities by other entities, noise, motion blur, uneven lighting and so on. This, coupled with the requirements of privacy and security, often render video-based capture unusable.

Proposed Framework for Assessment of Behaviors in Complex Systems

As both qualitative and quantitative methods have advantages and disadvantages, an improved solution for workflow monitoring can be obtained by combining the two types of methods.

Figure 2 depicts the framework of a mixed methods system for monitoring activities and behaviors in complex critical care environments. Such a framework supplements existing methods of data collection (qualitative observations) with quantitative methods (use of electronic tags and audio recording) to capture activities in the complex system. In addition to developing a quantitative activity models, a theoretical model of deviations (adaptations) is developed to support this mixed methodology framework. The qualitative analysis combined with quantitative metrics obtained from tags and audio recording can provide deeper insight into adaptations from which the emergent behavior of the complex system can be inferred.

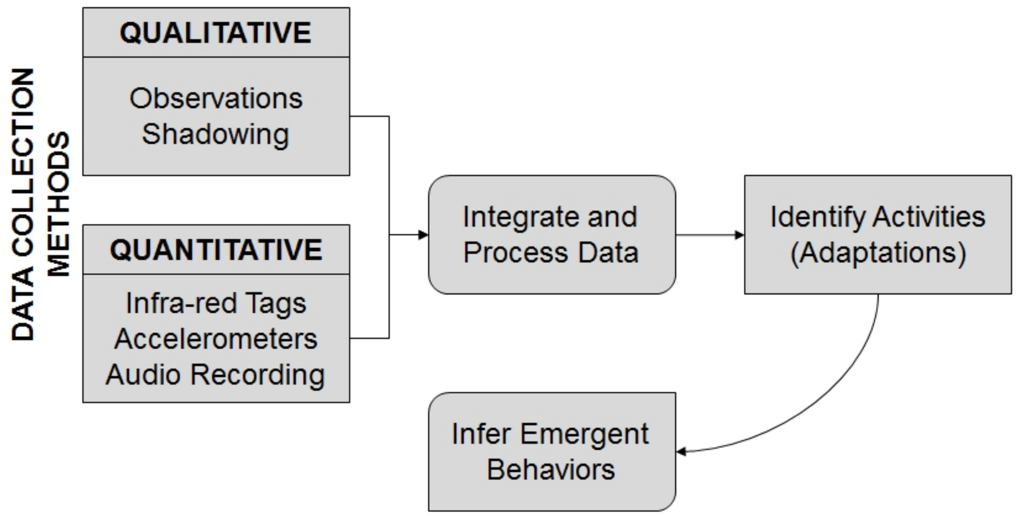

Figure 2. Conceptual framework of system for monitoring critical care

The work described in the manuscript can be broadly divided into three segments. First, an introduction to trauma critical care environments is provided. In this segment the guidelines and standards followed in trauma critical care is discussed, placing it within the context of the clinicians and teams interacting in trauma care. Following a description of the environment, the methods for developing quantitative models of activities in trauma care is discussed. This section presents the work dealing with the use of radio frequency identification tags for automatically detecting activities in trauma care. Unlike infrared tags, RFID tags do not require a line of sight with other tags to record information. Hence, RFID tags are utilized to gather the quantitative data. Observations gathered complement RFID data by providing a detailed description of communication and interaction activities that cannot be captured using the tags.

Finally the last segment deals with the research on the development of qualitative models of deviations from standards in trauma care. In order to assess the deviations from standards (as detected by quantitative models of activity), it is important to understand the various types of deviations than can take place in a complex system. In addition to describing preliminary (explorative) research in the domain, the extended study of deviations and experiments conducted to assess inter-rater reliability in classifying the deviations is elucidated in this section.

This research described in this work is one of the first to examine the cognitive basis of adaptive mechanisms of clinicians in medicine and present a methodology for studying the same. Through a deeper understanding of the cognitive decision-making processes that allow experts and teams to manage errors, healthcare systems designed around the principles of safety can be developed.

TRAUMA CRITICAL CARE

In critical care settings, teams of care professionals care for patients. These teams typically involve clinicians with varying backgrounds and expertise, working in a collaborative manner. A patient may interact with as many as fifty different employees (including nurses, physicians and technicians), during a typical 4-day stay at a hospital [49]. These teams operate in environments with dynamic social structures [39] and are required to adapt to varying task demands and coordinate their efforts to carry out activities necessary for task completion [50]. Team decision-making is a key factor that impacts co-ordination among individuals involved in the patient care process.

In trauma critical care, clinicians follow the Advanced Trauma Life Support (ATLS) guideline [51], developed by the American College of Surgeons (ACS). It is mandatory that this protocol be followed in every Level 1 trauma center for accreditation purposes. Research has shown that the ATLS protocol is effective in improving the quality of care in trauma centers across the United States [52]. The tasks and goals for "Initial Survey and Management" of the patient are common to both physicians and nurses (summarized in Table 1). The guideline can be divided into three sections: (i) primary survey and resuscitation, (ii) secondary survey and examination, and (iii) definitive care and transfer. In the primary survey, all immediate, life-threatening conditions are mitigated. Once the patient's vital signs stabilize, a thorough head-to-toe examination can be performed. Information obtained from examinations (and diagnostic tests) allows the trauma team leader to make decisions relating to the care of the patient.

Table 1. Key steps in Initial Assessment and Management ATLS protocol

(A) Primary Survey Assessment of ABCDE's	1. Airway with cervical spine protection 2. Breathing 3. Circulation with control for external hemorrhage 4. Disability with brief neurological evaluation 5. Exposure/Environment
(B) Resuscitation	1. Oxygenation and ventilation 2. Shock management and delivery of fluids 3. Management of life-threatening problems
(C) Adjuncts to Primary Survey and Resuscitation	1. Monitoring a. Arterial blood gas analysis and ventilator rate b. End-tidal carbon dioxide c. Electrocardiograph d. Pulse oximetry e. Blood pressure 2. Urinary and gastric catheters 3. X-rays and diagnostic studies a. Chest b. Pelvis c. C-Spine d. Diagnostic peritoneal lavage or abdominal ultrasonography
(D) Secondary Survey, Total Patient Evaluation: Physical Examination and history	1. Head and skull 2. Maxillofacial 3. Neck 4. Chest 5. Abdomen 6. Perineum/Rectum/Vagina 7. Musculoskeletal 8. Complete neurologic examination 9. Tube and fingers in every orifice
(E) Adjuncts to the Secondary Survey	1. Computerized Tomography 2. Contrast X-ray studies 3. Extremity X-rays 4. Endoscopy and ultrasonography
(F) Definitive Care	Based on the diagnosis, patient treated in trauma care (if applicable)
(G) Transfer	Based on the type of care needed, patient may be transferred (to the operating room or intensive care unit) or be discharged from the facility

In addition to providing a systematic way to treat patients, the ATLS guideline serves to establish a common vocabulary for multi-disciplinary trauma teams to function effectively. The guideline will now be described in detail, placing it within the context of the environment, tasks and goals.

Trauma Team Structure

Trauma teams aid in rapid identification and treatment of life-threatening conditions. They are responsible for: (i) assessment of the patient upon arrival, (ii) resuscitation and management of critical conditions, and (iii) diagnosis and transfer of the patient to the appropriate facility. The core team typically includes the attending surgeon, residents, an anesthesiologist, and nurses. Supporting members include a respiratory therapist, pharmacist and an X-ray technician.

Roles and responsibilities are well defined for team members. The trauma *team leader* supervises the trauma care, making major decisions and delegating work to other members of the trauma team. A resident physician may assist the trauma lead. The *assisting physician* performs hands-on evaluation and treatment. The *primary trauma nurse* is responsible for the immediate care of the patient. A nurse recorder who documents events in trauma workflow sheets may assist the primary nurse. The structure of the core team is often dynamic. Roles of the team leader and assisting physician may shift between residents and the attending trauma surgeon. In teaching hospitals, attending surgeons mostly play the role of a guide overseeing residents serving as the trauma leader.

Trauma Information Sources

In trauma critical care, the information available to the team evolves as new observations are made, tests are analyzed and consults are obtained. Trauma teams receive information from a variety of sources including pre-arrival patient information, trauma workflow sheets, the patient vital signs monitor, x-ray images, computerized tomography (CT) scans, diagnostic tools to analyze blood and urine samples, and information shared by other care providers [53].

Although team members follow the same guideline for treating the patient, the boundaries of an individual's role (within the team) impact the types of information processed and utilized by each team member. For example, x-rays and CT scans are always assessed by the trauma leader, which forms the basis for decisions about treatment and definitive care of the patient.

In such conditions, one of the main challenges faced by teams is decision making with evolving information. Often the complete medical history of the patient may not be available when critical decisions have to be made. Trauma teams may be required to adapt they decision making as more information emerges.

Trauma Scenario Walkthrough

Irrespective of type of trauma, certain key steps are performed (*in quasi-sequential order*), to evaluate the patient. In this section, a walkthrough of a typical trauma case scenario encountered at a Level-1 trauma facility is presented (workflow depicted in Figure 3). The workflow in trauma care can broadly be divided into (i) primary survey and resuscitation, (ii) secondary survey, and, (iii) tertiary survey and definitive care. In the following sub-sections, the activities performed by the team in these three categories are described. Note that this scenario is based on workflow observed on site at the Level-1 trauma center as well as on the existing literature on ATLS guideline implementation [51].

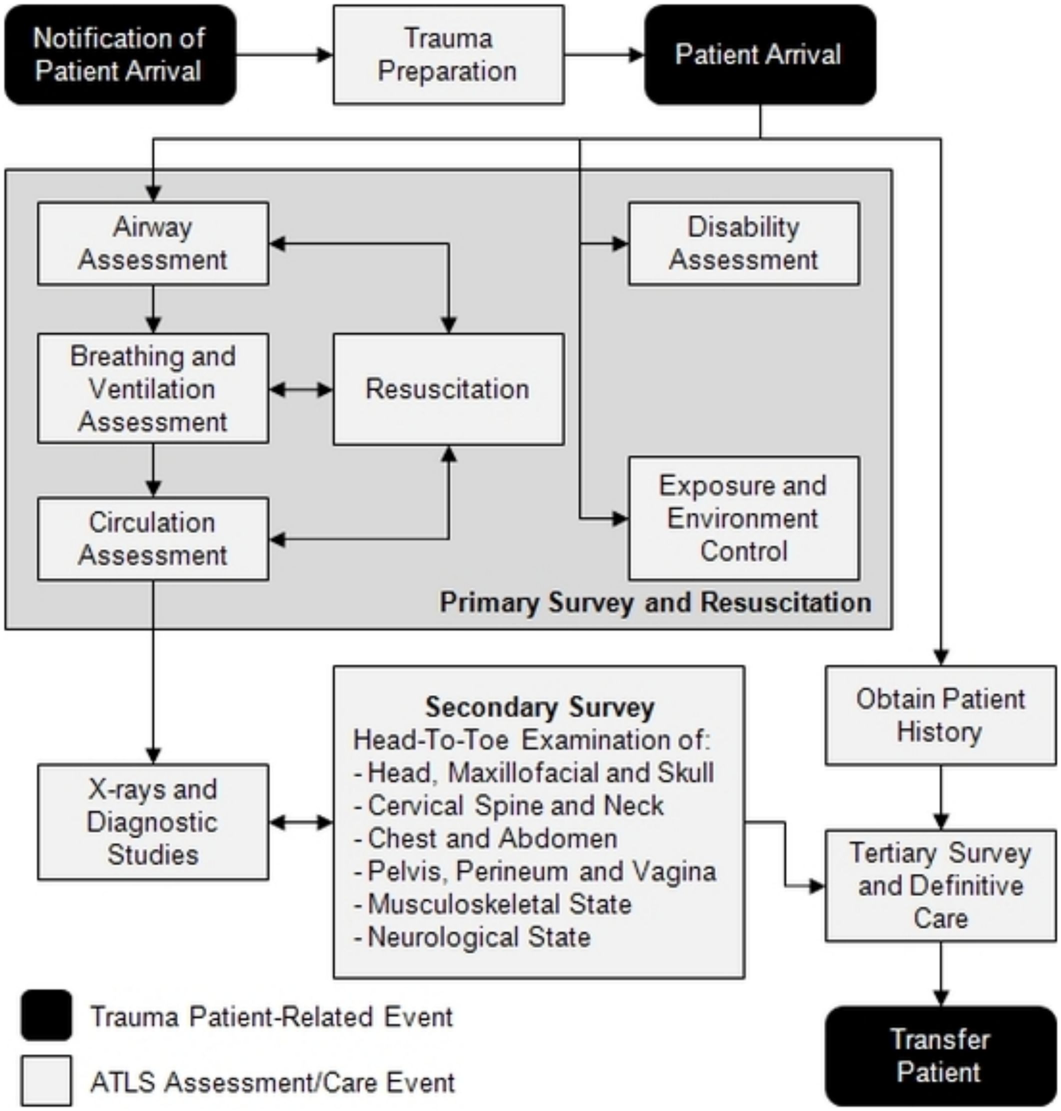

Figure 3. Trauma scenario walkthrough: typical workflow observed in trauma care

Trauma Care Preparation

A trauma care scenario typically begins with an announcement of trauma arrival with an acuity or case type indicator. This indicator is usually specific to the trauma care site. With respect to a representative venue, Banner Good Samaritan

Hospital (Phoenix, AZ), trauma cases that may require an anesthesiologist are classified as “trauma A”. Cases with lower severity are classified as “trauma B”. There may be other classifiers that are independent of severity. For example, any case involving a pregnant woman is classified as “trauma C”. Based on the trauma severity or type indicator, care provider teams assemble in the trauma unit. In the case of trauma C, two trauma teams assemble; one for the mother and the other for the child. As simple as this triaging scheme may be, it allows for resources to be managed effectively within the hospital.

Once the required team members assemble for the trauma care, the clinicians may have a brief window (often ranging from 2 to 10 minutes), in which they can perform activities to prepare for the case at hand. For example, clinicians may exchange information about the incoming case, or scrub and wear appropriate protective garments. When the patient arrives, emergency medical technicians transfer the patient to the trauma bay and provide a brief overview of patient history and treatment provided. At this point, the trauma leader takes charge of the trauma care and initiates the primary survey.

Primary Survey and Resuscitation

In the primary survey, the trauma leader evaluates the patient airway, breathing, circulation and neurological state (disability via Glasgow Coma Scale or Injury Severity Score metric [54]). This survey is usually quick and performed within the first two minutes of patient arrival. Resuscitative efforts (orders given by the leader) and patient exposure (removal of clothing) are typically performed in parallel by other team members (primary nurse and assisting physician). When all life-threatening conditions have been addressed, the team proceeds to utilize diagnostic tests (x-ray, CT scan, blood and urine sample testing) as needed to further diagnose the patient trauma and follow appropriate treatment.

Secondary Survey and Definitive Care

The secondary survey may be performed while awaiting the results of diagnostic tests and involves detailed head-to-toe examination of the patient. Once the patient is thoroughly examined and diagnostic test information is available, the trauma leader proceeds with formulating a treatment plan. At this stage, he/she may consult with the mentor (attending surgeon) or a specific specialty consult (for example, orthopedic or plastics consult). The team may then proceed with providing definitive care (management of conditions not treated at the end of the primary survey) and conducting tertiary surveys, if required. When the patient is ready to be transferred out of the trauma unit, the patient may be discharged or moved to a room for monitoring and extended treatment by a consult.

Protocols and Guidelines

The ATLS standard described (and tabulated in Table 1) is a *guideline* as opposed to being a fixed protocol. A guideline is defined as *"a statement or other indication of policy or procedure by which to determine a course of action"* [55]. In contrast, a protocol is *"a precise and detailed plan … for a regimen of diagnosis or therapy"* [56]. Since trauma care is a complex system that is inherently dynamic and unpredictable, providing clinicians with a rigid protocol would limit their ability to adapt to the situations at hand. A guideline, on the other hand, does not inherently penalize a clinician for not performing a particular step in order. This allows clinicians to adapt the guideline to suit the dynamic needs and requirements of the team.

For the purpose of this research, the ATLS guideline is considered to be a set of minimum specifications. The guideline provides general direction for the team and describes role boundaries, resources and constraints [57, 58]. The implementation of such a guideline, as opposed to detailed protocols, can result in the emergence of innovative and complex behaviors [59]. The key challenge here is to ensure that the deviations or novel adaptations made by the team members do not contradict the purpose of the guideline and consequently compromise patient safety. The following chapter discusses the proposed framework for studying complex critical care environments.

QUANTITATIVE ACTIVITY MODELING USING RFID TAGS

This chapter presents quantitative models of activity developed using Radio Frequency Identification (RFID) tags. The activities detected, combined with observations captured can provide a more complete picture of the workflow. RFID tags provide a means to automatic identify an entity, in addition to continuous monitoring and location sensing. Basic information about interaction among the entities, such as duration, proximity and location can be extracted from the tag data. These data, combined with qualitative measures allows one to construct an intermediate workflow that can be visualized in virtual reality environments. The end result is a system that augments existing methods of data collection to capture a comprehensive view of workflow in complex environments.

In general, workflow can be described by (a) the underlying cognitive processes that drive decision making, (b) gross physical movement, and (c) interaction and communication activities. The mixed methodology framework for workflow analysis system combines qualitative and quantitative methods of data collection to capture each of these three activities. RFID tags can provide quantitative information about movement activities, in addition to some basic interaction statistics such as proximity between two or more clinicians and time spent at particular locations. While, these statistics could be utilized to model the movement patterns of clinicians in the environment, RFID tags cannot gather information about specific details of communication or decision making of clinicians that result in a particular situation. Researchers rely on qualitative methods of observations for capturing this kind of information. Figure 4 depicts the types of activities that can be captured using the hybrid framework.

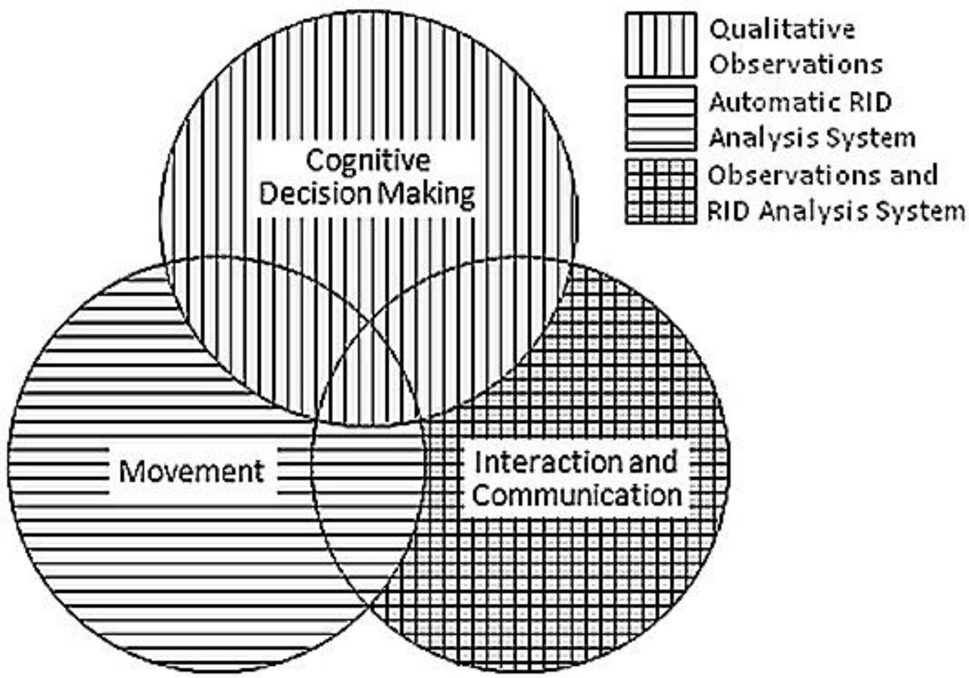

Figure 4. Overview of activities captured and tools utilized

Traditionally qualitative data collection involves observers following both cognitive and movement activities. This is in addition to collecting detailed information about the time of activity initiation and the sequence of activities in a workflow. Use of quantitative methods offers the means to offload the task of recognizing movement-based activities to the automatic algorithms that process incoming tag data. This system can theoretically capture any movement activities that require team members to move at least 8 inches [48].

With respect to the mixed method framework, two different streams of data are collected; (a) qualitative data from observers, and (b) quantitative data, gathered from the RFID tags. Both the qualitative data and quantitative data are obtained from standardized sources. While time-stamped quantitative data is retrieved from the RFID tags, observations are gathered by observers shadowing clinicians in the environment. Observations logged on a laptop are automatically dated and timed and stored in the output observation file. The saved time stamp is then used to synchronize the qualitative and quantitative data sources.

SNiF Radio Frequency Identification Tag System

Quantitative data is obtained using *active* RFID tags (depicted in Figure 5). Active RFID tags have an inbuilt power source, hence the name “active”. In addition to being portable, active tags use low levels of energy, ensuring that they do not interfere with other devices, such as telephones and other network connections found in a healthcare setting.

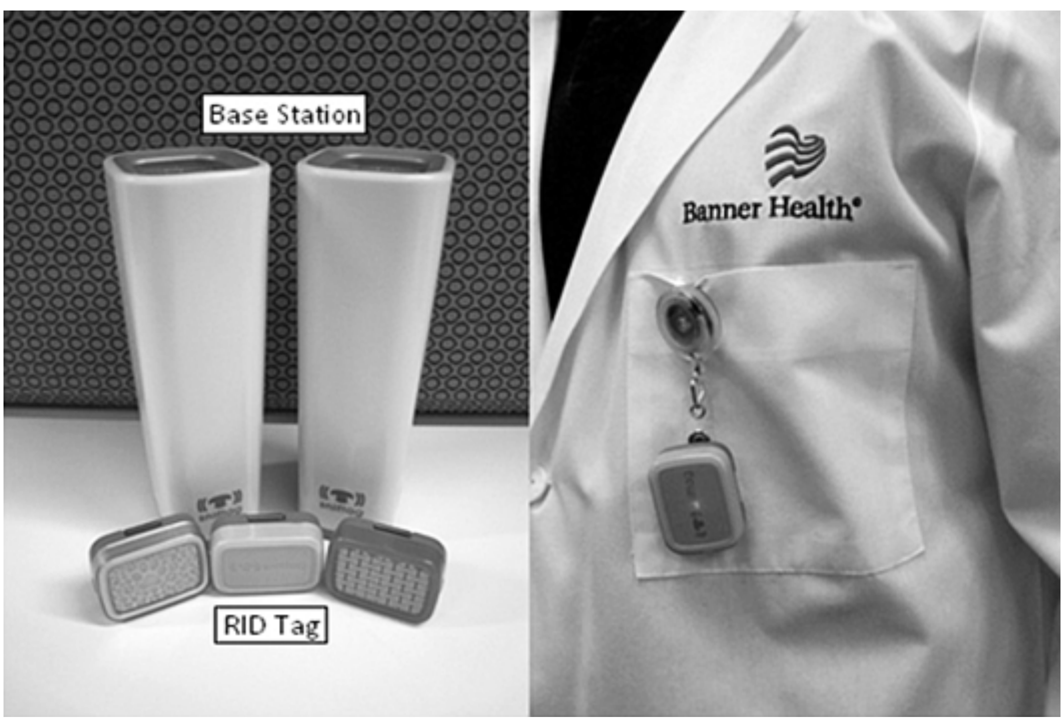

Figure 5. Active radio frequency identification (RFID) tag and base station

SNiF® (Social Networking in Fur) is an off-the-shelf active RFID tag. The SNiF tag system comes with portable tags and passive base stations. While the base-stations operate similar to the tags, they primarily serve as location beacons. They are placed at critical areas in the environment being studied. It was found that in a trauma unit, the trauma bays, nurses' station and entry and exit points are some of the key locations. It should be noted that these locations would vary from site to site, depending on variations in the workflow.

The tags record encounters with other tags (tag-tag encounter) and base stations (tag-base encounter). For each encounter or interaction, the tags record:

(i) identification number of the tag or base station detected,

(ii) time and date of encounter, and

(iii) received signal strength indication (RSSI) value.

The RSSI value provides proximity information. This value is inversely related to the distance between the interacting tags. Consequently, it can be used to measure approximate distances between tags and base stations involved in an interaction. Encounters are recorded at a rate of 0.33 pings/second. Temporal analysis of pings can provide information of duration of interaction, in addition to the proximity information. Using information about proximity and duration of interaction, it is possible to infer activities, such as a resident leaving trauma care, or a nurse documenting the case at the nurses' station. Utilizing proximity information from tag-tag and tag-base encounters, an abstraction of movement-based activities can be obtained.

Use of Proximity as a Proxy for Interaction and Communication

In addition to movements, some information about communication and interaction can be obtained from tags as well. Consider the scenario representing "patient arrival". When a patient arrives in a trauma room, trauma team members tend to converge at a trauma bay. Following this, an examination of the patient takes place. Eventually, a resident may move to the telephone for a consult or the nurse may move to the nurses' station to document details of the encounter. All these activities are linked to entities performing some type of movement in the environment and can be inferred from the tag data. Formally this sequence of activities can be expressed in terms of time as:

(i) At time t_1: Patient arrives at the trauma unit and is sent to the trauma bay.

(ii) At time t_2: The nurse and a resident check in on the patient.

(iii) At time t_3: The resident seeks a phone consult while the nurse heads over to the station to continue with documentation.

Assume the trauma room is equipped with base stations and tags, as depicted in Figure 6. In this diagram, 'P' refers to the patient; 'N' refers to the nurse and 'R' to the resident on call. The black solid dots denote the locations of base stations (B1 to B6). Base stations are placed at various key locations; one at each trauma bay, one near the phone and the other near the computer. Assuming that the trauma team members are carrying portable tags, the following are the trends seen in the extracted data:

(i) At time t_2: Tags R and N get close to B1.

(ii) At time t_3: Tag N is very close to B5 and Tag R is very close to B6.

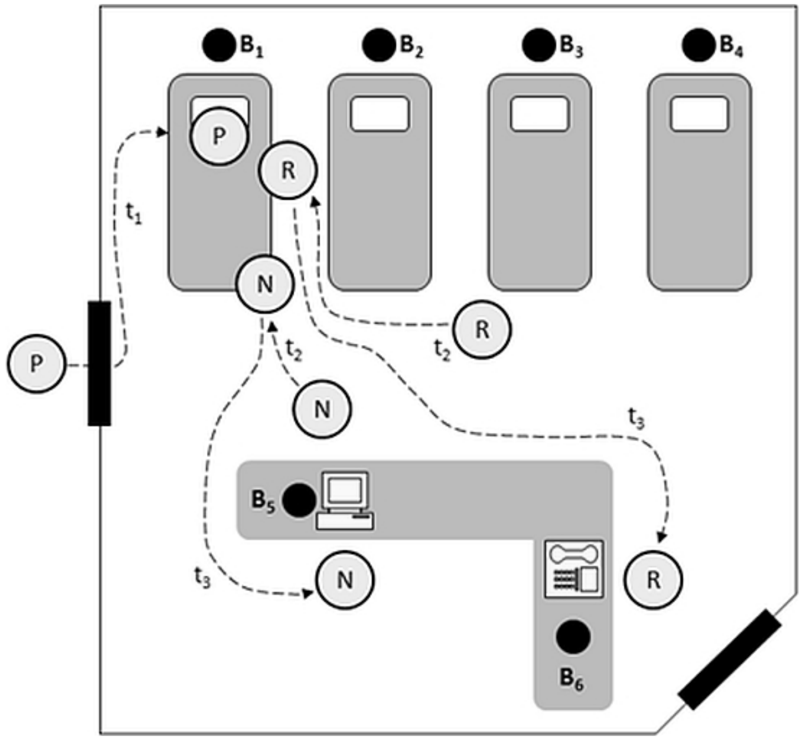

Figure 6. Scenario: Patient arrival at a trauma unit

When tags carried by the trauma team converge at a bay, it can be assumed that the patient is being examined (t_2) and that the patient arrived at the unit sometime before t_2. If the patient carries a tag as well, at time t_1 the data will show tag 'P' gets close to base station B1. At time t_3, the system can probabilistically estimate that the nurse was documenting the patient report, and that the resident was seeking a phone consult.

Although the scenario presented is a simplification of the total process, it provides a conceptual view of how interaction activities can be tracked using proximity information. As a general rule, any interaction and communication activity that is accompanied with measurable movement can be captured by this system and recognized. Following the same logic, any communication or interaction activity that is not accompanied by movement cannot be captured by the automatic analysis system. Additional sensors can give more detailed information on some of the activities. Incorporation of audio recording (while ensuring privacy of the subject is maintained) would facilitate automated tracking of communication between entities.

Activity Recognition using Hidden Markov Models

In this work, Hidden Markov Modeling is used to analyze the temporal data gathered by the tags and recognize known activities. It is a probabilistic modeling method used for temporal sequence analysis and model generation, and has been widely used in gesture and speech recognition [60, 61].

Hidden Markov Models

A Hidden Markov Model (HMM) is a finite set of states, each of which is associated with multidimensional probability distributions. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. There are three fundamental variables that must be determined to generate models of activity;

(i) Initial state probability, π – This is a set of probabilities π_i, which indicates the probability of the starting or initial state being i. π can be represented by a $Nx1$ matrix where N is the number of states.

(ii) Transition probability, A – A set of probabilities A_{ij} where a_{ij} indicates the probability of the operator transitioning from state i to state j. Hence, A is represented by a NxN matrix.

(iii) Bias probability, B - a set of probabilities $B_i(k)$ where $b_i(k)$ is the probability that symbol k is observed at state i. Hence, B is represented by an *NxM matrix*, where N is the number of states and M is the number of observation symbols.

The HMM is then represented as $\lambda = (\pi, A, B)$, where the observed sequence is modeled as a state machine, wherein the current state is dependent only on the previous state. Using HMMs requires solutions to the following basic problems;

(i) Given a model $\lambda = (A, B, \pi)$, what is the probability that a given observed sequence O belongs to λ, i.e., $P(O| \lambda)$?

(ii) Given, $\lambda = (A, B, \pi)$, what is the sequence of states $I = \{i_1, i_2, i_3, i_4 \ldots i_T\}$ (T is the number of observed symbols) such that $P(O, I| \lambda)$ is maximized?

(iii) How can the HMM parameters π, A and B be adjusted so as to maximize $P(O, I| \lambda)$? This is also known as a training problem or training an HMM.

The current problem at hand is activity recognition using HMMs. The observed sequence, in this case, is temporal data about encounters obtained from the tags. In order to develop robust activity HMMs, data that describe controlled samples of activity are obtained from the RFID tags. Multiple samples are captured for each activity of interest. A database of samples for each activity facilitates training the HMMs for each activity, thereby creating a library of HMM activity models for each activity. The following are the steps to train and test HMMs:

(i) Obtain data from tags for specific (marked) activities or motions. This is obtained from qualitative data collected (observations and interviews) in addition to tag data.

(ii) Use marked data to set the parameters of the HMM, i.e., train the model

(iii) Test the HMM, by evaluating if test data are appropriately recognized

Algorithm for Testing HMM (Forward-Backward Method)

This method defines a variable $\alpha_t(i)$ called the *forward variable* as follows:

$$\alpha_t(i) = P(O_1,O_2,O_3, \ldots O_t, i_t = i \mid \lambda)$$

This is the probability of the partial observation sequence up to the position *t,* At state *i* at position $\alpha_t(i)$ is given by,

(i) $\alpha_1(i) = \pi_i b_i(O_1), 1 \geq i \geq N$

(ii) For $t = 1,2,\ldots T\text{-}1$, $1 \geq j \geq N$,

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] b_j(O_{t+1})$$

(iii) Then,

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Step 1 refers to the probability for picking state *i* and generating O_1. The probabilities then generated by step 2 represent transitioning from a state at t to a state at t+1 and generating O_{t+1}. Inductively $P(O|\lambda)$ is found. A backward variable $\beta_t(i)$ is defined as:

$$P(O_{t+1}, O_{t+2}, \ldots O_T | i_t, \lambda)$$

This is the probability that a sequence from *t+1* to *T* is observed, given the state *i* at time *t* and λ. $\beta_T(i)$ is given by,

(i) $\beta_T(i) = 1, 1 \geq i \geq N$

(ii) For *t = T-1, T-2, …1,* $1 \geq i \geq N$*,*

$$\beta_t(i) = \left[\sum_{j=1}^{N} a_{i,j} b_j(O_{t+1}) \beta_{t+1}(j)\right]$$

(iii) Then

$$P(O|\lambda) = \sum_{i=1}^{N} \pi_i b_i(O_1) \beta_1(i)$$

Both the forward and backward procedure can solve for $P(O|\lambda)$ in N^2T time. Practically a test sequence is divided into two parts by breaking it in the middle. The first part is solved using the forward variable and the second part is solved using the backward variable. These probabilities are then combined to find the probability of a test sequence being close to the given HMM. Since a library of HMMs is available, it is possible to find the probability of the test sequence being close to each of the HMMs in the library. The HMM that generates the highest probability for a test sequence is the winning HMM for the given test sequence.

Algorithm for Training HMM (Baum-Welch)

This method is used in the training phase to find the HMM for a particular activity. All the tag data pertaining to a single activity are used to train a HMM for that activity. The function $P(O|\lambda)$ is called the likelihood function. Assume:

$$\gamma_t(i) = P(i_t = i|O, \lambda)$$

This is the probability of being in state *i* at time *t,* given sequence $O = O_1, O_2, \ldots, O_T$ and λ. From Bayes theorem,

$$\gamma_t(i) = \frac{P(i_t = i, O|\lambda)}{P(O|\lambda)} = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)}$$

Where $\alpha_t(i)$ and $\beta_t(i)$ are the forward and backward variables defined previously. A variable $\xi_t(i,j)$ is defined as:

$$\xi_t(i,j) = P(i_t = i, i_{t+1} = j|O, \lambda)$$

From the derivations, the following is obtained:

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\lambda)}$$

It can be seen that summing up $\gamma_t(i)$ from *t=1* to *T* provides the number of times state *i* is visited, or summing up only up to *T-1* provides the number of transitions out of state *i.* Similarly, summing $\xi_t(i,j)$ from t=1 to T-1, the number of transitions from state *i* to state *j* is obtained. Therefore,

$\sum_{t=1}^{T-1} \gamma_t(i)$ = Expected no. of transitions from *i*

$\sum_{t=1}^{T-1} \xi_t(i,j)$ = Expected no. of transitions from *i* to j

The re-estimation formulae are as follows:

$$\pi_i = \gamma_t(i), 1 \geq i \geq N$$

$$a_{i,j} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$b_j(k) = \frac{\sum_{t=1, O_t=k}^{T} \gamma_t(i)}{\sum_{t=1}^{T} \gamma_t(j)}$$

These are the updated parameters for the new HMM. Therefore the algorithm proceeds as follows. Obtain the initial HMM. Calculate *A, B* and π. Estimate $P(O|\lambda)$ until reaching a sequence length t. Re-estimate the model and the likelihood function. These steps are done repeatedly until the likelihood function is maximized.

Once a library of HMMs is built with one HMM for each activity, using the algorithms described, models can be developed and tested. As with any method, HMM-based activity recognition has certain advantages and disadvantages. The key disadvantage of HMMs lies in the fact that the amount of data that is required to train an HMM is very large. Another issue with HMMs is that they require positive data to train with, i.e., in order to effectively train an HMM to recognize a class of activities, researchers require a carefully constructed training set that best describes the activity. However, these disadvantages are outweighed by the capability of a trained HMM to handle variations in the style of execution of an activity. Activities can be performed in a different manner in critical care environments, and it is important that the model of activities accounts for these variations. By training the HMM system in this manner, it is possible to recognize the motion and some communication activities regardless of the deviations. In addition, HMMs scale well, since they can be trained to learn activities incrementally. New activities can be trained for without affecting models of previously learned activities. For these reasons, HMM was chosen for the development of activity models and activity recognition.

System Evaluation

The experiment conducted evaluates that the accuracy of the HMM-based activity recognition system in recognizing clinical activities involving movement patterns in a lab setting (setup depicted in Figure 7). Accuracy, in this work, is measured as the ratio of the number of correctly identified test sequences to the total number of test sequences.

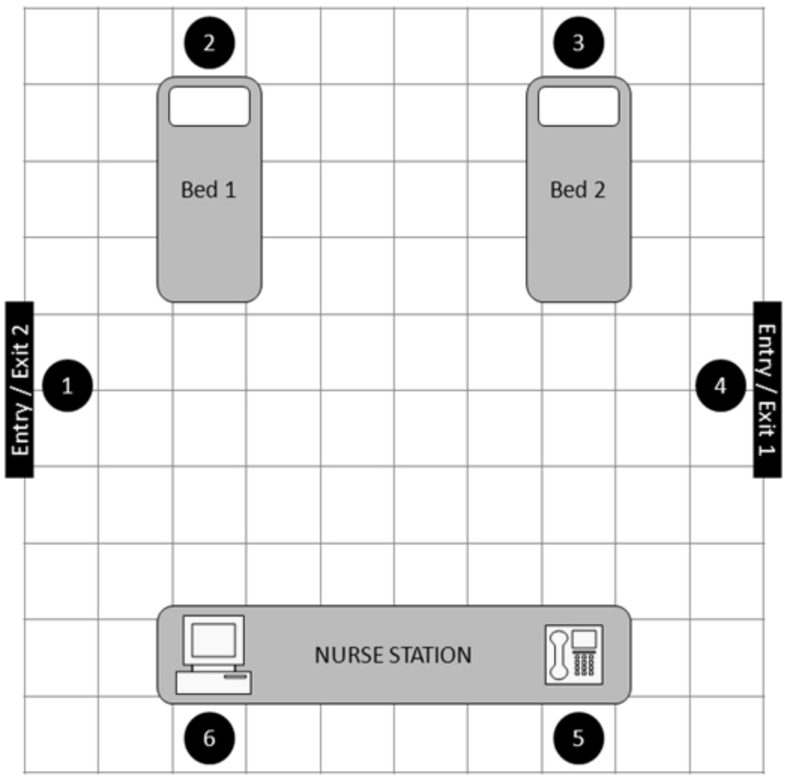

Figure 7. Test setup for simulated clinical activities

Evaluation Setup

In order to test the HMM-based activity recognition system, commonly occurring movement-based tasks in the trauma unit were identified; an example being *"physician moving to phone for a consult"*. These activities were then simulated in a lab setting.

The setup for the testing involved the creation of a 20 ft. by 20 ft. grid in a lab setting (setup depicted in Figure 7). Six base stations (depicted by black solid circles) are placed in predefined locations (Base 1 and 4 at Entry/Exit points 2 and 1 respectively; Bases 2 and 3 at Beds 1 and 2; Base 5 at the phone on nurse station; Base 6 at the computer on the nurse station). This is congruous with base station setup in the real-world scenario.

Data Collection

A total of 15 trauma activities (listed in Table 2) were simulated in a lab setting, with 10 tags and 6 base stations. Multiple samples of each activity were captured using the RFID tags. Each sample involved a tagged entity (researcher) following the movement pattern prescribed for the activity and was performed with 10 times with 10 different tags, totaling 100 samples for each activity. This ensured sufficient randomization of activity movements, accounting for inter-tag variability as well. Out of the 100 samples gathered for each activity, 50 samples were used to train the HMM for activity recognition, and the other 50 were used as a testing set to evaluate the algorithms' accuracy. A total of 1500 movement samples (15 activities x 10 samples x 10 tags) were gathered for this experiment.

Table 2. Activity list and corresponding clinical descriptions

Activity	Movement	Clinical Description
A1	1-to-2	Paged physician/nurse tends to patient on bed 1
A2	2-to-3	Physician/Nurse moves to treat patient on bed 2
A3	3-to-4	Physician/Nurse leaves Trauma through entry/exit 1 after visiting patient on bed 2
A4	4-to-5	Physician/Nurse enters Trauma through entry/exit 1 and attends to the phone
A5	5-to-6	Physician/Nurse after attending to a phone call move to use the computer at the nurse station
A6	6-to-1	Physician/Nurse leaves Trauma through entry/exit 2
A7	1-to-4	Physician/Nurse enter and leave Trauma
A8	4-to-6	Physician/Nurse enter Trauma through entry/exit 1 and move to use the computer at the nurse station
A9	6-to-2	After using the computer physician/nurse move to treat patient on bed 1
A10	2-to-4	After visiting patient on bed 1, physician/nurse leaves Trauma through entry/exit 1
A11	5-to-1	After attending a phone call, physician/nurse leaves Trauma through entry/exit 2
A12	1-to-3	Paged physician/nurse attends to patient on bed 2
A13	3-to-5	After visiting patient on bed 2 physician seeks a phone consult
A14	5-to-2	After completing a phone call physician/nurse moves to treat patient on bed 1
A15	3-to-6	After treating patient on bed 2 physician/nurse move to use the computer at the nurse station

For each RFID tag-base pair or tag-tag pair an encounter is recorded every 3 to 4.5 seconds. These data are captured in a time-modulated manner, i.e., encounter information is communicated by detecting differences in the time of the encounter rather than the frequency. This results in a sparse matrix when considering the entire tag-base station configuration. Figure 8 depicts a sample of the matrix generated. The encounters of a tag X with base stations A, B and C (gray filled boxes) are shown in a 60-second-long timeline. Linear interpolation is used to fill missing data in this sparse matrix. While this methodology provides an

RSSI value for all base stations at all instances, it adds some noise to the system that may affect the overall activity recognition accuracy.

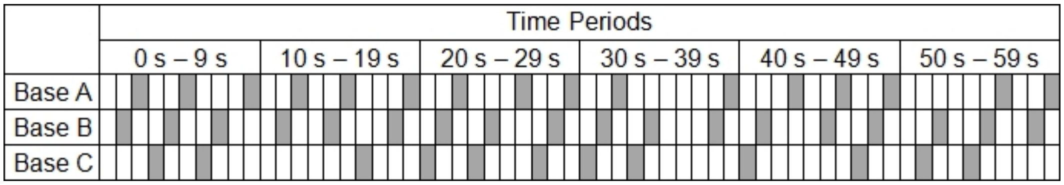

Figure 8. Sparse matrix of tag-base encounters

Results

Figure 9 summarizes the recognition accuracy for the 15 motion patterns (A1 to A15) elucidated in Table 2.

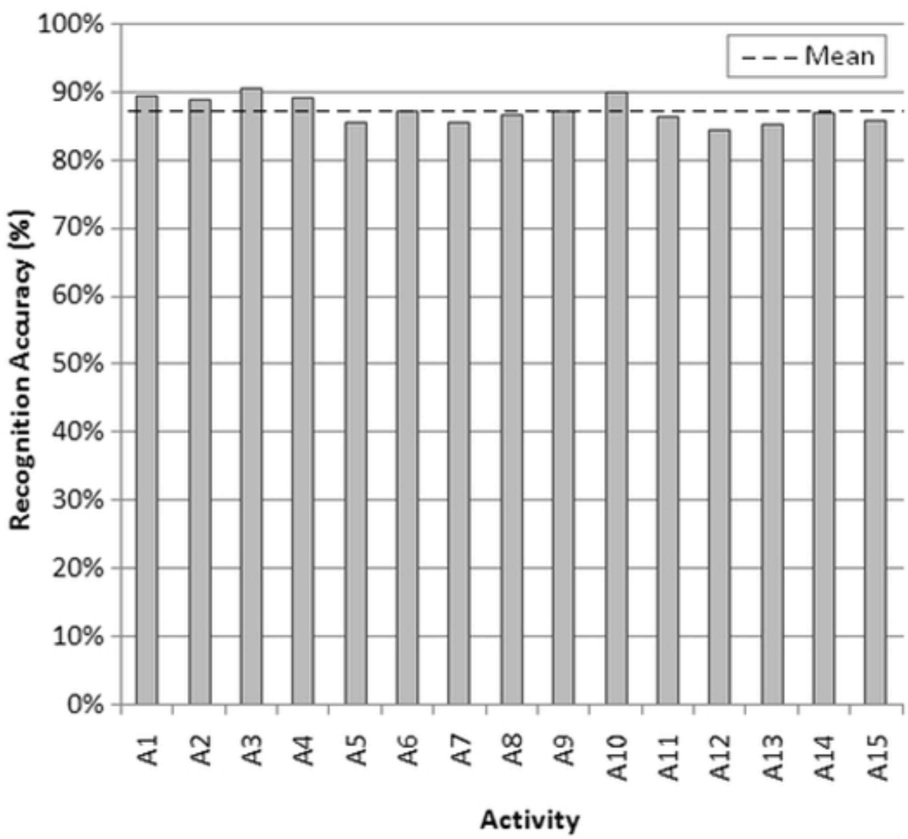

Figure 9. Hidden Markov Model (HMM) based activity recognition results

A mean recognition accuracy of 87.5% was obtained, with a maximum of 90.5% and minimum of 84.5%. The analysis of the incorrectly classified test samples revealed that misclassifications were a result of variations in the training set. As discussed previously, HMMs must be trained on a well-controlled sample that best represents the activity. Obtaining training data from real-world scenarios is bound to have variations that may compromise the quality of models generated. Additional sensors such as accelerometers could be utilized in conjunction with RFID tags to improve the activity recognition rates.

Study Limitations

The primary challenges to training HMMs for various activities lie in: (i) developing a controlled set of samples that best represent the activity being modeled, and (ii) the current limitations of RFID tags. The linear interpolation adopted for dealing with missing data introduces further errors into the system. The recognition accuracy of the system can be improved by: (i) increasing the sampling frequency of tags, (ii) using alternate methods of interpolation to fill the sparse matrix, and (iii) incorporating accelerometers with existing tags to refine data describing the movement. Since the current experiments were conducted in a lab setting, further evaluation and testing with multiple tags in critical care units would be required to complete the validation of this system.

A key limitation of this approach lies in reliance on movements and patterns of movements. Such an approach will noticeably miss the activities when the entities are not moving. However, in an environment such as critical care, a large percentage of activities do involve movements. An observer can capture activities that do not involve movements.

Although this system eases the burden on the observer (who can then capture high-level cognitive details of examination and leave the low level activity details to the automated system), in order to capture non-movement based activities in complex systems, and a classification of activities of interest is required. This research deals with studying adaptive behaviors in complex systems. The following chapter presents preliminary research in the development of a classification schema for deviations in trauma critical care.

DEVIATIONS FROM STANDARDS IN COMPLEX SYSTEMS

From the initial days when there was a simple doctor-patient relationship, healthcare today has expanded to include a multitude of factors that increase the complexity of the system. In order to cope with this complexity, clinicians tend to develop ad-hoc adaptations to function in an effective manner. It is these adaptations or "deviations" from expected behaviors that provide insight into the processes that shape the overall behavior of the complex system. In this chapter, a theoretical framework for assessing clinicians' deviations is presented.

Preliminary Classification of Deviations in Trauma Care

Deviations can be broadly defined as steps performed that are not on an accepted pre-defined standard. For the analysis of deviations in trauma care, the appropriate guideline or standard available is ATLS [51]. The preliminary classification is based on field observations of 10 cases conducted in a Level-1 trauma unit at Banner Good Samaritan Medical Center [62].

Deviations can be broadly classified as errors, innovations, and proactive and reactive deviations. Whereas errors are defined as deviations that potentially impact patients and their treatment outcome *negatively*, innovations are deviations from the protocol that may *positively* affect the patient's outcome. In addition to errors and innovations, there are some deviations that do not directly impact patient outcomes but rather are actions demanded by the dynamic nature of the complex environments. Deviations performed in reaction to patient-specific actions or condition changes are classified as reactive deviations, while steps taken to improve the efficiency of the trauma care by anticipating future needs are classified as proactive deviations. Using this analytic framework, individual (or unit) deviations identified using ATLS protocol for "Initial Assessment and Management" (detailed in Table 1), are classified to answer the following questions:

(i) How often do the trauma team members deviate from standard practice?

(ii) When clinicians deviate, what are the types of deviations made?

(iii) How do these types of deviations vary with the experience (level and type) of the members of the clinical team?

In the following section, the initial study conducted on deviations in trauma care and the associated results are elucidated.

Methods

Study Site Description

The field observations for this work were conducted in Banner Good Samaritan's trauma unit, one of 6 Level-1 trauma centers in the Phoenix metropolitan area. Approximately 3000 patients are treated annually in this 5-bed unit. The trauma center has dedicated hospital resources for the management of trauma patients throughout all aspects of care, including initial evaluation and resuscitation, acute care and rehabilitation. In addition, the trauma unit collaborates with surgeons from neurosurgery, cardiothoracic, vascular, orthopedic, plastics, ophthalmology, urology and internal medicine departments to provide the required care for incoming patients. The trauma team (present during every shift) includes 1 trauma resident, 2 trauma nurses, 1 trauma attending, 1 anesthesiologist, 1 to 2 junior residents, 1 to 2 medical students, and radiology and lab technicians. Trauma nurses supporting the trauma leader are experienced registered nurses (RNs) with 5-10 years of critical care experience.

Data Collection Methodology

This study was approved by the Institutional Review Board and the informed consents were obtained from the participants on each encounter. Field observations were gathered by one researcher over a period of 3 months from December 2009 to February 2010. Trauma cases that occurred between 9 am and 9 pm (Monday through Thursday) were observed. The researcher logged observations simultaneously as the trauma team treated the patient. All observations were gathered unobtrusively. Clarifications about the events that occurred were obtained from clinicians between trauma events.

Within the time period specified, a total of 10 trauma cases were observed with 7 attending trauma surgeons (experts), 7 junior trauma residents (novices in the first and second year of residency training) and 7 senior residents (in the third and fourth year of residency training). The trauma cases were of 2 types; trauma A and trauma B (trauma A refers to high criticality cases that require the presence of an anesthesiologist, while trauma B cases are those cases that are classified as low criticality). Out of the 10 cases observed, eight cases were trauma B cases and two were trauma A.

Analysis Methods

The ATLS standard for Initial Assessment and Management was utilized to assess these cases for deviations. Irrespective of the types of the cases, all steps of the Initial Assessment and Management are required to be followed by the core trauma team. This allows for a valid comparison between the 10 trauma cases.

The analysis of the data was performed by one researcher in collaboration with an expert trauma clinician (an attending). Deviations identified (through consensus) are classified as errors, innovations, proactive or reactive deviations based on the preliminary classification schema. The data set was then analyzed using statistical means and interpreted to answer the questions outlined in the previous section. Independent group t-test was used to find the differences between numbers and types of deviations in trauma A and trauma B cases. A p-value of $p<0.05$ was accepted as statistically significant.

Results

Mean Deviations per Trauma Case

The results are presented as mean (μ) ± standard deviation (σ). Figure 10 depicts the mean deviations that occurred in the 10 trauma cases for: (i) trauma A and trauma B (9.1 ± 2.14), (ii) trauma A (14 ± 1.41), and (iii) trauma B cases (7.5 ± 2.79). The mean numbers of deviations in trauma A cases were higher compared to the mean deviations in trauma B cases. Typically, trauma A cases involve unstable and unpredictable patients. Consequently, the trauma team makes a relatively larger number of deviations to adapt to the dynamic situation at hand.

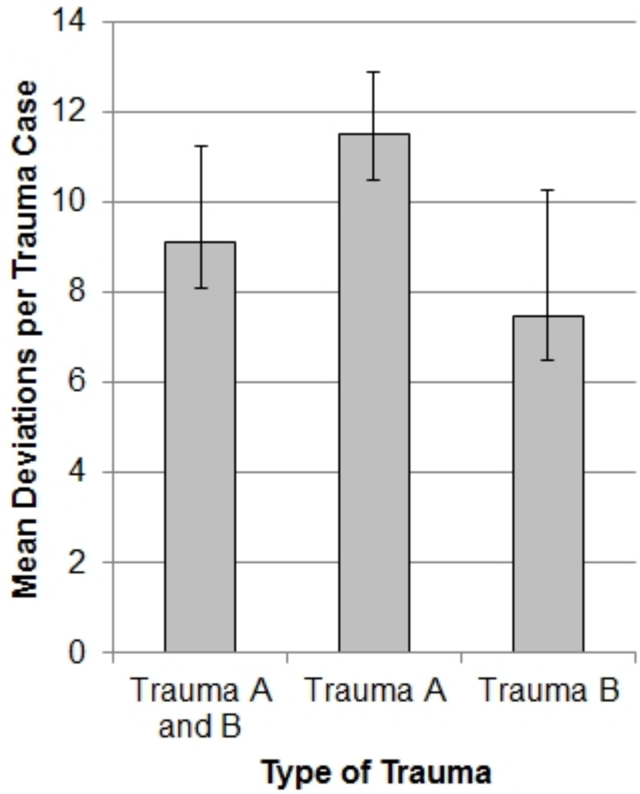

Figure 10. Mean deviations per trauma case

Deviation Distribution and Trauma Severity

Figure 11 shows the distributions of (i) errors (trauma A: $\mu = 1.5 \pm 1.06$, trauma B: $\mu = 2.63 \pm 1.1$), (ii) innovations (trauma A: $\mu = 0.5 \pm 0.35$, trauma B: $\mu = 0.75 \pm 0.7$), (iii) proactive deviations (trauma A: $\mu = 0.5 \pm 0.35$, trauma B: $\mu = 0.38 \pm 0.37$), and (iv) reactive deviations (trauma A: $\mu = 11.5 \pm 1.06$, trauma B: $\mu = 4.13 \pm 1.15$). From Figure 11, it can be seen that errors make up a small percentage (26.38%) of the total deviations in the 10 trauma cases. This is an important result from these observations, since it points to the limitations of the current strategy of marking most deviations as errors in assuring compliance to a standard.

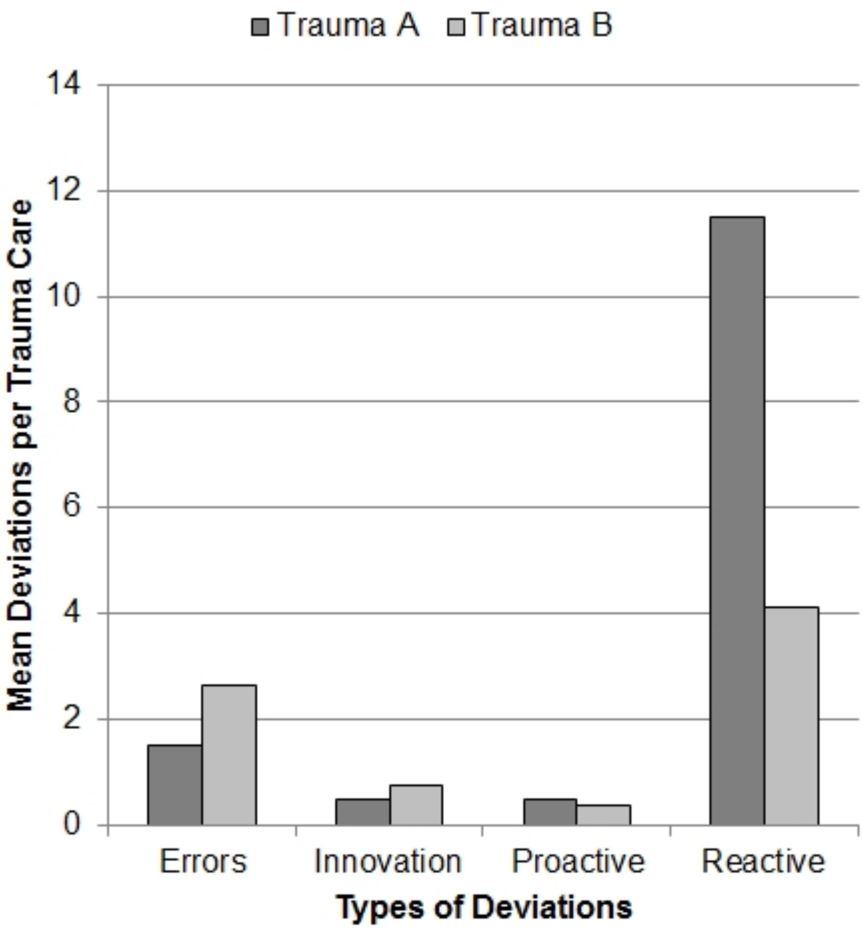

Figure 11. Deviation distribution in two trauma settings

The proactive and reactive deviations were significantly higher in trauma A when compared to trauma B cases ($p<0.05$). The critical condition of the patients in trauma A cases and the individual nature of the problem cause the trauma team to deviate often in order to manage the unique situation at hand. The analysis also showed that most of the deviations were reactive in nature, in both trauma A and trauma B cases. This can be attributed to the dynamic nature of the critical care environment. Clinicians are required to react quickly to the changes to ensure efficient operation in trauma care.

Deviation Distribution and Clinician Expertise

Figure 12 depicts the total number of errors and innovations made by core team members in the 10 trauma cases observed.

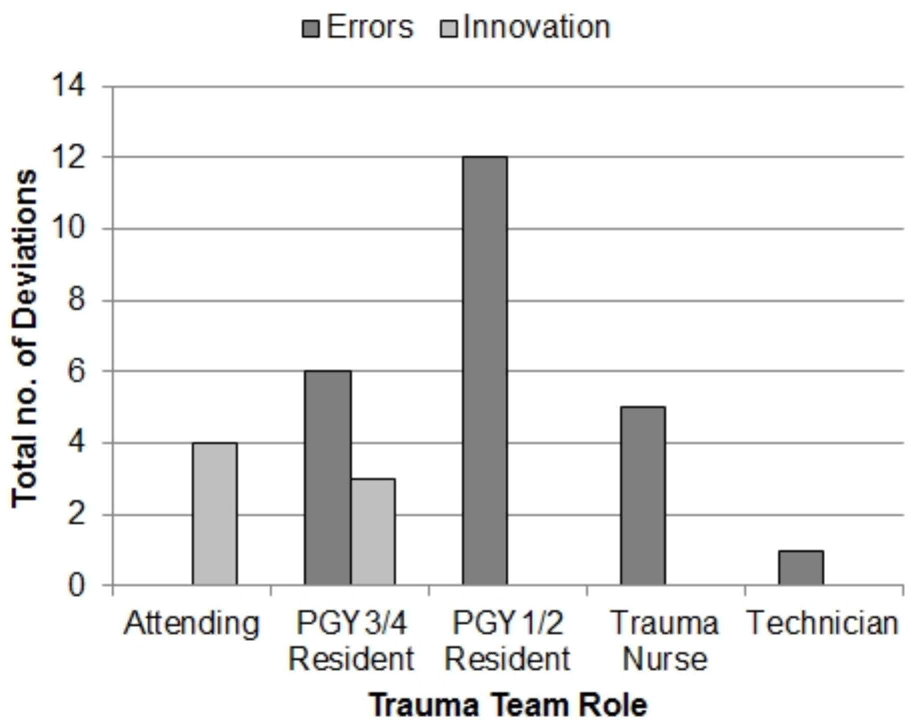

Figure 12. Error and innovation as a function of expertise

In this study, the experts made no errors as defined in the analytic framework. Care givers with lesser expertise (from the 3rd and 4th year resident to the 1st and 2nd year residents), made fewer innovations, when compared to the experts (attending trauma surgeons). While intermediate clinicians (3rd and 4th year residents) made more errors compared to the attendings, novices (1st and 2nd year residents) made more errors than any other group of clinicians. Trauma nurses and technicians show little evidence of innovation. Although this low frequency of innovation cannot be attributed to a lack of experience, it can be hypothesized that within the confines of their roles in interacting with a patient, there is not much scope for innovation. Nurses and technicians are trained to follow a strict protocol to support the trauma team, and that training may be responsible for the observed patterns.

Figure 13 provides a snapshot of the distribution of proactive and reactive deviations within the trauma team.

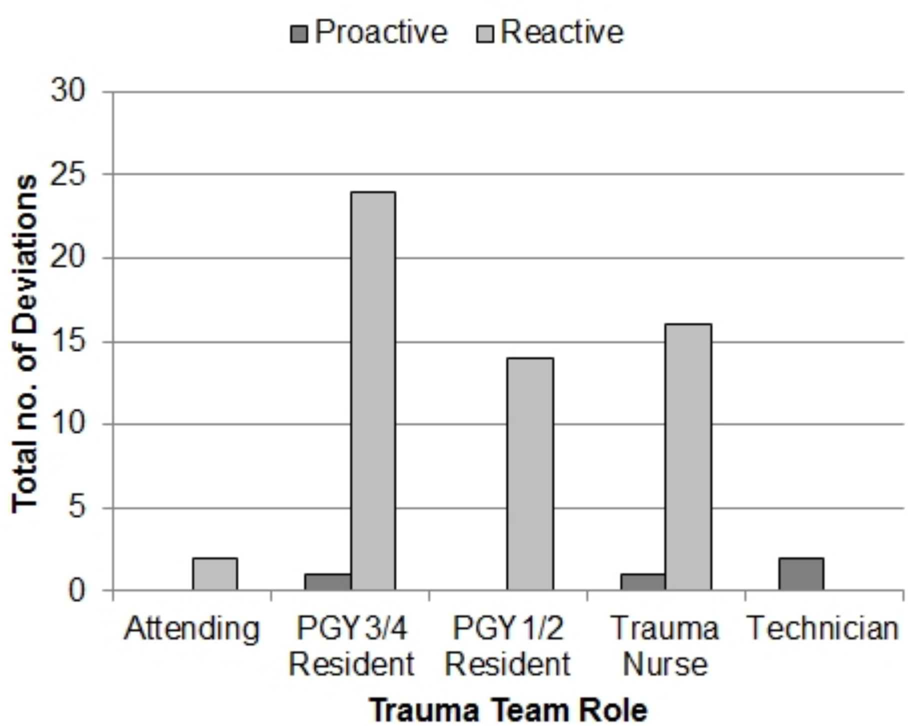

Figure 13. Proactive and reactive deviations as a function of expertise

It shows that senior residents make the most reactive deviations (because they are performing bulk of the tasks), followed by the trauma nurses. Junior residents who generally assist but may lead a few trauma cases also made a significant number of reactive deviations. These observations show that leadership role and associated tasks may be connected to generating deviations to the protocol.

Study Limitations

This study provides supportive evidence for the claim that deviations do occur in critical care environments and not all deviations are errors. Deviations from the standard can be important innovations and are tied to complex decision-making and judgment calls at the point of care. The results from this study show that expertise of the caregivers and criticality of a patient's condition influence the number and type of deviations from standard practice.

Although this research was a novel approach for assessing protocols and guidelines, there were not enough subjects studied to enable tests of significance. In addition, errors and innovations were defined in terms of patient outcome. The causal effect between deviations and specific patient outcomes may be difficult to track in critical care environments. For this reason, there is a necessity to define deviations in relation to protocols and guidelines instead. This will also enable definitions to be more generalizable to other critical care environments.

EXTENDED CLASSIFICATION OF DEVIATIONS IN TRAUMA CARE

This chapter describes research where deviations from standards are examined as a function of expertise and teams in complex critical care environment [63]. This work builds on research described in the previous chapter.

Deviations from Standards and Expertise

From a cognitive perspective, error, innovation and effectiveness in carrying out a protocol is intimately linked with expertise of the clinicians. Patel et al. studied the relationship between task difficulty and expertise [64]. The authors employed semantic analysis and found that experts were able to use a well-developed knowledge base and superior reasoning strategies in clinical reasoning. Patel and Groen [65], in another publication, isolated the reasoning process that physicians go through when diagnosing a clinical case, using techniques to identify knowledge structures. They showed that in medicine, experts tend to follow a top-down reasoning strategy wherein reasoning from a hypothesis is done to account for the case data, which seemed anomalous when compared to other domains. This is an important finding from the perspective of studying errors and innovations. In other domains wherein experts tend to gather data and assemble hypotheses, there is scope for significant amount of trial and error. On the other hand, in clinical decision-making, experts more often than not utilize a top-down approach to decision making. It has been shown that this methodology when combined with experience-driven cognitive constructs results in experts making fewer errors compared to novices. It is plausible that when experts do deviate, the deviations are more likely to be innovations.

Another aspect of cognition that needs to be accounted for is the capability of a clinician to generalize the given data into correct diagnoses. Cognitive research in medicine [66] has shown that clinicians can generate different levels of mental representations, from the very specific to the very general. The critical factor in determining generality is typically the degree of high-level expertise of the clinician, namely, specialized or specific expertise (i.e., knowledge of a particular sub-domain of medicine, such as endocrinology or cardiology). Higher-level representations are generated by these more expert clinicians; whereas lower-level and more detailed representations are typically generated by novices, or more commonly, intermediate level clinicians (e.g., senior medical students, recent graduates, and residents).

This condition points to the ability of experts to apply generic rules to a given case, giving them extra cognitive resources to apply innovations and limit errors. Research has shown that experts as a result of their practice, learn to associate individual items in working memory with the contents in long-term memory, which result in the development of conceptual organizations in memory called retrieval structures [67]. An expert can use these retrieval structures to provide selective and rapid access to long-term memory. On the other hand novices seem to occupy their working memory and long-term memory resources in the details of the case (due to the lack of mature retrieval structures), which may be irrelevant. In such type of workload, it may be challenging to innovate, and, depending on the workload, one may make extensive errors, as is the case in complex environments. In fact, research confirms that a key element of retrieval structures is their use by experts to eliminate irrelevant information [68] freeing working memory for innovative thinking.

In general the literature on clinical expertise, gives clues into the underlying mechanisms of the relationship between errors and innovations. One area of research that has explored the mechanisms of innovations is study of the cognitive basis of creativity [69]. This field explores the cognitive basis underlying creative thinking and reasoning. It identifies conditions that lead to creativity and innovation and is based on the hypothesis that creativity is supported by pre-invention structures and the explanation structures in experts. This is a very intriguing model for creativity and cognition, but its relevance to complex domains such as trauma care may be limited. In general, the theories from creativity tend to focus on a freethinking approach wherein timeliness of creativity is not a big factor. On the other hand, in complex environments such as trauma critical care, timeliness of decision-making may fundamentally alter the innovation process, and it is important to study the mechanisms underlying errors and innovations separately.

Deviations from Standards and Team Decision Making

In critical care settings, teams of care professionals care for patients. These teams typically involve clinicians with varying backgrounds and expertise, working in a collaborative manner. These teams operate in environments with dynamic social structures [39] and are required to adapt to varying task demands and coordinate their efforts to complete activities necessary for task completion [50]. Team decision-making is a key factor that impacts co-ordination between individuals involved in the patient care process.

Cannon-Bowers, Salas and Converse [70] define team decision making as a "team process that involves gathering, processing, integrating and communicating information in support of arriving at task-relevant decisions". It is a process that requires individuals to apply their expertise to filter data and communicate relevant information and recommendations to other team members. This can be affected by a number of environmental factors such as situation complexity, time pressures, multi-component decisions and evolving (at times ambiguous) information [71]. Effective decision-making relies on the emergence of shared mental models and cognition among all the providers involved in the care process [14].

Shared mental models reduce the communication required to co-ordinate decisions and activities required to complete a particular task. The team members perceive and interpret situations in a similar manner. This enables the team to make decisions and take action effectively. Providing teams with the tools to promote the development of shared mental models and cognition is critical to the development of coordinated healthcare delivery. Protocols, standards and guidelines are one such set of tools. Analysis of deviations from standards will allow researchers to evaluate whether adaptations made by the team members contradict the purpose of the guideline and consequently compromise patient safety or quality of care.

Extended Framework for Deviations from Standard Practice

Figure 14 depicts the hierarchy for an extended classification of deviations.

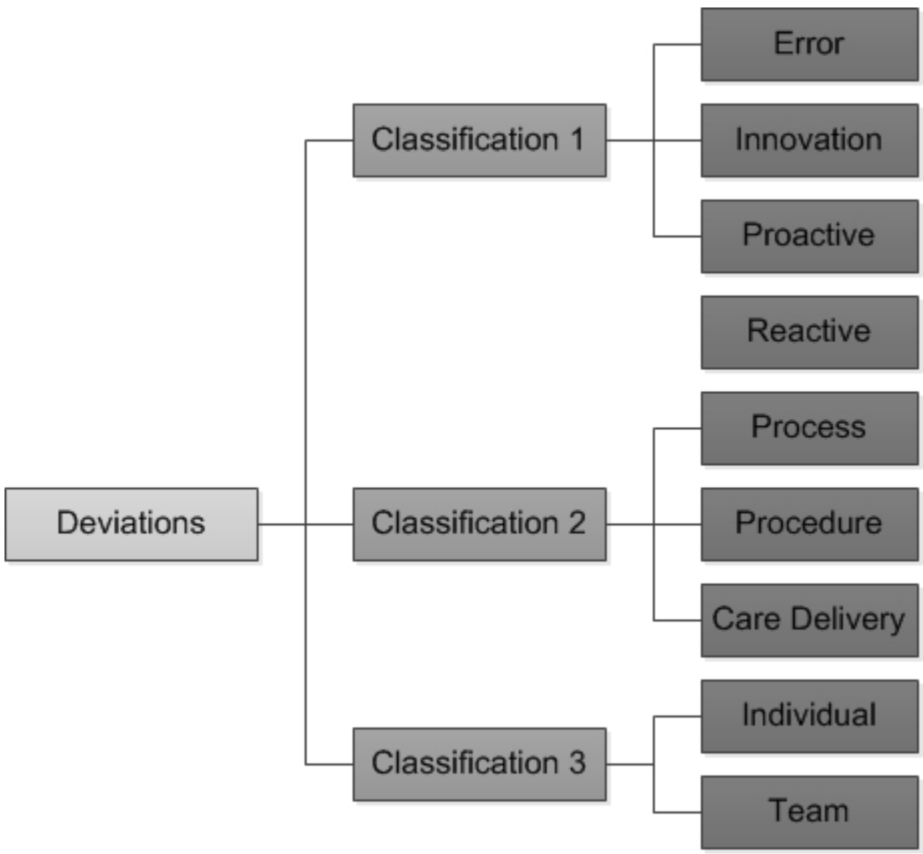

Figure 14. Extended classification of deviations in trauma care

In previous research [62], deviations were classified as (i) errors, (ii) innovations, and (iii) proactive and (iv) reactive deviations. Further examination showed that deviations could also be classified by how they affected the trauma care (Classification 2), and how many members were involved in making the decision (Classification 3). In this section, the previous classification of deviations is revisited (providing more concrete definitions for the ideations of error and innovation) and an extended classification is presented.

Classification Schema 1

Deviations as Errors

An error is defined as a deviation from the standard, if it: (i) violated a prescribed order of activities with a negative impact on workflow, (ii) resulted (directly or indirectly) in compromising patient care, or (iii) resulted in an activity being repeated due to failure in execution or a loss of information. Examples of errors encountered in the trauma cases observed in this study include:

- A resident completed the secondary survey prior to ordering chest and pelvis x-rays. Consequently, obtaining these x-rays for diagnosis was delayed. In this case, the sequence in which the tasks were performed violated the order prescribed in the ATLS standard. Since this deviation caused a delay in receiving information critical to treating the patient in a timely manner (and thereby negatively impacting workflow), it was classified as an error.
- A junior resident attempted to remove the spine board before the patient's spine was cleared (confirmed not to be injured). This deviation directly compromised patient care and consequently was classified as an error.
- The lab technician needed to redraw a sample for blood work when additional tests were ordered. The previous sample had been discarded. A

lack of communication within the team resulted in this deviation. While not as severe as the previous error, the repetition of a task by a team member due to a failure in communication was classified as an error.

Deviations as Innovations

Innovations are defined as deviations that potentially benefit the individual, team or patient by bringing novelty to the situation at hand [72]. Some examples of innovations identified in this study are given below:

- A patient required a translator in order to communicate with the resident. The team was unable to find a translator. The attending asked the trauma nurse to see if the patient's family could help. The patient's sibling was able to come into trauma facility and act as a translator. This allowed the resident to continue with his examination, leading to successful assessment and treatment of the patient. The standard protocol of seeking an in-house translator was violated. A novel step (that resulted in a positive outcome) was introduced in the workflow, which qualifies as an innovative deviation.
- A patient was nervous about the damage done to his face due to an accident. In order to calm the patient, the nurse provided him with a small mirror so that he could assess the damage for himself. The patient then relaxed. For such a case, the guideline provides no instruction on how to deal with a difficult patient. The clinician deviated by introducing an

action outside the scope defined by the guideline to successfully care for the patient.

- The resident examined a patient's leg injury (in fewer than 15 seconds), and ordered an x-ray of the extremity along with chest and abdomen x-rays. By introducing a brief examination of the injury site, the resident was able to anticipate a future need and advance a step in the standard. The results were relayed back to the team more promptly than if the prescribed order of steps had been followed. The introduction of a novel step that resulted in a positive outcome on the workflow was considered to be an innovation.

Proactive Deviations

A proactive deviation occurs when (i) an activity is performed (without compromising patient care) in anticipation of a future requirement (or lack thereof) when treating a patient or (ii) an activity (which may be out of the bounds of an individual's role in the trauma team) is performed in order to correct or prevent error occurrence. Some examples of proactive deviations encountered in the trauma cases observed include:

- A radiology technician set up the x-ray sensor board for a chest x-ray prior to the trauma arrival, since the trauma team had been notified about the nature of the trauma case.

- A trauma nurse called the radiology unit to let the unit know that the technician would not be required, since the scans had already been taken in the previous facility.
- The trauma nurse reminded a junior resident that c-spine results have to be received prior to removal of the spine board.

Reactive Deviations

Reactive deviations occur when an activity is performed in reaction to an unanticipated event or change in patient condition, diagnostic process or treatment plans. Examples of reactive deviations found in this study include:

- A patient was violently reacting to pain and needed to be held down by the trauma team in order to complete the primary survey and intubate the patient (if necessary).
- The results of the x-ray ordered were inconclusive. As a result, the resident ordered an angiogram.
- A patient, concerned about his facial injuries, requested a plastics consult. The treatment plan had to be altered to accommodate the patient's request.

While in this study, errors, innovations, proactive and reactive deviations are treated as mutually exclusive, in reality there may exist an overlap between these categories. While further investigation is required is assess of the schema should be modified to examine inter-relationships between the categories, for this exploratory study the categories are treated as mutually exclusive groups.

Classification Schema 2

In addition to classify deviations by the impact they may have on workflow, deviations may also be classified by how they impact the steps of the trauma standard. Based on the granularity of the step deviated from and the type of activity performed, deviations may also be classified as (i) process-related, (ii) procedure-related, or (iii) care delivery-related deviations.

Process-related Deviations

Deviations that may be related to how the guideline is implemented are classified as process-related deviations. Examples of process-related deviations include log roll not being performed correctly or an x-ray being ordered after the secondary survey. In both examples, clinicians deviated from the recommended method for guideline implementation.

Procedure-related Deviations

In contrast to process-related deviations, procedure-related deviations deal with how a specific step in the guideline is performed. An example of a procedure-related deviation is a clinician making an error in stapling a wound. The key difference between process- and procedure-related deviations lies in the granularity of unit activities in trauma care. Changes in order or presence/absence of activities are considered as a process-related deviation, whereas changes made to the unit activity itself are procedure-related.

Care Delivery-related Deviations

Any deviation dealing with the care provided to the patient (not specified in guidelines) is classified as a care delivery deviation. These deviations include a nurse providing a mirror to a patient concerned by facial injuries or providing medications for a patient in pain. Whereas procedure related deviations typically involve medical interventions, care delivery-related deviations involve activities performed that support the trauma team and patient.

Classification Schema 3

Finally, deviations may be differentiated by the number of trauma team members involved in the decision making process that ultimately resulted in the occurrence of the deviation. Deviations may be classified as (i) individual, or (ii) team deviations.

Individual Deviations

Deviations initiated by a single clinician are classified as individual deviations. Examples of individual deviations include a resident making an error in a procedure, or an attending suggested a novel methodology for a step in the protocol or a trauma leader proactively performing certain steps in the protocol. In each of these cases the deviations were initiated by a decision made by a single individual.

Team Deviations

Whereas an individual may initiate many deviations, some deviations occur at the team level. Such deviations involve more than one clinician participating in the event. For example, a resident may decide on an alternate course of treatment based on a discussion with his attending or the team. Such a deviation is classified as a team deviation. Table 3, Table 4 and Table 5 summarize the terminology involved in classifying deviations as described in this section.

Table 3. Summary of classification schema 1

Error	Related to standard practice: ▪ Task order violation ▪ Task omission ▪ Task repetition due to communication or execution failure Impact on workflow: **NEGATIVE** ▪ Causes delays ▪ Compromises patient care
Innovation	Related to standard practice: ▪ Novel task addition Impact on workflow: **POSITIVE** ▪ Improves workflow efficiency ▪ Improves quality of patient care
Proactive	Related to standard practice: ▪ Task advancement ▪ Error prevention ▪ Out of role expectations Impact on workflow: **NUETRAL** ▪ No observable impact on workflow or patient care
Reactive	Related to standard practice: ▪ Common task addition in response to random event Impact on workflow: **NUETRAL** ▪ No observable impact on workflow or patient care

Table 4. Summary of classification schema 2

Process	Related to standard practice: ▪ Deviations (including task order change, omission and addition) that related to how the standard is implemented
Procedure	Related to standard practice: ▪ Deviations (including tasks repeated due to execution failure) that related to medical interventions provided to the patient
Care Delivery	Related to standard practice: ▪ Deviations related to supportive care interventions provided to the patient

Table 5. Summary of classification schema 3

Individual	Initiated by decision making process of a single clinician in the team
Team	Initiated collaboratively by two or more clinicians in the team

It should be noted that the three types of classification schema are treated as independent of one another. A team deviation can be an error or an innovation, for example. Such a classification allows researchers to examine the context of various types of deviations. This can further the understanding of various factors that contribute to deviations.

In the research described in this chapter the following are explored; (i) various types of deviations that occur in trauma care, (ii) how they relate to expertise, and, (iii) whether they were initiated by an individual or by a team. The following section describes the ethnographic study performed in a Level-1 trauma unit and presents the results of applying the described deviation classification schema on the various deviations identified.

Methods

Field observations for this work were conducted by one researcher from September 2010 to December 2010 at Banner Good Samaritan's Level-1 trauma unit. A total of 20 trauma cases were observed. This, added to the 10 trauma cases previously observed, resulted in a total of 30 cases with 15 cases being led by 4th or 5th year (senior) residents and 15 cases led by 2nd or 3rd year (junior) residents. Out of the 30 cases, 6 cases were categorized as trauma A (patient in critical condition) and 23 cases as trauma B (moderate criticality of patient). One case was classified as trauma C as it involved a pregnant woman. As patient identifiers such as Glasgow coma scale (GCS) and injury severity score (ISS) were not captured (the protocol involved shadowing clinicians alone), the classification of the trauma is used as a proxy to assess severity of the incoming patient.

The trauma cases were observed by one researcher using the A(x4) model [73]. This model requires contextual observations (snapshots) to be captured by highlighting 4 key parameters, namely, actors, activities, atmosphere and artifacts. Observations captured in this manner provide rich contextual descriptions of the situation, which is required for analysis of deviations.

Each time-stamped observation was compared to the corresponding step in the ATLS guideline [51] in order to determine (i) if a deviation had occurred, (ii) the type of the deviation and (iii) if the deviation resulted from individual or team-level processes. The data were analyzed iteratively until the number and type of deviations stabilized. The analysis methodology is similar to the methods described in the preliminary analysis of deviations [62].

This study was approved by the Institutional Review Board and the informed consents were obtained from the participants on each encounter.

Results

A total of 165 deviations were identified from the 30 trauma cases observed. Of these deviations, 4 were found to be related to auxiliary activities in trauma care. The activities corresponding to these deviations included (i) attendings teaching residents specifics of trauma care, and (ii) clinicians gathering evidence in trauma cases that resulted from criminal activities. These deviations are unrelated to trauma team expertise or guideline implementation. Consequently they were omitted from the analysis.

The 161 remaining deviations are described categorically using the variables (i) training of the resident leading trauma care (Variable - Leader), (ii) role played by clinician initiating the deviation in the trauma team (Variable – Role), (iii) phase of the trauma standard at which the deviation took place (Variable – Phase), (iv) deviation type based on classification schema 1 (Variable – Class1), (v) deviation type based on classification schema 2 (Variable – Class2), and (vi) deviation type based on classification schema 3 (Variable – Class3). The severity of the trauma case was not considered as a variable as a disproportionate number of trauma B cases were observed compared to trauma A during the duration of the study.

For each pair of variables, Chi-square analysis was performed to tease out relationships that may exist. Table 6 summarizes the results of pair-wise relationship tests conducted for the variables described. Significant relationships (p-value <0.05) are indicated by bold font.

Table 6. Chi-square p-values of pair-wise relationships between variables

Variables	Leader	Role	Phase	Class1	Class2	Class3
Leader	-	**<0.0001**	0.2975	**0.0150**	**0.0258**	0.8469
Role	**<0.0001**	-	**<0.0001**	**<0.0001**	**0.0174**	0.8129
Phase	0.2975	**<0.0001**	-	**<0.0001**	**<0.0001**	0.0648
Class1	**0.0150**	**<0.0001**	**<0.0001**	-	**0.0002**	0.0720
Class2	**0.0258**	**0.0174**	**<0.0001**	**0.0002**	-	0.7919
Class3	0.8469	0.8129	0.0648	0.0720	0.7919	-

From Table 6 it is seen that (i) expertise of the trauma leader, (ii) the phase in which the deviation occurs, and (iii) the role played by the clinician have significant relationships with types of deviations made. There is also an indication of a strong association between classification schema 1 and schema 2. It should be noted that near-significant relationships are found between classification schema 3 and schemas 1 and 2. This indicates a possible relationship that may need additional data to verify its validity. In the following sections, the individual significant relationships are further characterized.

Deviations and Expertise of Trauma Leader

Although no significant difference was found in the frequency of deviations, the types of deviations made were found to be related to the experience level of the clinician leading the trauma. Chi-square analysis between team leader and deviations classified using schema 1 showed significant relationship between these variables (Chi-sq = 10.4608, df = 3, p = 0.0150). Figure 15 depicts the relationship between the experience level of the trauma leader and errors, innovations, proactive and reactive deviations.

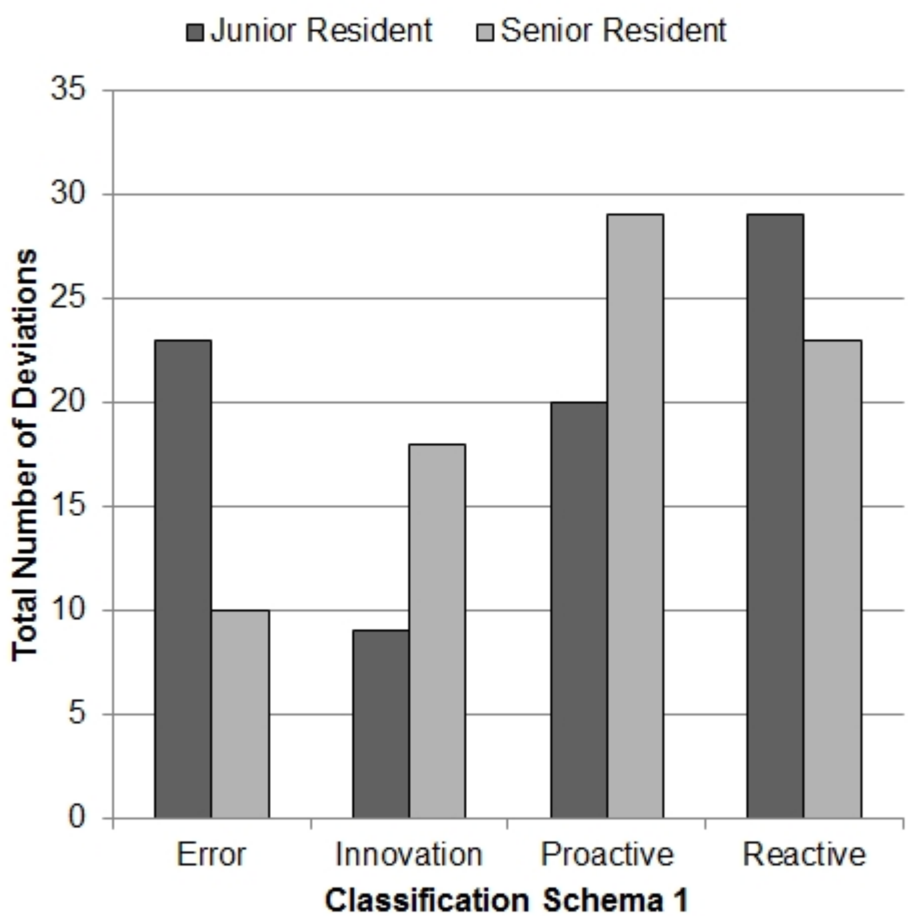

Figure 15. Deviations (classification schema 1) and expertise of trauma leader

Trauma cases led by senior residents had more proactive deviations and innovations compared to cases led by a junior resident. Errors and reactive deviations were found to be greater in cases led by junior residents. These finding suggests that (i) trauma leaders with more experience are able to adapt (making innovations) to the dynamic environment while minimizing errors, and (ii) experience enables leaders to guide a more proactive trauma team. Thus, it can be hypothesized that the proactive nature of expert trauma leaders enables them to anticipate future needs and possible errors, thereby minimizing resource wastage and unnecessary negative impact on patient outcomes.

A significant relationship was also found between the experience level of the team leader and deviations classified using schema 2 (Chi-sq = 7.3179, df = 2, $p = 0.0258$). Figure 16 depicts the relationship between leader expertise and process-, procedure-, and care delivery-related deviations. Cases led by junior residents had fewer care delivery-related deviations and more procedure-related deviations compared to cases led by a senior resident. Junior residents focused more on specific procedures. This is indicative of their level of training. Senior residents have mastered procedures, and can focus on developing other skills, such as communication. The number of process-related deviations was found to be similar for the two groups.

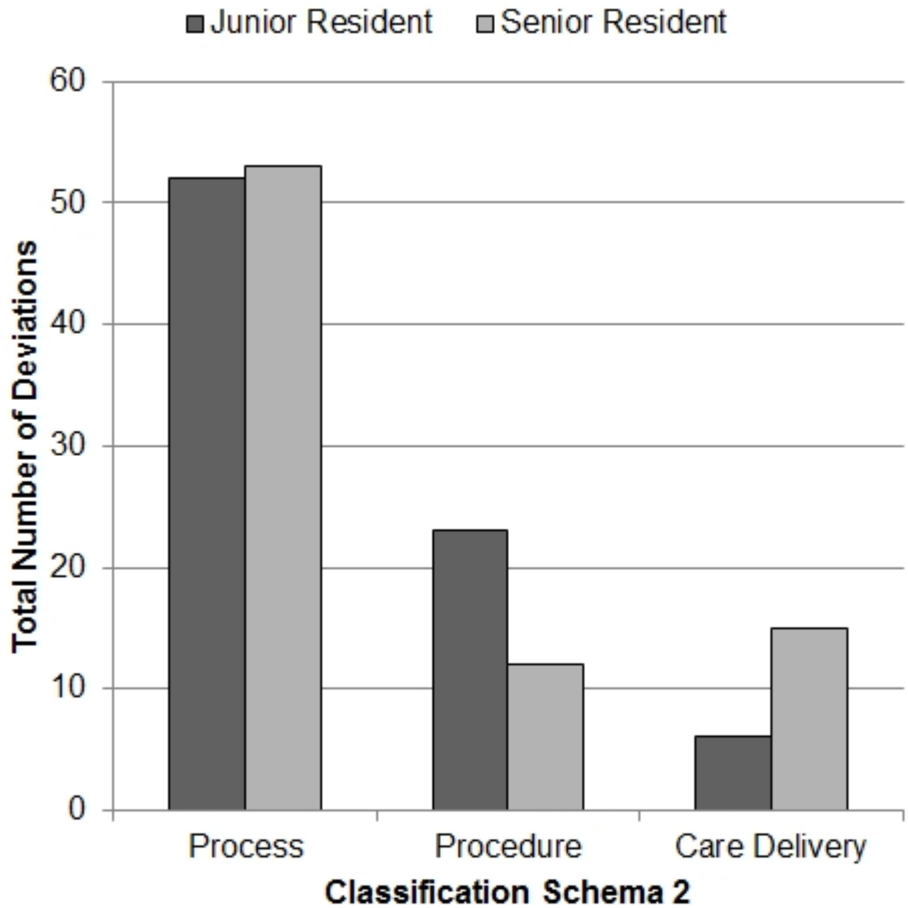

Figure 16. Deviations (classification schema 2) and expertise of trauma leader

Finally, Figure 17 depicts the significant relationship (Chi-sq = 83.7175, df = 4, p = <0.0001) between role of the clinician in the trauma team (junior resident, senior resident, attending, nurse and technician) and expertise of trauma leader. Whereas the statistics indicate a strong association between the variables, this could largely be attributed to the importance of the trauma leader handing a case. As seen in Figure 17, most deviations are made by the leader. Consequently, it is difficult to draw conclusions about flexibility of leadership based on the data available. However, it can be seen that the attending plays a larger role in cases led by a junior resident. This is expected in a teaching setting.

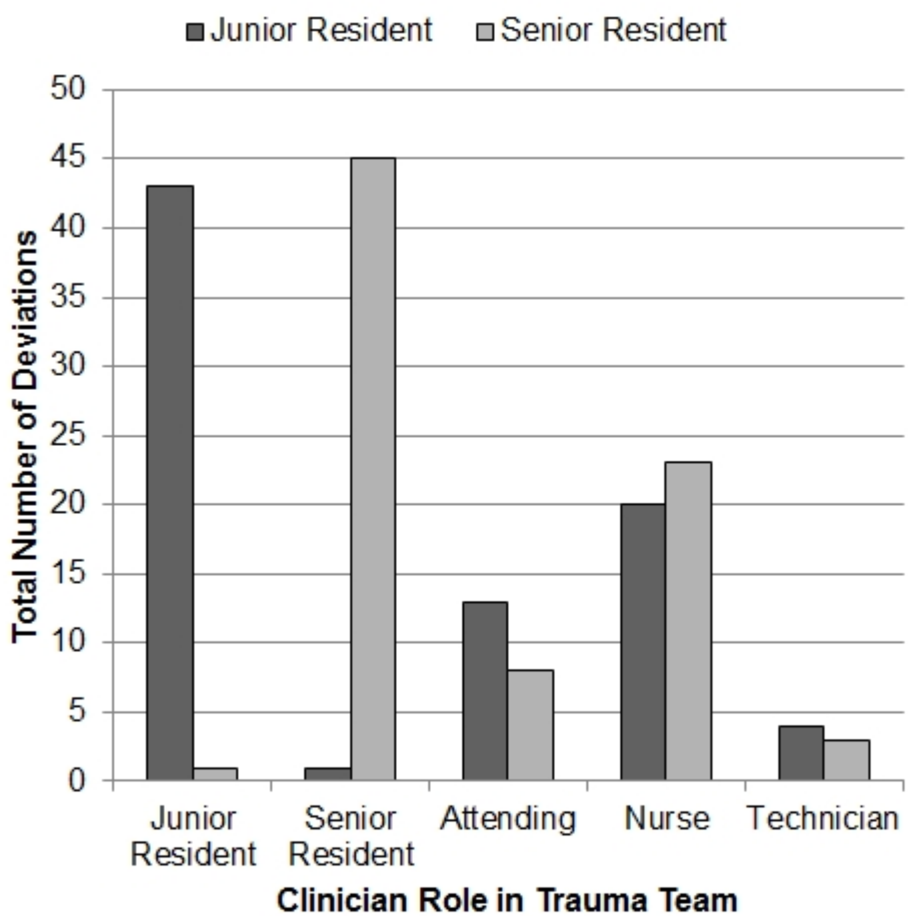

Figure 17. Deviations (clinician role) and expertise of trauma leader

Deviations and Phases of Trauma Standard Protocol

Figure 18 shows total number of deviations identified at each key stage in the trauma management standard (Phase 1: Trauma Preparation, Primary Survey and Resuscitation, Phase 2: X-ray and Diagnostic Studies, Phase 3: Secondary Survey, Phase 4: Tertiary Survey and Definitive Care). A greater number of deviations were found to occur in the phases following trauma preparation and primary survey and resuscitation (Percentage of deviations in Phase 1: 13.04%, Phases 2-4: 86.96%).

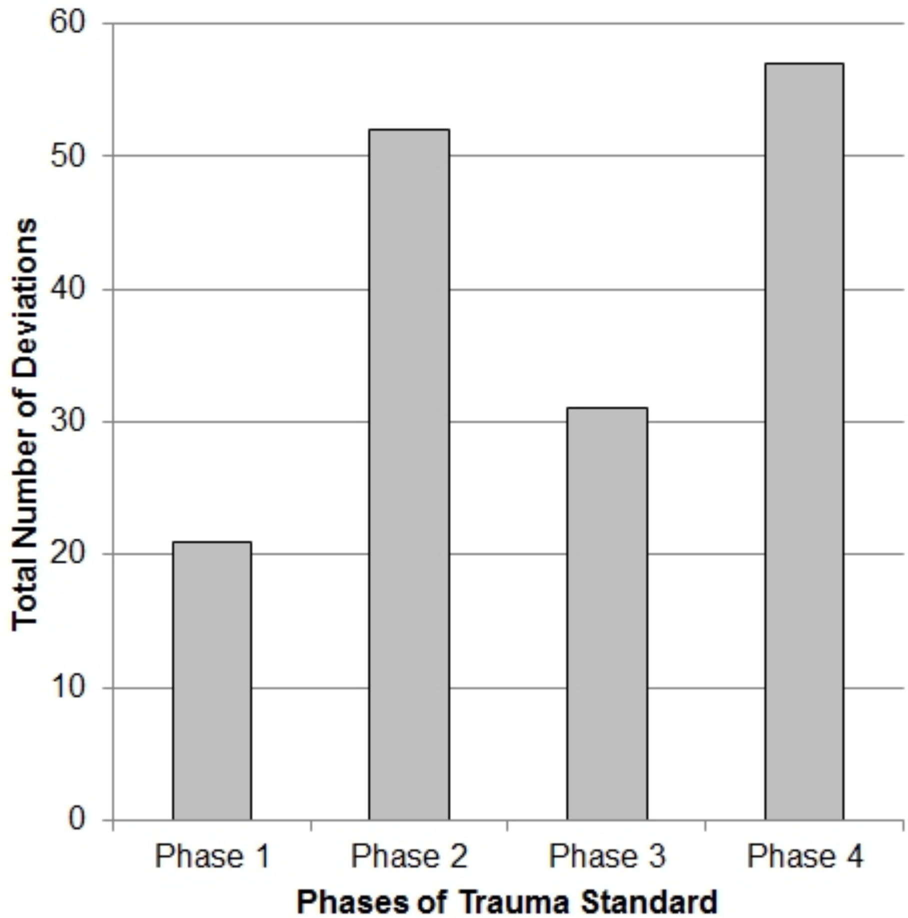

Figure 18. Total number of deviations in phases of trauma standard

Using chi-square analysis, a significant relationship was found between the phase in the standard and deviations classified using schema 1 (Chi-sq = 44.255, df = 9, $p < 0.0001$). As seen in Figure 19, errors occur throughout the various stages of the trauma care, whereas innovations only occur once the primary survey is completed. This is indicative of the level of adaptability the guideline allows for in the earlier stages of trauma treatment. The primary survey is protocol-driven, whereas the secondary survey and definitive care are more flexible, allowing the trauma team to deviate and adapt to the case at hand.

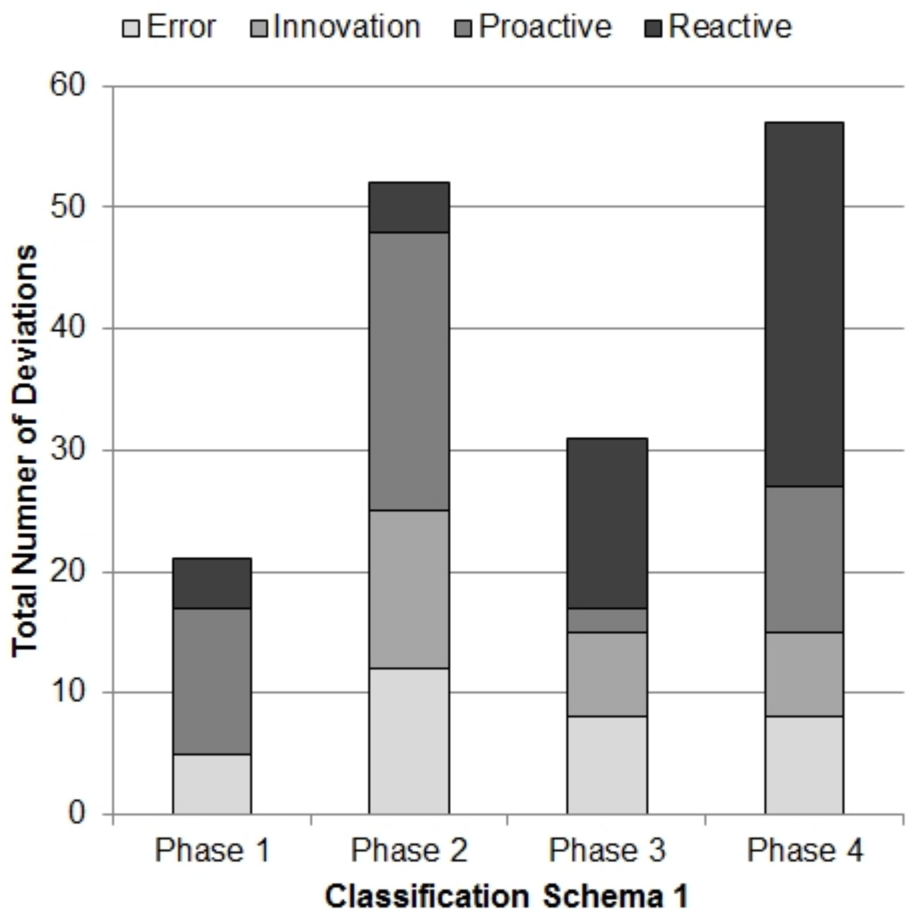

Figure 19. Deviations (classification schema 1) and phases of trauma standard

The key difference between an expert clinician and a novice is that expert clinicians deviate within the flexible portions of the guidelines, resulting in innovations. Novices, on the other hand, do not possess the necessary knowledge to understand the broader implications of their actions. Deviations made in critical steps, such as the primary survey, would result in error.

In addition to errors and innovations, it can be seen that more proactive deviations occur in the earlier stages of the trauma standard, while reactive deviations occur in the tertiary survey and definitive care stages. This is expected. As more information becomes available to the team, decisions about care of the patient may be altered in a reactionary manner.

Figure 20 shows the relationship between phases of the trauma standard to deviations classified using schema 2 (Chi-sq = 40.0974, df = 6, $p < 0.0001$).

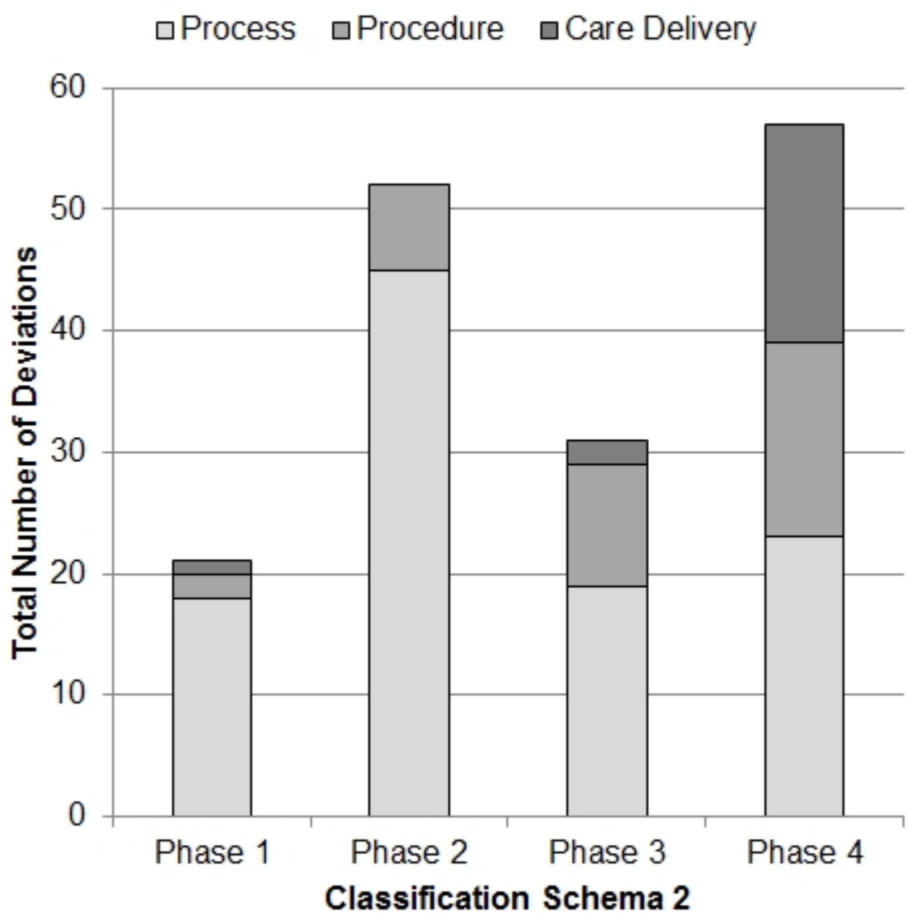

Figure 20. Deviations (classification schema 2) and phases of trauma standard

The total number of process-related deviations is higher when x-ray and diagnostic tests are ordered (27.95% in Phase 2). This indicates that certain steps in trauma treatment may be more adaptable than others. Identifying such critical steps and monitoring the deviations that occur could provide more information that will help direct guideline updates. In addition to the differences in process related deviations, it is interesting to note that procedural deviations linearly increase as trauma care proceeds through the various phases. This is expected, because the initial phases of the trauma care are more focused on examination of the patient. Once a diagnosis is made and results from x-rays and diagnostic tests are obtained, interventions to treat the patient trauma are performed. It should also be noted that supportive care delivery deviations occur largely in Phase 4. In Phases 1-3, the focus of the team is in examining the patient. Supportive case is usually provided after these phases are completed.

Figure 21 shows the relationship between phase of trauma standard and deviations classified role played by clinician in the trauma team (Chi-sq = 51.3650, df = 12, $p < 0.0001$). It can be seen that for each role deviations are biased in a certain phase of the standard. For senior residents, most deviations are made in Phase 2 (X-ray and Diagnostic Studies), whereas nurses make most deviations in care delivery. This indicates the shift in activity control between clinicians involved in trauma care. Experienced clinicians (senior residents and nurses) also show restraint in the phases in which they deviate. This supports the previous statement that expert clinician possess the knowledge base to deviate with the flexible portions of the guidelines alone.

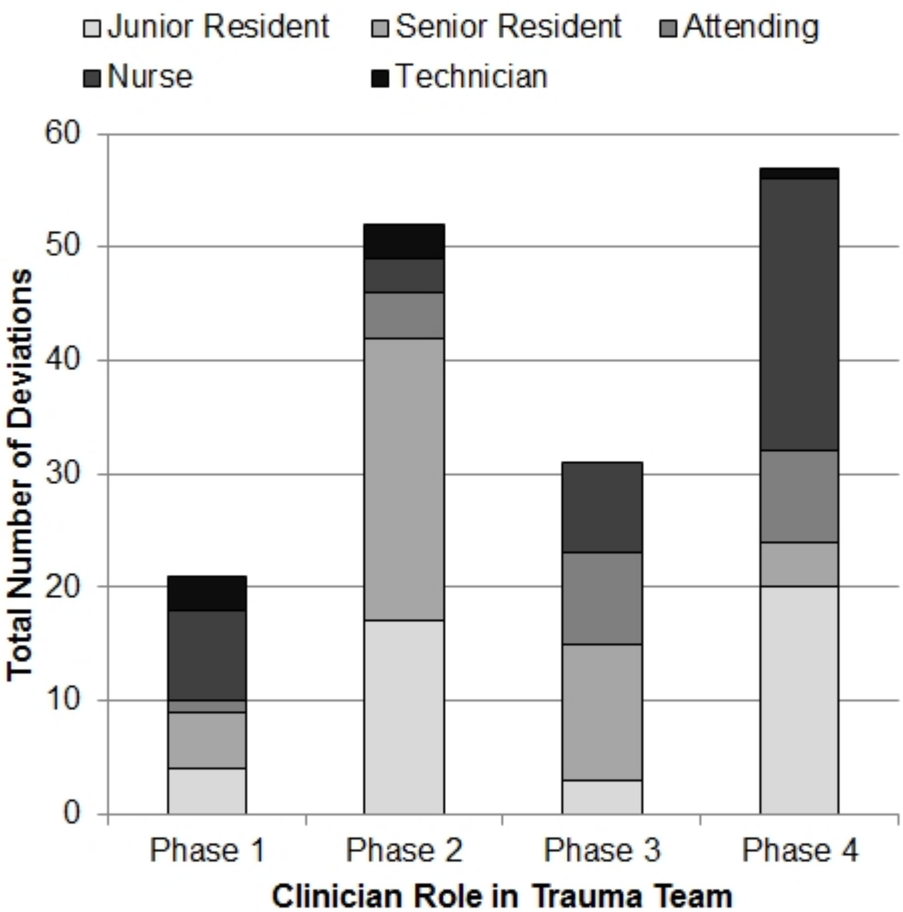

Figure 21. Deviations (classified by role) and phases of trauma standard

Deviations and Clinician Role in Trauma Team

Figure 22 and Figure 23 show the total number of deviations made by the individual members of the trauma team for classification schema 1 (Chi-sq = 65.7722, df = 12, $p < 0.0001$) and classification schema 2 (Chi-sq = 18.5554, df = 8, $p = 0.0174$). A statistically significant relationship was found between types of deviations and role played by the clinician in the trauma team. Expert clinicians made more innovations (attendings: 44.44%, PGY4/5 residents: 33.33%, trauma nurses: 14.81%) when compared to junior residents (7.41%). Junior residents on the other hand made a greater number of errors than any other group (63.64%). These findings substantiate the preliminary study on errors and innovations [62].

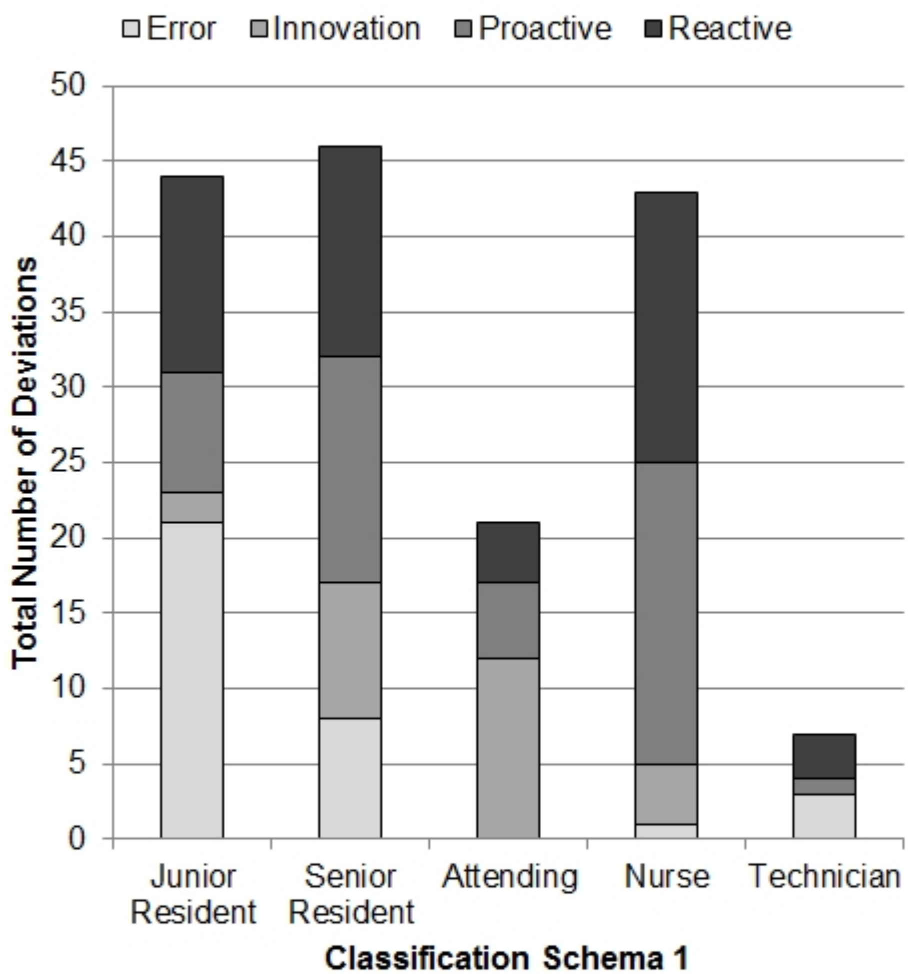

Figure 22. Deviations (classification schema 1) and clinician role

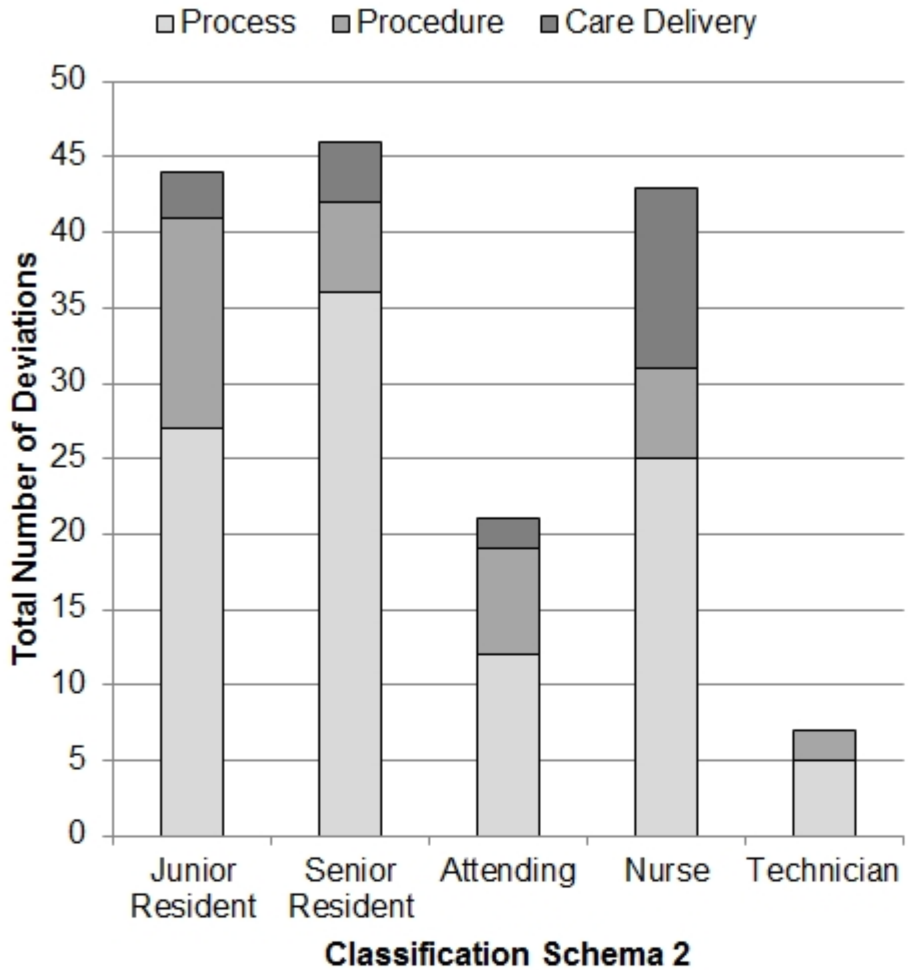

Figure 23. Deviations (classification schema 2) and clinician role

It can also be seen that most of the deviations performed by residents are process- and procedure-related. As mentioned earlier and corroborated by Figure 23, junior residents' deviations are more biased towards procedures. It is not unusual that deviations made by nurses are predominately care delivery-related. Trauma teams have well-defined role boundaries. This enables teams to function effectively in chaotic situations.

Expertise is critical to formation of adaptive teams in trauma critical care. The results show that trauma leaders with more experience are able to adapt to the dynamic environment while minimizing errors. Novices, on the other hand, are preoccupied by procedural aspects of trauma care and fail to achieve the necessary levels of communication needed to facilitate team innovations. Another key difference between experts and novices lies in their ability to recover from errors and unexpected events. Patel and colleagues [74] showed that experts' knowledge is adapted to recognize familiar patterns of stimuli. However, their heuristic reasoning from the pattern recognition strategy may not be effective in some complex situations [75]. Experts may make errors, but are adept at correcting them before negative consequences occur. Novices on the other hand fail to perceive the consequences of their decisions until it is too late [6, 66].

Relationship between Classification Schemas

A significant relationship was also found between classification schema 1 and schema 2 (Chi-sq = 25.9012, df = 6, p = 0.0002). Figure 24 depicts this relationship. A majority of the proactive deviations were process-related. Proactive deviations often involved task advancement. This could account for the observed relationship. Since the number of process-related deviations is high, it is difficult to assess the nuances of the relationships between other variables. However, one observation is that there are no care delivery-related errors. In order to assess the validity of this finding, further data collection to increase the sample size of trauma cases will be needed.

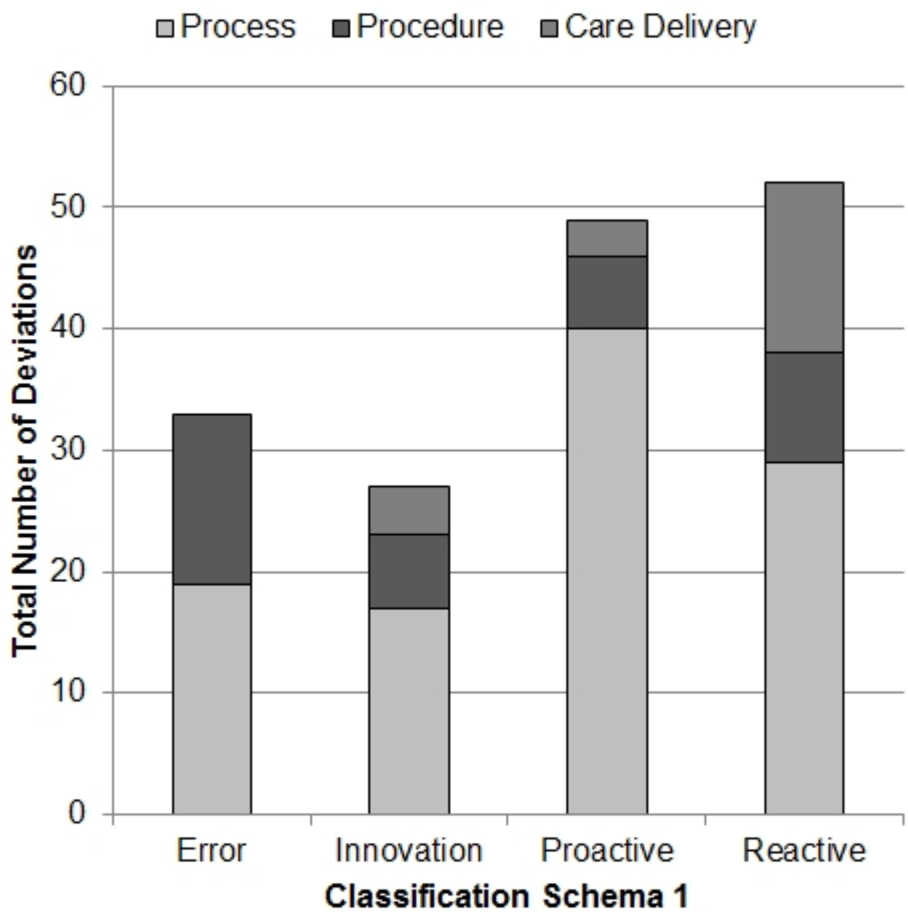

Figure 24. Deviations classified with schema 1 and 2

Study Limitations

This work attempts to provide definitions and structure to a subjective form of analysis. Classifying deviations using the methodology described (based purely on observations) is difficult since there may not be enough contextual information to make a concrete decision. Video recording of trauma cases or using data gathered using the hybrid framework described in this manuscript will enable capture of all the activities that take place in trauma care. This will especially be useful in cases where it is difficult to identify the clinicians involved in initiating the chain of events that resulted in a particular deviation.

In addition to data collection methods, the validity and generalizability of the classification schema will need to be assessed. Deviations will need to be classified by independent raters or coders. Sufficient inter-rater reliability will provide concrete support for the value of this work. In addition to validation, the current data set does not contain enough examples of team, procedure and care-delivery type deviations. In order to ascertain if this is because of low sample size or the absence of such deviations in general, more trauma cases will need to be observed. In the following chapter, research efforts to assess validity and completeness of the classification schema are described.

EVALUATING GENERALIZABILITY OF CLASSIFICATION SCHEMAS

This chapter describes two independent experiments conducted to assess the generalizability of the classification schemas presented. Based on the limitations described in the previous chapter, experiments are designed to (i) assess the replicability of the classification by independent raters, and (ii) concordance of their rating/coding with the original classification. The results of the experiments described will help guide future work in this domain.

Replicability of Classification Schemas

In previous work 30 trauma cases were observed in Banner Good Samaritan's Trauma Center. These observations were de-identified and utilized to develop the current classification schema. These observations were used in the experiment to assess the replicability of classification by other raters.

Methods

This study was approved by the Institutional Review Boards of Arizona State University and Banner Good Samaritan Medical Center. Fifteen trauma cases were randomly chosen from the existing pool of thirty trauma cases. Deviations from five of these cases were used for training two raters. The deviations in the remaining ten cases served as the test set. The raters chosen for this experiment had prior clinical environments experience (having spent 30 to 60 hours observing clinicians). Raters with experience were chosen due to the contextual nature of the task.

The training phases consisted of a PowerPoint® slideshow that provided a brief introduction to trauma critical care and the various classification schemas. Raters were then asked to code each deviation in the training set (a total of 17 deviations). After every classification, the answers from the current classification were present followed by a discussion about the deviation. Upon completion of the training phase, raters proceeded with the test.

In the test phase, raters were presented with deviations from the randomized test cases (a total of 38 deviations). For each deviation, raters marked the type of deviation for classification 1, 2 and 3. They were provided with a not applicable (N/A) option, if they were unsure of how to classify the deviation.

Among the 38 deviations, one rater marked N/A for one deviation. This sample was omitted from the analysis as an anomaly. Following the coding, the data were analyzed to assess (i) inter-rater agreement between the two raters, and (ii) concordance with existing classification though a similar agreement measure. A simple Cohen's Kappa statistic was used for the analysis. As the classification schema is not ordered, all categories were given the same weight (one).

There are a number of guidelines available for interpreting Kappa statistics. For example, Fleiss's [76] guidelines consider Kappa >0.75 as excellent, 0.40 to 0.75 as fair to good, and < 0.40 as poor agreement. Landis and Koch [77], on the other hand present a more granulated scale for measuring agreement. They consider Kappa values of 0.81 – 1.00 as almost perfect agreement, 0.61 – 0.81 as substantial, 0.41 – 0.6 as moderate, 0.21 – 0.40 as fair, 0.0 – 0.20 as slight agreement and <0 as poor agreement. Since the nature of the classification task is subjective, the scale proposed by Landis and Koch [77] is used to interpret the results of the Kappa tests performed. In the following section the results of this experiment are presented.

Results

Inter-rater Agreement for Classification Schemas

For each of the classification schemas (1: Error, Innovation, Proactive, and Reactive; 2: Process, Procedure, and Care Delivery, and 3: Individual and Team), the rating or classification provided by the two raters was analyzed using Cohen's Kappa. Table 7 summarizes the statistics for the inter-rater reliability test between Rater A and Rater B.

Table 7. Kappa statistics for test between Rater A and Rater B

Classification	Kappa	95% Lower Conf. Limit	95% Upper Conf. Limit
Schema 1	0.7743	0.6125	0.9361
Schema 2	0.7632	0.5079	1.0000
Schema 3	0.5355	0.1344	0.9366

There is substantial agreement for classification 1 and 2. However, there is moderate agreement for classification schema 3. One reason for this result could be the lack of sufficient examples of team deviations in the current data set. Another reason could be the difficulty in defining what constitutes a team deviation in trauma care. Take, for example, the case where the log roll step in trauma care is missed. One could argue that the trauma leader is responsible for how trauma care is conducted. Hence is it an individual error. On the other hand, there were a number of other team members who could have prevented the error. In that sense it could be a team error. Such a difficulty could be resolved by studying individual and team interactions further in trauma care.

The results of the inter-rater reliability test are promising. For classification schemas 1 and 2, the relatively high Kappa score indicates that independent raters can use the classification schema.

Concordance with Original Classification

Table 8 and Table 9 show the results of tests conducted between (i) Rater A and the original classification, and (ii) Rater B and the original classification. Rater A had very high (almost perfect) agreement with the original classification in all three schemas. Such high levels of agreement are unexpected. Rater B, on the other hand had substantial agreement for schema 1 and moderate agreement for schema 2 and 3.

Table 8. Kappa statistics for test between Rater A and original classification

Classification	Kappa	95% Lower Conf. Limit	95% Upper Conf. Limit
Schema 1	0.7743	0.6125	0.9361
Schema 2	0.7632	0.5079	1.0000
Schema 3	0.5355	0.1344	0.9366

Table 9. Kappa statistics for test between Rater B and original classification

Classification	Kappa	95% Lower Conf. Limit	95% Upper Conf. Limit
Schema 1	0.7729	0.6090	0.9367
Schema 2	0.5068	0.2252	0.7885
Schema 3	0.5355	0.1344	0.9366

These results indicate the natural differences between raters. The high agreement with rater A and moderate to substantial agreement with rater B validates the categories developed to assess deviations. Combined with the results of agreement between Rater A and Rater B, this indicates that the classification schema is replicable and can be effectively used by other researchers.

Study Limitations

The key limitation of this study is that raters obtain their contextual information from tertiary observations. The process of immersing oneself in an environment provides information about several nuances of behavior that may be completely missed in written observations. Reproducing the study with data from the hybrid framework or video recording of trauma cases will provide the raters with all the information they would need to make a classification. It is also possible that the Kappa scores will improve even further if raters were provided with comprehensive data.

Classifying deviations to understand cognitive decision-making processes is a very subjective process. One example from the test set is an attending asking a nurse if there is a tuberculosis protocol to follow, after it was discovered that the patient might be infected. The classification schema stated that it was a proactive deviation. Rater A marked it as an innovation and Rater B marked it as a reactive deviation. All three cases can be argued. It is a proactive deviation, since the attending went out of the bounds of his role in requesting the information (possibly in anticipation of steps to follow). It can be considered to be reactive, since it is a common task addition in reaction to the patient being infected (a random event). If thought of as a novel task addition that greatly improves patient and team safety, then it is an innovation. These arguments are based on (i) what the rater finds is accepted, or common behavior, and (ii) what they perceive the impact of the deviation might be. Prior to classification and analysis, researcher will need to develop a rubric for addressing these two factors.

DISCUSSION

The findings from the research described in this work have a number of scientific and practical implications to the field of biomedical informatics. In this section, the broad implications of this work are discussed under the themes of (i) protocols and guidelines in complex systems, (ii) expertise and innovation, and (iii) training clinicians in complex systems.

Protocols and Standards in Complex Systems

Protocols and standards are important for ensuring process consistency and patient safety in healthcare. It has been shown that linear systems and processes are aided by protocol and checklist deployment. For example, Pronovost and colleagues showed that implementation of a checklist for central line placement decreased the rate of catheter related blood infections from 2.7% to zero in the first three months of deployment [78]. Such protocols limit errors by reducing the workload on human memory and automating the care process [79]. Most critical care environments, however, are characterized by non-linear interactions and dynamic emergent behavior [80]. In such environments, clinicians need to make dynamic adjustments to protocols and guidelines, in order to adapt to the operational conditions and to achieve high accuracy and efficiency. The analysis of 30 trauma cases in this work showed that an average of 5.37 deviations occur during each case. Therefore, complex systems similar to trauma critical care, *cannot be treated as a zero-tolerance environments*. While protocols and guidelines serve to control complexity and errors through standardization, the importance of adapting standards safely to adjust to the environments needs to be recognized by clinicians and researchers alike.

Expertise and Innovation

Protocols and standards are based on observations and evidence gathered from practices. New information and novel findings from practice need to be incorporated into the guidelines and protocols. So how do such novel ideas get generated from practice? When regular or standard patterns do not fit or match the current problem, possible alternative ideas get generated. This is the process of innovation, and innovation is not possible without deviations. As practitioners gain experience in the execution of a task, their performance become increasingly smooth and efficient. While developing proficiency with attention-demanding complex tasks, some component skills become automatic, so that conscious processing can be devoted to reasoning and reflective thought with minimal interference in the overall performance. A great deal of experts' knowledge is finely tuned and highly automated enabling them to execute a set of procedures in an efficient manner. Yet they can perform such tasks in a highly adaptive manner, which is sensitive to shifting contexts.

The findings from this research showed that expert clinicians (senior residents and attending surgeons) do make errors. However, they are able to correct errors made before they result in a critical failure. The analysis of deviations also showed that the expertise of the trauma team leader impacted the types of deviations made. Expert teams were more innovative, compared to teams led by a novice resident. Not only are these finding consistent with emerging knowledge about medical errors and expertise [74], it also indicates that *expertise is critical to the formation of adaptive clinical teams*.

Training clinicians in Complex Systems

There is a strong need for informatics tools that will enable novices to adapt to the trauma environment in following certain standards, allowing for fewer errors. The classification of deviations could allow for a scientific framework for modification of protocols and enable protocol developers to leverage a data-driven approach to modifications. Currently available tools such as checklists and protocols need to allow for note takers to mark and document deviations, errors and innovation. In protocol-driven environments, checklists have been found to be a valuable tool in minimizing error rates. However, since experts' deviations are important for education and practice, these checklists would have to be flexible enough to be automatically updated. For a dynamic environment like trauma, these checklists when implemented would need to be adaptable as well. In order to develop such a tool, one would need to know the general decision process in trauma and the various types of deviations that may occur. Using the classification of deviations presented in this work, it may be possible to create such a checklist, one that is customized to the expertise and the role of the individuals in a trauma team.

In addition to supporting dynamic checklists, the classification schema can also enable the development of simulators driven by real-world data that provide training to maximize innovation and minimize error occurrence. Such an educational tool will be critical in developing decision-making skills of residents and caregivers. It would allow for a comprehensive evaluation of the skills of the caregivers as well as a means to train teams for not only adherence to a protocol but enabling recognition of circumstances where innovation is needed.

The classification schema developed is generic and can be utilized to study deviations in other environments where similar complexity is experienced. Such environments include emergency departments and intensive care units.

The recognition of deviations utilizing a schema that classifies deviations as errors, innovations and procedural deviations can significantly alter compliance procedures. In addition, such a classification can lay the foundations for an adaptive framework for the modification of existing protocols. For example, if deviations are consistently seen on a particular step in a protocol, then that step may have to be re-analyzed. Similarly if innovations are continuously seen and replicated in multiple sites, then it could be incorporated into the next version of a protocol. Therefore, the analysis of deviations as described in this work can help guide efforts to update existing protocols and guidelines in meeting the requirements of complex adaptive systems.

CONCLUSIONS

Clinicians deviate from protocols when managing patients. The studies discussed in this manuscript show that clinical teams in critical care environments make a significant number of deviations per case, and that not all deviations are errors. The study of these deviations can provide new insight into how teams operate in complex environments and what distinguishes experts from novices. The results are in coherence with existing literature on exploring the cognitive basis of clinical expertise. It can be hypothesized that existence of retrieval structures in experts and top-down information processing allows for time-critical thinking that supports innovation by experts. This is supplemented by the information filtering that the retrieval structures support. On the other hand, novices are driven by bottom-up reasoning mechanisms and, without retrieval structures and filtering, are overwhelmed by the data and often make errors. Although only further experimentation can investigate this hypothesis, the observations clearly point to the plausibility of such mechanisms.

An analysis of deviations can enable the building of models of expertise and workflow that can be then used to design the next generation of effective interventions. Interventions could be standardized communication tools, and uses of information technology that supports innovations by effective presentation of information and cognitive decision support through educational efforts such as simulations. Simulations offer an exciting means of teaching clinical caregivers to learn how to effectively innovate in complex environments. The Accreditation Council of Graduate Medical Education recognizes simulation as an effective means of promoting critical thinking, professionalism and clinical knowledge [81]. It is generally seen only as an effective means of promoting standardization and adherence to a protocol [82]. This study, however, shows that simulation should be used for teaching clinical caregivers the nuances of errors and innovations. Simulation offers a safe environment to achieve such goals. Simulations that are not just a means of achieving standardization but also help develop certain knowledge structure fairly quickly (through practice that would make any deviations safer) can be developed.

The data presented in this paper suggests that there is a strong link between innovations, errors and expertise. Expert caregivers deviate from the protocol almost as often as novices but make significantly more innovations. This seems to suggest that expert have a strong mental model of how and when to innovate and can employ their knowledge and application abilities to innovate on the fly. Such innovations and recognizing them should be an important part of clinical practice as it helps is redesigning protocols and procedures.

The next steps for this research include studies to explore in detail the underlying mechanisms of expertise and innovations in trauma. The methodologies described by Arocha and Patel [83] will be employed for these studies. Focusing on semantic analysis as a means of studying the innovations process in experts and novices will greatly add to the conclusions of this work. Semantic analysis will yield important insights into how information is assimilated and processed by clinical caregivers. This would be crucial in understanding how to develop novel protocols and standards. For example, given the seriality of information as it passes from working memory to long term memory [84], one may include markers within the case description that may invoke the correct knowledge structures in long-term memory that support creativity. Continuation of this research will enable testing such interventions (including simulations mentioned above) and evaluating the same.

REFERENCES

1 Reason J. Human Error. Cambridge, England: Cambridge University Press; 1990.

2 Leape LL, Woods DD, Hatlie MJ, Kizer KW, Schroeder SA, Lundberg GD. Promoting Patient Safety by Preventing Medical Error. Journal of American Medical Association (JAMA). 1998;**280**(16):1444-7.

3 Institute of Medicine (IOM), U.S. Committee on Quality of Health Care in America. To Err Is Human: Building a Safer Health System. USA: The National Academies Press; 2000.

4 American Hospital Association. Hospital Statistics. Chicago; 1999.

5 Consumers Union. To Err is Human – To Delay is Deadly. Available at: http://safepatientproject.org/safepatientproject.org/pdf/safepatientproject.org-ToDelayIsDeadly.pdf (Last Accessed on 20 April 2012); 2009.

6 Preventing Fatal Medical Errors. Editorial, New York Times. December 1, 1999.

7 Institute of Medicine (IOM), U.S. Committee on Quality of Health Care in America. Crossing the Quality Chasm : A New Health System for the 21st Century: National Academy Press; 2001.

8 Leape LL, Berwick DM. Five Years After To Err Is Human: What Have We Learned? Journal of American Medical Association (JAMA). 2005 May 18, 2005;**293**(19):2384-90.

9 Joint Commission on Accreditation of Healthcare Organization (JCAHO). Sentinel Event Trends: Potassium Chloride Events by Year. Available at: http://wwwjcahoorg/accredited+organizations/ambulatory+care/sentinel+events/set+potassiumhtm, (Last Accessed on April 2, 2005).

10 Kelly JJ, Sweigard KW, Shields K, Schneider D. John M. Eisenberg Patient Safety Awards. Safety, Effectiveness, and Efficiency: A Web-based Virtual Anticoagulation Clinic. Journal on Quality and Patient Safety. 2003;**29**(12):646-51.

11 Whittington J, Cohen H. OSF Healthcare's Journey in Patient Safety. Journal of Quality Management in Health Care. 2004;**13**(1):53-9.

12 Pronovost P, Needham D, Berenholtz S, et al. An Intervention to Decrease Catheter-Related Bloodstream Infections in the ICU. New England Journal of Medicine. 2006;**355**(26):2725-32.

13 Goldspink C. Modelling Social Systems as Complex: Towards a Social Simulation Meta-Model. Journal of Artificial Societies and Social Simulation. 2000;**3**(2):Available at: http://jasss.soc.surrey.ac.uk/3/2/1.html, (Last Accessed on April 20, 2012).

14 Plsek PE. Some Emerging Principles for Managers of Complex Adaptive Systems (CAS), (Working Paper). 1997.

15 Plsek PE, Greenhalgh T. The Challenge of Complexity in Health Care. British Medical Journal (BMJ). 2001;**323**:625-8.

16 Waldrop WM. Complexity: The Emerging Science at the Edge of Order and Chaos. New York: Simon and Schuster; 1992.

17 Holland J. Hidden Order: How Adaptation Builds Complexity. New York: Addison-Wesley; 1995.

18 Lorenz EN. The Essence of Chaos: CRC Press; 1995.

19 Sibthorpe B, Glasgow N, Longstaff D. Complex Adaptive Systems: A Different Way of Thinking About Health Care Systems. Australian Primary Health Care Research Institute. 2004.

20 Plsek PE, Greenhalgh T. The challenge of complexity in health care. British Medical Journal. 2001 September 15, 2001;**323**(7313):625-8.

21 Lewin R. Complexity: Life at the Edge of Chaos: University Of Chicago Press; 2000.

22 Horsky J, Gutnik L, Patel VL. Technology for Emergency Care: Cognitive and Workflow Considerations. American Medical Informatics Association (AMIA) Annual Symposium Proceedings; 2006; 2006. p. 344–8.

23 Malhotra S, Jordan D, Shortliffe E, Patel VL. Workflow Modeling in Critical Care: Piecing Together Your Own Puzzle. Journal of Biomedical Informatics (JBI). 2007;**40**(2):81-92.

24 Gilmour D, Woodwards S. The Surgical Safety Checklist: Legacy and Next Step. Journal of Preoperative Practice. 2010;**20**(3):78.

25 Duane TM, Brown H, Borchers CT, et al. A Central Venous Line Protocol Decreases Bloodstream Infections and Length of Stay in a Trauma Intensive Care Unit Population. Annals of Surgery. 2009;**75**(12):1166-70.

26 Semel ME, Resch S, Haynes AB, et al. Adopting A Surgical Safety Checklist Could Save Money And Improve The Quality Of Care In U.S. Hospitals. Health Affairs. 2010 September 1, 2010;**29**(9):1593-9.

27 Donchin Y, Gopher D, Olin M, et al. A Look Into the Nature and Causes of Human Errors in the Intensive Care Unit. Quality and Safety in Health Care. 2003;**12**(2):143–7.

28 Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of Medical Error British Medical Journal (BMJ). 2000;**320**:774-7.

29 Rothschild JM, Landrigan CP, Cronin JW, et al. The Critical Care Safety Study: The Incidence and Nature of Adverse Events and Serious Medical Errors in Intensive Care. Critical Care Medicine. 2005;**33**(8):1694-700.

30 Brennan TA. The Institute of Medicine Report on Medical Errors--Could It Do Harm? New England Journal of Medicine. 2000;**342**(15):1123-5.

31 Helmreich RL, Foushee HC. Why Crew Resource Management? Empirical and Theoretical Bases of Human Factors Training in Aviation. In: Wiener E, Kanki B, Helmreich R, editors. Cockpit Resource Management. San Diego: Academic Press; 1993. p. 3-45.

32 Helmreich RL. On Error Management: Lessons from Aviation. British Medical Journal (BMJ). 2000;**320**(7237):781-5.

33 Helmreich RL, Merritt AC, Wilhelm JA. The Evolution of Crew Resource Management in Commercial Aviation. International Journal of Aviation Psychology. 1999;**9**:19-32.

34 Langewiesche W. Fly by Wire: The Geese, the Glide, the Miracle on the Hudson. New York: Farrar, Straus and Giroux; 2009.

35 Wise J. What Really Happened Aboard Air France 447. Popular Mechanics. December 6, 2011.

36 Vankipuram M, Kahol K, Cohen T, Patel VL. Visualization and Analysis of Activities in Critical Care Environments. American Medical Informatics Association (AMIA) Annual Symposium Proceedings. 2009:662-6.

37 Vankipuram M, Kahol K, Cohen T, Patel VL. Toward Automated Workflow Analysis and Visualization in Clinical Environments". Journal of Biomedical Informatics (JBI). 2011;**44**(3):432-40.

38 Brixey JJ, Robinson DJ, Turley JP, Zhang J. The Roles of MDs and RNs as Initiators and Recipients of Interruptions in Workflow. International Journal of Medical Informatics. 2008;**In Press, Corrected Proof**.

39 Laxmisan A, Hakimzada F, Sayan OR, Green RA, Zhang J, Patel VL. The Multitasking Clinician: Decision-making and Cognitive Demand During and

After Team Handoffs in Emergency Care. International Journal of Medical Informatics. 2006;**76**(11-12):801-11.

40 Alwan M, Dalal S, Mack D, et al. Impact of Monitoring Technology in Assisted Living: Outcome Pilot. IEEE Transactions On Information Technology in Biomedicine. 2006;**10**(1):192-8.

41 Østbye T, Lobach DF, Cheesborough D, et al. Evaluation of an Infrared/Radiofrequency Equipment-Tracking System in a Tertiary Care Hospital. Journal of Medical Systems. 2003;**27**(4):367-80.

42 Stanford V. Using Pervasive Computing to Deliver Elder Care. IEEE Pervasive Computing. 2002;**1**(1):1536-268.

43 Matthai P, Kenneth PF, Mike P, et al. Inferring Activities from Interactions with Objects. IEEE Pervasive Computing. 2004;**3**(4):50-7.

44 Schrooyen F, Baert I, Truijen S, et al. Real Time Location System over WiFi in a Healthcare Environment. Journal on Information Technology in Healthcare. 2006;**4**(6):401-16.

45 Chen C, Liu C, Li Y, et al. Pervasive Observation Medicine: The Application of RFID to Improve Patient Safety in Observation Unit of Hospital Emergency Department: IOS Press; 2005.

46 Fry EA, Lenert LA. MASCAL: RFID Tracking of Patients, Staff and Equipment to Enhance Hospital Response to Mass Casualty Events. American Medical Informatics Association (AMIA) Annual Symposium Proceedings. 2005:261-5.

47 Hauptmann AG, Gao J, Yan R, Qi Y, Yang J, Wactlar HD. Automated Analysis of Nursing Home Observations. Pervasive Computing, IEEE. 2004;**3**(2):15-21.

48 Islam GS. An Indoor Location Finding and Fall Detection System Using Sensor Integration with RFID Network and Accelerometer Data. Tempe: Arizona State University; 2009.

49 Hughes RG. Patient Safety and Quality: An Evidence-Based Handbook for Nurses; 2008.

50 Salas E, Dickinson TL, Converse SA, Tannenbaum SI. Toward and Understanding of Team Performance and Training. In: Swezey RW, Salas E, (eds) Teams: Their Training and Performance. 1992(3–29).

51 American College of Surgeons. Advanced Trauma Life Support for Doctors. 7 ed. Chicago, IL: American College of Surgeons; 2004.

52 Collicott PE, Hughes I. Training in Advanced Trauma Life Support. Journal of the American Medical Association (JAMA). 1980 March 21, 1980;**243**(11):1156-9.

53 Sarcevic A, Burd RS. What's the story? Information needs of trauma teams. Proc of AMIA. 2008:641-5.

54 Pal J. The Value of the Glasgow Coma Scale and Injury Severity Score: Predicting Outcome in Multiple Trauma Patients with Head Injury. Journal of Trauma: Injury, Infection, and Critical Care. 1989;**29**(6):746-8.

55 National Assocation of EMS Physicians. Emergency Medical Dispatching. Prehospital and Disaster Medicine. 1989;**4**(2):163-6.

56 American Society for Testing and Materials. Standard Practice for Emergency Medical Dispatch. Annual Book of ASTM Standards. 1990.

57 Burns JP. Complexity Science and Leadership in Healthcare. Journal of Nursing Administration. 2001;**31**(10):474–82.

58 Zimmerman BJ, Lindberg C, Plsek PE. Edgeware: Insights from Complexity Science for Health Care Leaders. Irving, TX: VHA Publishing; 1998.

59 Plsek PE, Wilson T. Complexity, leadership, and management in healthcare organisations. BMJ. 2001 September 29, 2001;**323**(7315):746-9.

60 Kahol K, Smith M, Ferrara J, Tripathi P, Leibowitz A, Panchanathan S. Gesture Based Hand Movement Analysis and Haptic Feedback for Surgical Training [selected as best poster]. Poster presentation at the 14th Meeting of Medicine Meets Virtual RealityMedicine Meets Virtual Reality 14. 2006.

61 Rabiner LR. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE. 1989;**77**(2):257-86.

62 Kahol K, Vankipuram M, Patel VL, Smith ML. Deviations from Protocol in a Complex Trauma Environment: Errors or Innovations? Journal of Biomedical Informatics. 2011;**44**(3):425-31.

63 Vankipuram M, Ghaemmaghami V, Patel VL. Adaptive Behaviors of Experts in Following Standard Protocol in Trauma Management: Implications for Developing Flexible Guidelines. Submitted to AMIA Annual Symposium; 2012; 2012.

64 Patel VL, Groen GJ, Arocha JF. Medical expertise as a function of task difficulty. Memory and Cognition. 1990;**18**(4):394-406.

65 Groen G, Patel V. Medical Problem-Solving: Some Questionable Assumptions. Medical Education. 1985;**19**:95-100.

66 Patel V, Arocha J, Kaufman D. Diagnostic Reasoning and Medical Expertise. Psychology of Learning and Motivation: Advances in Research and Theory. 1994;**31**:187-252.

67 Ericsson K, Patel V, Kintsch W. How Experts' Adaptations to Representative Task Demands Account for the Expertise Effect in Memory Recall: Comment on Vicente and Wang (1998). Psychological Review. 2000:578-92.

68 Joseph G, Patel V. Domain Knowledge and Hypothesis Generation in Diagnostic Reasoning. Medical Decision Making. 1990;**10**:31-46.

69 Finke R, Ward T, Smith S. Creative Cognition: Theory, Research and Applications. Cambridge, MA: MIT Press; 1996.

70 Cannon-Bowers JA, Salas E, Converse S. Shared Mental Models in Expert Team Decision Making; 2001.

71 Salas E, Rosen MA, Burke CS, Nicholson D, Howse WR. Markers for Enhancing Team Cognition in Complex Environments: The Power of Team Performance Diagnosis. Aviation, Space, and Environmental Medicine. 2007;**78**(Supplement 1):B77-B85.

72 West MA, Wallace M. Innovation in Health Care Teams. European Journal of Social Psychology. 1991;**21**(4):303-15.

73 Anderson L, Rothstein P. Creativity and Innovation: Consumer Research and Scenario Building. In: Kahn BE, Luce MF, editors. Advances in Consumer Research; 2004: Association for Consumer Research; 2004. p. 747-52.

74 Patel VL, Cohen T, Murarka T, et al. Recovery at the Edge of Error: Debunking the Myth of the Infallible Expert. Journal of Biomedical Informatics (JBI). 2011;**44**(3):413-24.

75 Patel VL, Groen GJ, Arocha JF. Medical expertise as a function of task difficulty. Mem Cognition. 1990;**18**(4):394-406.

76 Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley; 1981.

77 Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Journal of Biometrics. 1977;**33**:159-74.

78 Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. New Engl J Med. 2006;**355**(26):2725-32.

79 The Institute of Medicine (IOM). To Err is Human: National Academies Press; 2000.

80 Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. Br Med J. 2001;**323**(7313):625-8.

81 Accreditation Council for Graduate Medical Education. ACGME Competencies: Suggested Best Methods for Evaluation. 2000 [cited 2007 December 20]; Available from: http://www.acgme.org/Outcome/assess/ToolTable.pdfaa

82 Smith CC, Huang GC, Newman LR, et al. Simulation Training and Its Effect on Long-Term Resident Performance in Central Venous Catheterization. Simulation in Healthcare.**5**(3):146-51 10.1097/SIH.0b013e3181dd9672.

83 Arocha JF, Wang D, Patel VL. Identifying Reasoning Strategies in Medical Decision Making: A Methodological Guide. Journal of Biomedical Informatics. 2005;**38**(2):154-71.

84 Groen G, Patel V. The Relationship between Comprehension and Reasoning in Medical Expertise. In: Chi M, Glaser R, Farr M, editors. The Nature of Expertise. Hillsdale, NJ: Erlbaum; 1988. p. 287-310.