NASA-CR-198027

p. 85

T · H · E OHIO STATE UNIVERSITY

Grounding Explanations in Evolving, Diagnostic Situations

Leila J. Johannesen, Richard I. Cook, and David D. Woods
Cognitive Systems Engineering Laboratory

(NASA-CR-198027) GROUNDING EXPLANATIONS IN EVOLVING, DIAGNOSTIC SITUATIONS Final Report (Ohio State Univ.) 85 p

N95-25335

Unclas

G3/63 0045760

NASA-Johnson Space Center
Houston, Texas 77058

Grant No. NAG9-390
Final Report
RF Project No. 760678/727376

December, 1994

GROUNDING EXPLANATIONS IN EVOLVING, DIAGNOSTIC SITUATIONS

Leila J. Johannesen, Richard I. Cook, and David D. Woods

Cognitive Systems Engineering Laboratory
The Ohio State University
Columbus, OH 43210

CSEL REPORT 1994-TR-03
December, 1994

Sponsored by NASA Johnson Space Center
under Grant NAG9-390

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES

PAGE

ACKNOWLEDGMENTS

This research was sponsored by NASA Johnson Space Center, Houston, TX. Our deep appreciation goes to Dr. Jane Malin, technical monitor, for her interest and support in our work.

Collection and analysis of the operating room data was made possible in part by grants from the Anesthesia Patient Safety Foundation and the Mandalfino Research Fund of the Ohio State University. The authors also acknowledge the generous assistance of the faculty and residents of the Departments of Anesthesiology and Neurosurgery of the Ohio State University, without whose cooperation the study would not have been possible.

EXECUTIVE SUMMARY

This research was motivated by the desire to further understanding on how artificial intelligence (AI) systems may effectively support practitioners engaged in fault management in dynamic situations. One standard approach for diagnostic assistance is to provide retrospective explanations, but these are not well suited to the demands of dynamic fault management. Such explanations occur as interruptions in the flow of work and result in data overload. A field study of human practitioners in one dynamic fault management application (anesthesiologists during neurosurgical operations) was undertaken in order to gain insight into effective diagnostic support among team members. The conceptual framework that guided the field study drew from research on cooperation in communication, and particularly on work from conversation analysis on the "common ground" maintained during coordinative activity.

The findings indicate that team members assist one another in maintaining accurate interpretations of the process by helping one another keep track of influences on the process. Two ways they do this are by providing unprompted reports of their relevant activities on the process and by providing informative responses that go beyond answering explicitly posed questions. Episodes of management and diagnosis show that causal explanations among team members are better described as joint interpretations (in which both team members are involved in the process of attaining a mutual interpretation), rather than as retrospective explanations given from one team member to another. Explanations of assessments and activities that are found are typically brief and embedded in the flow of activity.

The general implications for the design of intelligent systems intended to support practitioners in dynamic fault management are that such systems should not be "dark boards" concerning their activities and assessments. But, because they lack many of the sophisticated capabilities displayed in human communication, intelligent system design must avoid distracting their human partners in an effort to maintain the common ground. Instead, the focus should be on providing intelligent system assessments and information about activities in the context of (i.e., relative to) events in the dynamic process.

CHAPTER I
INTRODUCTION AND OVERVIEW

Introduction

Certain fields of practice involve the management and control of complex dynamic systems. These include flight deck operations in commercial aviation, control of space systems, anesthetic management during surgery or chemical or nuclear process control. Fault diagnosis of these dynamic systems generally must occur with the monitored process[1] on-line and in conjunction with maintaining system integrity.

Some of the demands of fault management in a dynamic process include: the need to form interpretations of the situation with faulty data or before all the data are available, the need to continuously update these interpretations as data comes in or is changed, and the need to act based on these interpretations (in order to prevent possible dire consequences or to find out more about possible faults, e.g., Woods, 1994). Artificial intelligence (AI) and automation is increasingly being used to assist practitioners who manage and control complex dynamic systems. One common role for AI systems is to function as advisors, presenting diagnoses and recommendations. In addition they may also function as intelligent subordinate agents taking actions on the monitored process, but supervised to some extent by humans. The nature of the interaction (or support paradigm) between humans and intelligent system (IS) is critical for the success of the joint human-intelligent system.

One problem with standard forms of assistance in dynamic fault management is a lack of coordination between AI explanations and the demands of evolving, diagnostic situations (Malin et al., 1991). Poor design of the interaction and coordination between automation and human have been implicated in problems in managing automated resources and in failures (Billings, 1991; Norman, 1990). In particular, forms of interaction that take the human out of the problem solving loop, while still leaving the person responsible for the problem's solution have been shown to be ineffective (Roth, Bennett and Woods, 1987; Billings, 1991). There is a need to create more cooperative intelligent systems-- systems that assist the human practitioner in the process of problem solving and decision making, rather

---

[1] We refer to the dynamic process that is monitored and managed as the "monitored process" in order to distinguish it from other possible processes.

than simply providing a solution or recommendation (e.g., Layton et al., 1994).[2]

This research seeks to understand in more detail what it means for an intelligent system to function cooperatively, or as a "team player" in complex, dynamic environments. The approach taken was to study human practitioners engaged in the management of a complex, dynamic process: anesthesiologists during neurosurgical operations. The investigation focused on understanding how team members cooperate in management and fault diagnosis and comparing this interaction to the situation with an AI system that provides diagnoses and explanations. Of particular concern was to study the ways in which practitioners support one another in keeping aware of relevant information concerning the state of the monitored process and of the problem solving process.

An Illustrative Example

In the expert system model for problem solving assistance, the AI system provides diagnoses and/or recommendations, typically accompanied with an explanation for the diagnosis or recommendation. The diagnosis provided by the system, is also a type of explanation; it is usually a causal explanation that relates observed symptoms to some underlying cause(s). The explanation of the diagnosis, on the other hand, is generally a description of how the system arrived at the diagnosis. Chandrasekaran, Tanner and Josephson (1989) refer to this distinction as "explanations of the world" and "explanations of decisions" respectively.[3] This distinction will be relevant throughout this report. We will often refer to the first as *interpretations* (rather than diagnoses, which has a connotation of completeness), and *explanations of interpretations* (or simply, explanations).

To motivate discussion of the problems with the standard expert system form of diagnostic explanation for dynamic fault management, we use the following example. Though the example is hypothetical, the characteristics and cognitive implications are based on actual systems (see Woods, 1988; Malin et al., 1991; Woods, 1994).

Imagine that you are able to observe a practitioner at work on managing some process. For illustrative purposes, you will also have access to the person's cognitive activities. What the process is, doesn't really matter: it might be chemical, nuclear, or even a physiological process like a patient undergoing surgery. The important aspects of the situation are that the process is dynamic, that it is important to avoid certain states, and that the

---

[2]The term practitioner is used as opposed to "operator" or "user" in order to emphasize the person's role as working within a field of practice, not simply interacting with a machine.

[3]There are many distinctions among explanations that one could make. For example, Chandrasekaran et al. (1989) distinguish the latter class into 1) explaining why certain decision were or were not made, 2) justifying system's compiled knowledge by linking it to deep knowledge from which it was derived, and 3) explaining control behavior and problem solving strategy.

process cannot be taken off-line while diagnosis proceeds. Information about the state of the process is available via a monitor that displays raw parameter values. There is also an expert system, which receives data directly from the process, and provides assessments on another monitor to the practitioner. The expert system's main display has three main windows that provide: status messages, diagnoses and explanations of the diagnoses. Now, consider the following hypothetical scenario.

The practitioner has been monitoring the process display when he notices an unexpected anomaly in a process parameter. Because the anomaly (deviation from a standard condition) is unexpected, he[4] begins diagnosis-- thinking about what fault or faults could account for the anomaly. In order to better assess the situation, he calls up a few other windows on the process display that contain some relevant parameter values. From the raw data in the different windows he pieces together a picture of what may be going on-- two hypotheses jump to mind. He glances over to the expert system but its diagnosis window is blank. Shifting attention back to the process display, he notices that the initially anomalous parameter is continuing its trend. He also notices that another parameter is anomalous. He quickly considers the safing actions[5] that should be taken, and then goes back to the expert system display and retrieves the automated subsystem command window. He gives instructions for it to execute the safing action. He turns back to the process display in order to evaluate the action on the process.

The safing action has both therapeutic and diagnostic value. Seeing its effect on the process leads the practitioner to investigate another set of parameter values. In order to see potential patterns, he needs to call up several windows at once. As he begins to do this, he notices an anomalous trend; another hypothesis springs to mind--one with more severe consequences than the two that initially occurred to him. He searches through the interface to gather more information. Suddenly the expert system display beeps. He makes a note to look at it as soon as he finds the relevant data he is looking for. When he does, he looks away from the process data screen to see what hypothesis has been posted in the diagnostic window. It is one of the lower priority ones he initially considered. He re-checks the current parameter values in the process display. He wonders whether the expert system has taken into account the other, more dire, possibility. He looks back at the explanation window on the ES display to try to understand the data used and the reasoning followed in arriving at its diagnosis. The explanation scrolls beyond the window. He begins to read it, but then turns to the process display to check on how things are progressing, and to determine what the next

[4]Note that the male pronoun is used only for simplicity's sake.
[5]Actions to place the system in a "safe" configuration.

management action should be. As he looks at the raw data, he realizes a new event occurred. Was this due to an action taken by the automated system? He turns back to the expert system to check the status messages; there is a long list of messages. He searches and scrolls among the messages to find one that might explain the current pattern. Meanwhile, data on the process display are changing...

"Clumsy" Explanation

This example attempts to give a sense of the flow of cognitive activities involved in managing disturbances.[6] A model of the cognitive activities in disturbance management is shown in Figure 1 (from Woods et al., 1991). The first step is for the practitioner to recognize, out of all the potentially changing indications, which are anomalies. Recognizing one or more anomalies leads to cognitive activity about how to cope with the disturbance(s), i.e., what safing responses to execute. When these actions are taken, the practitioner needs to monitor to see if the responses have occurred as expected and whether they are having the desired effect. If unexpected anomalies are, diagnostic cognitive activity is triggered. When a diagnosis is reached, the lines of reasoning concern developing corrective responses. Note how the practitioner in the example needs to cope with the consequences of faults and perform fault diagnosis in parallel (Woods, 1988; 1994).

The example illustrates a lack of coordination for joint problem solving that can exist between a human practitioner and an "intelligent system." The first problem has to do with the time period during which the expert system's diagnosis and explanation arrives. A retrospective explanation typically comes at a time when the practitioner is likely to be engaged in multiple activities which may include several of the following: evaluating hypotheses, dealing with a new event or the consequences of the fault(s), planning corrective actions or monitoring for the effects of interventions, and attempting to differentiate the influences due to faults and those due to corrective actions. The expert system does not coordinate its attention direction behavior with the activities of others by judging interruptibility the way people do (Pavard et al., 1989). In effect, the system's message occurs as a potential interruption to these on-going lines of reasoning and monitoring.

---

6 Disturbances are abnormal conditions in which process state deviates from the desired function for the particular operating context.

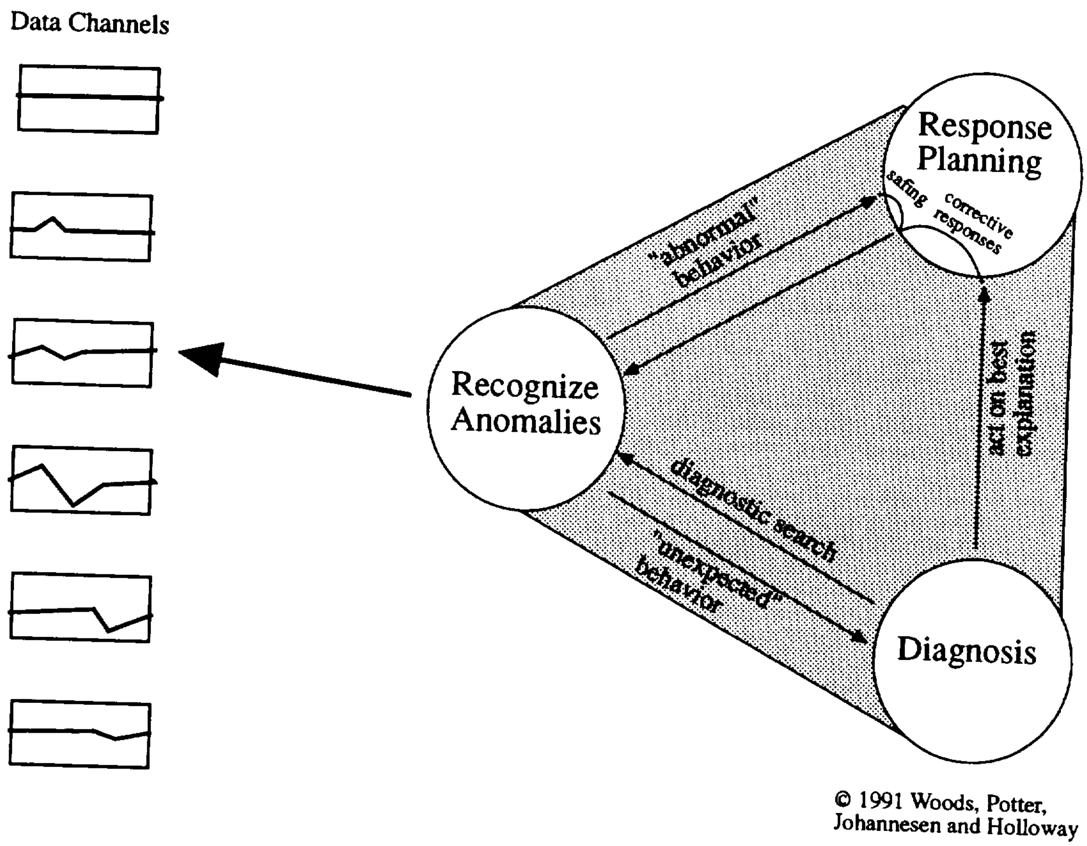

Figure 1. Model of anomaly-driven information processing in dynamic fault management.

A second problem concerns the amount of data present in the diagnosis/explanation. A long explanation requires the practitioner to pay attention to the expert system for some time, demanding cognitive resources to follow the reasoning. In addition, interface management tasks may be required such as scrolling, in order to read the complete message, or such as searching among other windows for relevant information in order to understand the explanation.

A problem that relates to both of the above results from a dissociation of the diagnosis/explanation from the process views. This is a problem because it means that establishing relationships among the process data and the explanation needs to be undertaken as extra tasks by the practitioner. Also, it means that examining the diagnosis and explanation will take the practitioner away from what is currently going on in the process, possibly resulting in missed events. In general, the dynamics and form of this expert system's explanation make it a "clumsy" type form of interaction. This term is based on "clumsy automation," a phrase coined by Earl Wiener to refer to automation that provides its benefits during low workload periods, but creates new burdens during high workload periods (Wiener, 1989). A clumsy explanation style is one that creates extra workload during high tempo periods. It is a form of data overload; it interrupts the practitioner at a busy time and creates extra tasks and cognitive burdens. Figure 2 illustrates this concept.

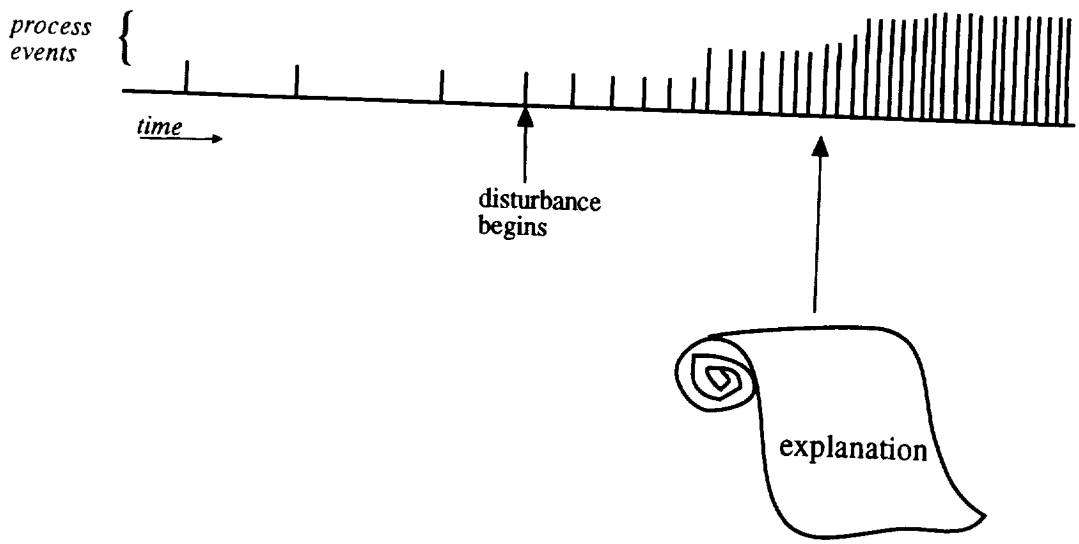

Figure 2. "Clumsy Explanation."

Feedback

In the illustrative example used earlier, the practitioner was unsure about whether the intelligent system had taken an action on the process. Effective feedback on the automation's activities is critical in dynamic fault management applications because diagnosis may involve disentangling the many different influences on the process, some of which may be due to actions taken by other agents (Woods, 1994). "Strong, silent" types, both among humans and automation are not team players (Foushee and Manos, 1981; Norman, 1989; Malin et al., 1991). Studies of the flight management system in the cockpit indicate that inadequate feedback can lead to difficulties in anticipating and knowing when uncommanded mode changes have occurred. Ineffective feedback can increase cognitive workload by increasing the demands on pilots to remember information about how the system functions (Sarter and Woods, 1992; 1994). Inadequate feedback on automation's activities also contributes to outcome failures; it has been a factor in aviation incidents.[7]

[7] As an example, consider the China Airlines 747 incident: the autopilot silently compensated for an engine's loss of power until it reached its limits, whereupon the unaware crew was forced to deal with the situation.

Some Characteristics of Typical Expert System Explanations and New Directions

Recent AI research has pointed to some of the problems associated with typical expert system explanations. Many of these criticisms and subsequent developments in the research labs have been inspired by studies of human explanations (e.g., Pollack et al., 1982; Cawsey, 1992; Paris, 1988). Much of the work in explanation concerns developing better explanations for diagnostic systems (e.g., Hasling, Clancey and Rennels, 1984; Chandrasekaran, 1989) or better explanations for how some device works for tutoring purposes (e.g., Cawsey, 1992, Feiner and McKeown, 1990; Suthers et al., 1992). These task domains have important characteristics that distinguish them from process control: 1) the underlying system is static and unchanging, and 2) time is not a significant factor. It is worth noting that there seems to be no research that investigates what constitutes effective explanations in *dynamic* fault management applications. We can expect the characteristics of the task demands to impact cognitive functioning and hence to impact the nature of a cognitive support system. In particular, in this work we are concerned with how the demands of process control impact the nature of effective AI diagnostic support.

Below we summarize some of the characteristics of typical explanations that are problematic for dynamic fault management. We also point to AI research that has also indicated the limitations of these characteristics for other domains.

- <u>Explanation as Retrospective.</u> Expert system explanations are typically given at the end of some problem solving activity (Cawsey, Galliers, Reece and Jones, 1992). These authors point out that, by contrast, when people are engaged in collaborative problem solving, they tend to provide justifications of their beliefs or reasoning as part of the problem solving.

In dynamic fault management, practitioners often cannot wait for a full-fledged diagnosis before taking corrective actions; though the picture of the underlying fault(s) may become clearer with time, the ability to recover or safe the system will tend to decrease with time (Perrow, 1984). Hence, some situations demand that management actions be taken before all the data becomes available or with uncertain data. Some understanding of the state and evolution of the process is needed in order to move the situation towards stabilization. This suggests that there should be collaboration in the diagnosis *process*, that is, with hypothesis revision and refinement over time (Malin, et al., 1991).

The plane went into an uncontrolled roll and lost thousands of feet in altitude before the crew recovered the aircraft.

In addition, keeping the operator out of the loop on the interpretation process of the AI system may result in a higher workload because it will mean taking the practitioner away from the monitored process at an inappropriate time in order to read a long explanation and to evaluate its soundness.

• Explanation as a One-Shot Process from Explainer to Explainee. A related aspect of typical AI explanations is that they are provided in "one shot" or as one long chunk (Moore and Swartout, 1989; Cawsey, 1991, Mastaglio and Reeves, 1990). By contrast, empirical studies of how people provide explanations show that it often requires a dialogue consisting of, for example, clarifications or elaborations (Moore and Swartout, 1990; Cawsey, 1991). In a sense then, an explanation is a negotiation process, in which the explainee is involved in the shaping of the explanation (Pollack et al., 1982). Recent research directions include how to make machine explanation interactive (e.g., Cawsey, 1991; Moore and Swartout, 1990; Suthers, Woolf and Cornell, 1992).

Mastaglio and Reeves (1990) point out that the one-shot approach puts too much into an explanation. They draw a parallel with this type of explanation and the *man page* summaries provided by the Unix system that attempt to provide everything a user might want to know about a command. By contrast people provide "minimal" or tailored explanations--providing just the information needed (Mastaglio and Reeves, 1990; Grice, 1975).

The man page approach to explanation results in data overload; it forces the user to search for and extract the relevant information, and there is no guarantee that the information the user needs is even there. This is particularly problematic in domains of practice in which time constraints are present, and in which the consequences of erroneous actions can be severe.

• Explanation as Response. Explanations can be thought of (at least implicitly) as a response to some query (e.g., "how..?" "why...?"). This is taken quite literally in some systems that provide interpretations or explanations for some process disturbances only when explicitly requested by the operator (Malin et al., 1991, vol 2). By contrast, people often provide unprompted or "spontaneous" explanations (Karsenty and Falzon, 1992). A problem with providing explanations only when queried is that people may not know quite what question to ask (what information they seek), or in the extreme case, may not even know that they should be asking for an explanation (that they need some information). This is particularly relevant in complex dynamic situations in which several agents may be involved in managing the process and in which situations can evolve quickly.

• Explanation as given by Expert to Novice. A premise behind the typical explanations given by expert systems is that the system is the authority on what is being explained and provides the person with a final answer. This perspective is inappropriate for supporting cooperative fault management in which practitioners are highly knowledgeable about the monitored process and are active in its management. Rather than substituting for expertise, many researchers believe that a more appropriate interaction paradigm is to support practitioner expertise (Roth, Bennett and Woods, 1987; Billings, 1991). Insight may be gained from studying collaborating team members with different levels of experience and expertise, and who have overlapping expertise.

• Explanation as Context-Independent. Traditionally, the goodness of an explanation has been seen as context-independent, i.e., that there is a valid, best explanation for some event (Leake, 1991). Leake points out that in this approach, determining an explanation relies on certain criteria such as explanatory coherence or testability, completeness, and brevity. Work on explanation in philosophy and psychology (Hilton, 1990) and more recently in AI (Leake, 1991, Cawsey, 1991; Moore and Swartout, 1990; Suthers, Woolf and Cornell, 1992) indicates that effective explanations need to take into account the goals and information needs of the explainee.

Another sense in which typical AI explanations are context-independent is that they are dissociated from the process views. For example, it is common to present IS interpretations and explanations in a separate "message list" window, i.e., a list of text messages (Malin et al., 1991). Message lists often fail to highlight monitored process events and relationships (e.g. temporal, causal) among these events (Potter and Woods, 1991). This lack of coordination among the IS assessments and the process view imposes the integration task on the practitioners. In such cases, practitioners simply have ignored the intelligent system (Remington and Shafto, 1990).

Diagnostic support systems for dynamic fault management that provide explanations with the above characteristics for dynamic fault management are likely to lead to "clumsy" interactions (Malin et al., 1991; Remington and Shafto, 1990). Woods (1989) points out that disturbance analysis systems developed in the nuclear industry in the 1980's, failed for the same basic reason. Though these systems did not incorporate AI, they attempted to automate diagnoses. The actual result was clumsy explanation and data overload-- operators had to sort through the diagnostic information, interpret it, and integrate it with other sources, all at the same time that the demands for monitoring, assessing and responding to events in the monitored process were the highest. Similarly other studies have found that technology purported to aid practitioners in fact created new cognitive burdens during critical and high tempo portions of the task (Woods, 1993a).

Overview of Approach and Scope

The research undertaken was an empirical investigation of how practitioners in one dynamic fault management domain (anesthesiology) support one another in management and diagnosis. The general approach is one of studying human-human interaction to provide insight to aid the design of human-computer interaction. (For another example of this approach see Coombs and Alty, 1980). The assumption is that by examining experienced practitioners, one can form higher level principles that are applicable to the human-intelligent system case. The basic logic of the research described in this report is shown in Figure 3.

The study performed is a field study because we observed and analyzed the performance of practitioners engaged in actual work, without taking any interventions. The study is what might be called a "focused field study" as opposed to a cognitive task analysis of anesthesiology (see Xiao, 1994 for this kind of field study of anesthesiologists). The study described herein focuses on one particular aspect-- how practitioners cooperate in supporting one anothers' situation assessment in dynamic fault management.

Contributions

This research begins to fill a gap in our understanding of how practitioners support one another in dynamic fault management. Data is provided on the kinds of explanations seen among practitioners and the information exchanges seen during episodes of management and diagnosis and episodes of updating. By applying the conceptual framework of the common ground to information exchanges among team members in dynamic fault management, a new interpretation of the nature and function of explanations in these applications is possible. The study indicates some ways that human team members maintain the common ground. Based on the study and constraints of current AI technology, implications are drawn for the design of intelligent systems intended to support practitioners as "team players" in dynamic fault management.

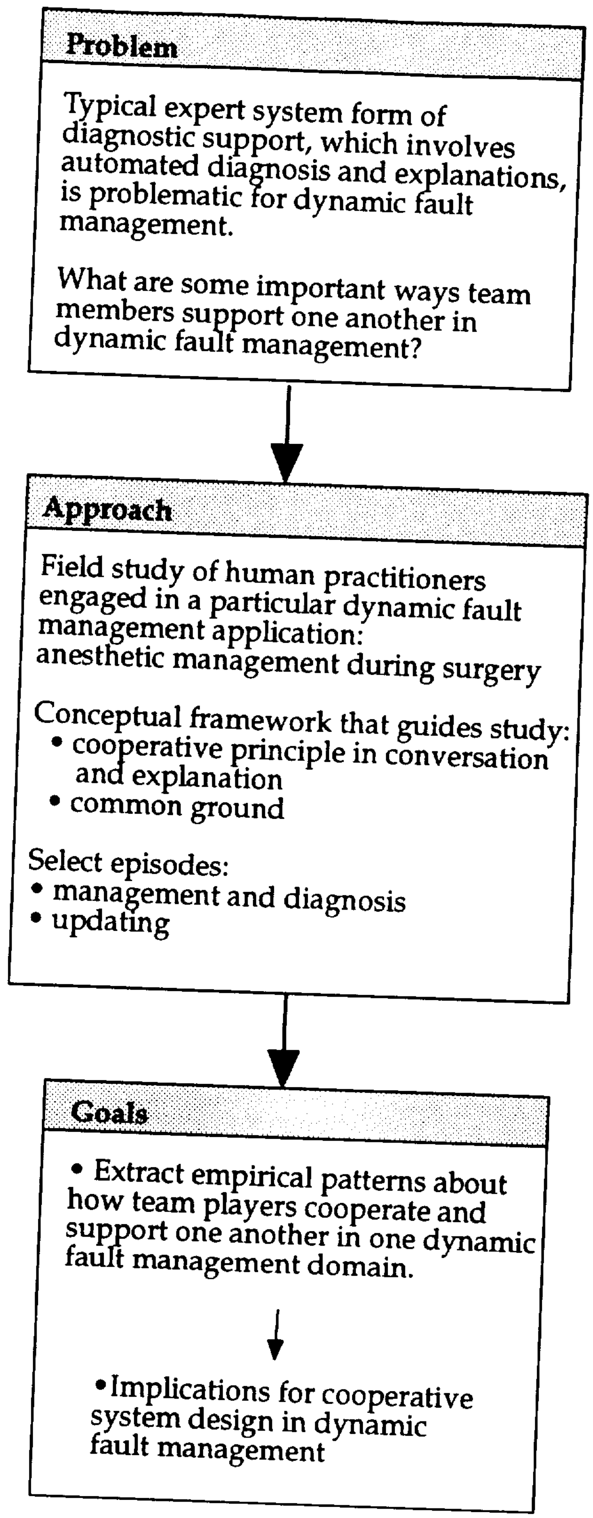

Figure 3. Basic Logic of the Research.

Overview of the Rest of the Report

Chapter 2 presents the concept base (or conceptual framework) used to analyze the raw data. This framework consists of concepts that highlight the cooperative nature of communication, i.e., the context sensitivity of explanations and the notion of a "common ground." Chapter 3 presents the research strategy of the study. Chapter 4 presents the findings and Chapter 5 discusses these findings.

CHAPTER II
FRAMEWORK FOR ANALYSIS

The challenge of field studies is to extract valid, generalizable patterns from the study of complex behavioral situations (Woods, 1993b). The basic analytic process involves taking a description of the actual performance (raw transcript), which is concept-free and highly context-dependent and converting it to a description that is concept-specific and domain-independent and which can generalize to other domains. In order to do this, one needs to have a framework or "conceptual looking glass" that guides the concepts of importance (Hollnagel et al., 1981; Woods, 1993b).

The problems with the standard forms of diagnostic support and explanation capabilities can be thought of, very generally, as communication problems. Therefore, the conceptual framework used for the field study draws on research and theory from different areas relating to effective communication and explanation. The major concepts of the framework are: conversation and explanation as a cooperative endeavor and the "common ground" that is maintained during communication. This framework makes explicit some concepts that seem to be relevant to understanding communication and cooperation in distributed problem solving.

Cooperative Communication

Grice (1975) pointed out that conversation is cooperative. In order for interlocutors to recognize the intentions (illocutionary acts) behind utterances, speakers need to adhere to certain principles or maxims. Grice referred to these as maxims of:

- *manner*: speak as clearly as possible, avoid ambiguity.
- *quantity*: provide as much information as needed in a context but not more
- *quality*: speak true information
- *relation*: make your contribution relevant to the context in which you are speaking

Adhering to these maxims implies that speakers need to take into account information about their listeners when formulating messages. People tailor their communications to what they perceive to be the information needs of their listeners (e.g., Krauss and Fussell, 1991).

Grice's maxims, when applied to human-computer communication, allow one to anticipate problems in interaction. Woods (1992a) illustrates how they apply to alarm systems. For example, inscrutable alarm messages (e.g., "Error code: 22345") violate the maxim of manner; they do not express clearly what the problem is. Group alarms, in which the same alarm indicates several different underlying problems, violate the maxim of quantity. There are cognitive consequences of violating the maxims of cooperative communication; for example, in the case of the group alarm, the person will need to seek out the missing information to understand what event in fact has actually occurred perhaps during a high tempo period.

The Cooperative Nature of Causal Explanations

Any particular event may be "explained" in different and valid ways. There are potentially many contributing factors-- each necessary but only jointly sufficient to produce the event that is to be explained. We often speak as if there was a single cause for the event, but we are selecting one of the multiple necessary factors to focus on. Which of these multiple contributors we select depends on the purposes of our inquiry.

Causal explanations always have an assumed contrast case, a "rather than" built into them (Hilton, 1990). For example, an explanation to the question "why did x occur?" contains some counterfactual contrast case, even though it may remain unarticulated. The complete question, if articulated, could be one of several, such as: "why x *rather than not x?"*, "why x *rather than the default value for x?"* or "why x *rather than y?"* The factors that are seen as mere conditions and those that are viewed as causal will vary depending on the frame of reference adopted, i.e., on the contrast case. Rather than explaining an event per se, one explains why the event occurs in the target case and not in some counterfactual contrast case. What is taken to be the cause depends on the causal counterfactual contrast case (Mackie, 1965).

According to Hilton (1990) causal explanations are like other forms of dialogue in the sense that they adhere to the general rules of (cooperative) conversation, or Gricean maxims (Hilton, 1990). To be relevant, explanations must close a gap in the explainee's knowledge concerning the issue in question (Turnbull, 1986). Some conditions are background (mere conditions) and some are foreground (causes) relative to the purposes of the explanation, which generally depends on the information needs of the explainee. The relevant explanation is one that "refers to the factor that makes the difference from the questioner's point of view between the target event and the backgrounded contrast case" (Hilton, 1990). Another way to say it is that some conditions are taken as shared or common knowledge, whereas the explanation itself focuses on the factor(s) that the explainer believes the explainee needs to know. Therefore, the goodness of an explanation depends on whether it provides the information that is needed to satisfy some goal (Leake, 1991).

Studies indicate that people provide explanations that attempt to close the gap in someone's knowledge and furthermore that they will change the explanations they give depending on the knowledge that they perceive to be shared among themselves and the explainee (see Turnball and Slugoski, 1988 for an overview). So if explainer and explainee share knowledge of some contributing factor x, then factor y which was also involved but which the explainer thinks the explainee does not know, will be emphasized. These results are interpreted to mean that explainers act cooperatively, complying with Gricean maxims when they explain.

Mutual Knowledge for Explanation

How does the explainer know about the knowledge, intentions, goals of the explainee that may be relevant to the explanation? One way is that explainees provide information about what they need to know in their queries. In studying advisory interactions, Aaronson and Carroll (1987) found that verification requests (utterances that effectively ask "is this idea correct...") are a good way to communicate one's understanding and for eliciting the explanation type and level. Similarly, Pollack et al. (1982) observed that people seeking advice actively participate in the definition and resolution of the problem; they may offer information, in order to gain some assurance that the advisor has used all facts. Linguistic means are also available for conveying what should be taken as foreground and background for an explanation. The explainee's question can be used as a cue for determining the contrast case and causal background (Haviland and Clark, 1974). For example, some part of an utterance can be foregrounded by emphasis (e.g., "Why did BILL go to the store?" implies that the requester wants to know why Bill went as opposed to other people who could have gone). Also, a sentence like "why isn't John home yet?" conveys the explainee's expectation that John should have been home already.

Another way to figure out the implicit contrast case for an explanation is to use mutual knowledge, goals and expectations. People engaged in collaborative activity have mental models of what is mutually known (Lewis, 1969; Turnball and Slugoski, 1988). This mutual knowledge is used in giving explanations, just as it is used in conversations in general (Searle, 1992). This mutual knowledge can be built up by having undergone some relevant shared experience, or taking stock of what is in the "cognitive environment"--the facts and assumptions that are capable of being perceived or inferred (Sperber and Wilson, 1986). Collaborating on some task provides a shared experience that informs participants about mutual goals and expectations.

Grounding in Human-Human Communication

Communication, being a collaborative activity relies on a "common ground"--the set of beliefs and presuppositions that each participant assumes are held by both, i.e., what they take to be their mutual knowledge and mutual beliefs (Stalnaker, 1978; Clark and Schaeffer, 1989). How common ground is built up, its role and functions has been studied mainly in conversation. During the course of the conversation, participants attempt to establish whether their utterances have been understood well enough for their current purposes. An utterance or, more precisely its meaning, is added to the common ground if it is accepted. Acceptance of an utterance occurs when interlocutors provide evidence of understanding; some ways people do this in conversation are by: 1) using acknowledgments such as "uh-huh" or nodding, and assessments (e.g., "gosh," "really"), 2) supplying a relevant next turn, i.e., an appropriate response which gives evidence that a participant has understood the utterance, 3) continued attention, and 4) demonstrating all or part of what has been understood or repeating verbatim.

Participants may actively seek positive evidence of understanding (Clark and Brennan, 1991). One way this is done is via the try marker, a rising intonation followed by a pause (Sacks and Schegloff, 1979). Try markers allow for the confirmation that a particular part of an utterance or reference has been understood, or allow the opportunity for correction.

Non-acceptance of an utterance leads to repair. In everyday conversations, repair is almost always initiated on the spot and is completed quickly.

Factors Affecting Grounding

Grounding is affected by the purposes and the medium of communication (Clark and Brennan, 1991). Different techniques for grounding may be required depending on the purposes of the conversation. Task-oriented dialogues may require that the criterion for grounding be higher. For example, if an important, complicated piece of information needs to be imparted, the speaker may present it in installments, expecting his or her partner to respond or sometimes repeat verbatim after each installment (Clark and Schaeffer, 1989).

The medium of communication affects the effort required to ground. For example, the acknowledgment "okay" is easy in face to face or telephone conversations, but in keyboard communication (e.g., via the full duplex Unix™ 'talk' facility) it is difficult to time it precisely so as not to interrupt the other typist. Hence, the cost is higher for using this kind of acknowledgment in keyboard communications. Some of the costs of grounding include: formation, production, reception, understanding, start-up (of new discourse), delay, asynchrony, speaker change, display, fault (utterance mistakes), and repair (Clark and Brennan, 1991). For example, delay costs are high in face-to-face or cotemporal and simultaneous media, e.g., long pauses are interpreted in various ways, or forgetting may occur. According to Clark and Wilkes-Gibbs (1986), people operate via the Principle of Least Collaborative Effort. That is, they try to use as little combined effort as possible. Their prediction is that people should ground with those techniques available in a medium that lead to the least collaborative effort.[8] Some aspects of the situation

[8] A study by Cohen (1984) demonstrates the way the medium affects grounding in a way that is consistent with this principle. In his study, tutors explained to a partner (student) in another room, how to put together a pump. The tutor, but not the student, had the instructions. When communicating over the phone the tutors tended to get their students to first identify an object and only when they had confirmed its identification did they ask the student to do something with it. In contrast, in keyboard conversations, tutors would identify an object and instruct students on what to do with it all in a single turn. This result is interpreted as being due to the different grounding costs in the two media; repairs are more costly over keyboard.

(constraints) that affect the grounding process in conversations between two people, A and B, are (from Clark and Brennan, 1991):

- copresence (A and B share the same physical environment),
- visibility (A and B are visible to each other),
- audibility (A and B can hear each other),
- cotemporality (B receives a communication at the same time as A produces it),
- simultaneity (A and B can send and receive communications at the same time),
- sequentiality (A and B's turns are not de-sequenced by external intrusions),
- reviewability (B can review A's communications),
- revisability (A can revise B's messages).

In this theoretical framework, the work required to ground in different media varies because the media vary on these grounding constraints. For example, reference is less costly in media that allow for copresence and visibility, where cost is measured in terms such as production, reception and understanding.

Grounding in Cooperative Problem Solving

Not only is grounding necessary for conversation; all coordinative activity among agents requires moment to moment updating of the common ground (Clark and Brennan, 1991). Of particular interest for the purposes of this work is how team members establish common ground necessary for effective joint problem solving.

Some work has investigated the role of copresence and visibility in establishing common ground within the context of performing some task. Findings by Grosz (1981) and McCarthy, Miles and Monk (1991), showing grounding difficulties in the absence of a shared visual field, are consistent with the theoretical framework outlined in the previous section. Grosz (1981) found that the absence of a shared visual field for an expert and apprentice engaged in a disassembly task led to confusions about common referents. Grosz (1981) points out that one way misunderstandings arise is that participants can think they have established common ground when they actually have not. McCarthy et al. (1991) found that when participants who communicated via text to solve a problem (concerning the layout of a bank) and lacked a shared visual space, they experienced grounding difficulties. Grounding was better among participants who had a shared report space in which solutions and arguments for these solutions could be jointly posted and seen.[9] The authors postulate that the effect could be due to several mechanisms, such as "the value of the public report as a shared memory aid;

[9]Grounding difficulty was measured by the disagreement in recalling the solutions and arguments between the two members of the problem solving team. Grounding of the arguments, not solutions, was what suffered in the private report space condition.

the communication efficiencies afforded by the increased use of [deitic reference, i.e., pointing]; the visibility of a partner's task relevant action constraining the range of meanings attributable to an utterance." (p. 212).

This work is consistent with findings about the importance of open (or observable) interactions in team performance (Hutchins, 1990). The particular interactions afforded by the task environment and tools affect the nature of grounding. In the domain of ATC operations, Hughes et al. (1992) point out that the tools used (e.g., flight strips) allow for open interactions-- they allow all the relevant participants to easily see the state of the system and to see what actions others take on the system. On the other hand, some characteristics of work environments may inhibit the ability to ground. For example, Woods et al. (1994) point out that multifunction controls and displays used in the cockpit tend to suppress cues about the activities and intent of the other human crew member. This disrupts their ability to maintain a common situation assessment, which can degrade communication and coordination across the crew.

Grounding in dynamic fault management

In dynamic fault management there is an external reference: the process being managed. Not only must team members maintain an up-to-date common ground about the state of problem solving, they must also maintain common ground about the changing state of the monitored process. Many monitored processes cannot be observed directly, but only via data derived from sensors, transducers, etc.

Another difference is that in dynamic fault management applications there are typically several data sources which provide different levels of processed data; raw data may be available from computer-based displays (monitored process displays), and processed or integrated data (in the form of assessments) may be provided by "intelligent systems." Maintaining common ground about the state of problem solving and about the state of the monitored process requires knowing about the relevant activities of other team members because their activities may impact the process. This is important in order to be able to manage the process effectively (because knowing what to do depends in part on knowing what has been done, what is expected and what is planned for). It is also clearly important for diagnosis (in order to know what may be the causes(s) of an anomaly). Furthermore, team members also need to be grounded about relevant assessments of others because these can potentially affect expectations and plans.

CHAPTER III
RESEARCH STRATEGY

The goal of this investigation was to understand more about how human team members engaged in dynamic fault management support one another. An important consideration was to observe behavior in varied situations and under the actual constraints faced by practitioners. For this reason, a field study was done, as opposed to a simulator study. Particularly relevant for this domain are the constraints of time pressure, complexity and high consequences of failure.

Anesthesiologists in practice were studied because this offered the opportunity to investigate issues of explanation and more generally, cooperation and communication within a dynamic fault management application. It should be pointed out that the field study's goal was not to characterize the range of cognitive activities taken by anesthesiologists (Xiao, 1994). Rather, it was to investigate the above issues using the anesthesiologists in practice as a vehicle. We expand on the issues that anesthesiology in the operating room allows one to investigate below, but first we present a brief overview of what an anesthesiologist does.

General Goals and Activities of Anesthesiologists

The anesthesiologist's main goals during an operation are to maintain the health and safety of the patient and to create appropriate surgical conditions. From the anesthesiologist's point of view, the operation is divided into the following basic phases: preinduction, induction, maintenance, emergence and recovery. Preinduction involves preparation of the patient for anesthesia, which includes establishing intravenous access, placement of the patient on the operating table, placement of the monitoring sensors for the electrocardiogram, blood pressure, pulse oximetry, etc. During induction, the patient is put to sleep, intubated[1] and artificially ventilated. Thus, the beginning of a case, before the surgeon makes an incision, is a busy period for the anesthesiologists; they must undertake several activities such as attaching the equipment that monitors the patient's vital signs, placing catheters in the patient (for delivery of drugs and fluids and for monitoring critical parameters), getting drugs ready, administering drugs to patient, intubating the patient. In some settings, especially teaching hospitals, more than one practitioner is involved in many of these tasks. During maintenance, drugs and fluids are administered to keep the patient

---

[1] Insertion of an endotracheal tube in order to provide a clear airway and protect the patient's lungs against aspiration of gastric contents.

anesthetized for the duration of the operation and to maintain normal physiological function (e.g., intravenous fluid to replace blood loss). During emergence, when the surgical procedure is finished, the administration of drugs is discontinued and the patient is awakened and is extubated.

Anesthesiologists are the physicians responsible for managing the physiological process and diagnosing unexpected anomalies during the operation. To do this they need to monitor the patient's physiological signs and the effects of anesthesia and surgery. The major functions and signs that anesthesiologists must monitor are: depth of anesthesia, circulatory function, blood loss, respiratory function, respiratory and anesthetic gases, renal function, neuromuscular function, body temperature and other system functions depending on the type of surgery or the health of the patient (e.g., blood sugar, electrolytes, hemoglobin.) Clinical means (e.g., inspection, palpation, auscultation) as well as several instruments are used to provide indications of these functions and signs. Some important devices and monitors used include:

- electrocardiogram (ECG) to monitor cardiac rhythm,
- pulse oximeter to measure blood oxygen saturation and pulse rate,
- arterial cannula or "A-line", which measures arterial blood pressure, and is used to sample arterial blood which is sent to a lab for analysis of partial pressure of oxygen and carbon dioxide in the blood,
- automated sphygmomanometer (i.e., inflatable cuff for measuring blood pressure),
- central venous pressure (CVP) catheter, which provides an indication, together with blood pressure, pulse rate and urine output, of blood volume and hence a guide to fluid replacement,
- capnograph to measure airway concentration of carbon dioxide,
- mass spectrometer to measure and distinguish the concentration of various other gases,
- thermistors to monitor body temperature,
- Swan-Ganz catheter for measuring pulmonary artery and capillary wedge pressure and cardiac output.

Data from many of these measurements are available on an integrated computerized display (which we refer to in the text as the vital signs monitor). While some of these data are continuously available (e.g., the heart rate) others are available at intervals (cuff blood pressure) and still others require some explicit activity by the practitioner (e.g., cardiac output). Also, not all data is immediately available. For example, an arterial blood gas sample requires analysis in a remote lab and ten minutes may elapse between drawing the sample and receiving the results.

Management actions, such as administering drugs, blood, or fluids, are taken on the process depending on its state. Many drugs, each with specific actions, side effects and contraindications, are available to the anesthesiologist. Some types of drugs that are typically used during the maintenance phase of the operations observed are: inhalation anesthetics (for maintaining unconsciousness), narcotic analgesics, muscle relaxants, hypotensive agents, vasopressors and vasodilators.

For more on the cognitive activities of anesthesiologists see Cook, Woods and McDonald (1991) and Xiao (1994).

Practitioner Roles and Relationships

The operations observed took place at a large teaching hospital and involved at least two anesthesiologists: an attending anesthesiologist (or simply "attending") and one or two residents. The attending is a senior member of the anesthesiology staff, who holds a faculty position. In all cases the attendings observed here were board-certified, which means that they had completed a special examination in anesthesiology. The attending is responsible for overseeing several operations concurrently. He or she is always present during the induction phase of the operation, and typically returns periodically throughout the case. The attending adapts his schedule of visits depending on expectations about how the case will go and on assessments of the resident's competence to handle the case alone; for a relatively routine case he may only be present during induction.

The resident, an anesthesiologist gaining practical experience for four years after medical school, is present throughout the case, and in general manages the case. He is typically a senior resident (in his third or fourth year of residency). The operations with two residents had a senior resident and a junior resident (usually in his second year). Occasionally, a medical student was also present, but typically he was not involved in the management of the case (although the attending or resident might have him do some things, such as assist in intubation or in the placement of a catheter.)

General Communication Issues

Present in this domain are general issues of coordination and communication among agents managing some process and having different areas and levels of expertise. The attending and resident(s) must communicate and coordinate with one another, as well as with other personnel, such as surgeons and nurses who have different tasks, and who have the same high-level goal of preserving the integrity of the physiological process, although their lower level goals may be quite different (and may even conflict in some situations; see Woods et al., 1994).

Supervisory Control Issues

The attending is the more experienced, generally more knowledgeable, team member. He or she oversees the process and the resident, setting the general strategy and specifying certain actions and or decision choices. The resident defers to the attending in these decisions. Both attending and resident monitor and take actions on the process, but the resident is present during the whole operation, while the attending is present only some of the time (since he supervises other cases as well). When he returns to the operation, he needs to update his situation assessment (e.g., determine what events have occurred, how certain vital signs are proceeding); the resident will typically assist him in this process.

In these characteristics the relationship between attending and resident resembles a supervisory control relationship. In supervisory control, the human supervisor oversees some process and intelligent computer subordinates, each with local scopes of responsibility. These intelligent subordinates take control actions on the process, and may provide assessments to the supervisor. Figure 5 shows the similarities among the resident-attending relationship and the supervisory control relationship. In the operations involving a senior and a junior resident, the relationship among them was similar to that of an attending and resident in the sense that the senior resident directs strategy while the junior resident typically defers to the senior resident's decisions.

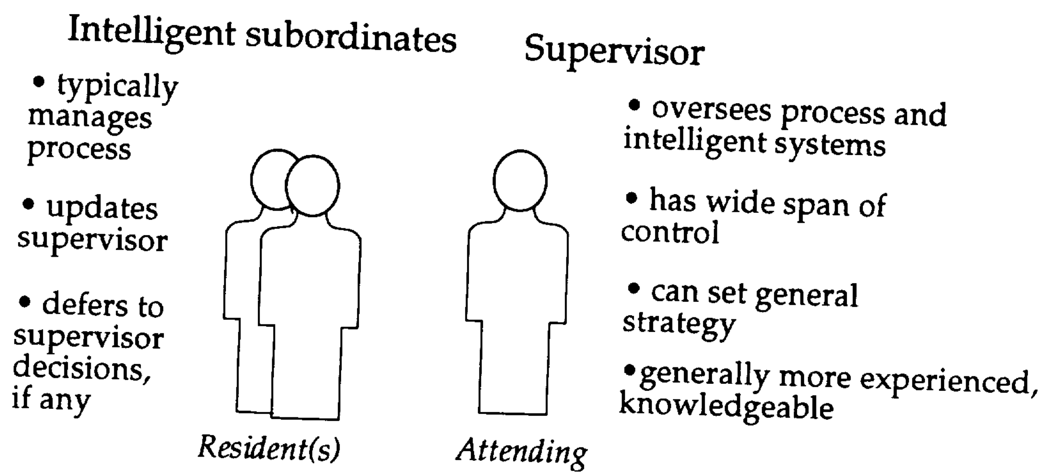

Figure 5. Supervisory control relationship among attending and resident.

The supervisory control form of interaction is important in applications in which a complex process must be managed and monitored, such as NASA space applications (Woods et al., 1991). The technological trend is to add more automation to complex systems, giving the human a wider span of subordinates to monitor.

Because of these relationships, the interactions among the anesthesiologists may provide a model of how intelligent subordinates can effectively support supervisors; or in the language of the introduction, how they can function as effective "team players" in supporting a team leader. Studying interactions among the resident and attending can provide insight into how an intelligent subordinate:

- receives and implements instructions,
- provides feedback on actions,
- offers assessments,
- updates the supervisor when his attention returns to the process.

Types of Joint Cognitive Activity and Information Exchanges

Anesthesiologists continually monitor and manage the patient's physiological process and diagnose unexpected anomalies. In order to perform the high level goals of management and diagnosis effectively, practitioners in this domain must form expectations about the future, plan courses of action, keep track of what has occurred, and evaluate past management actions and interventions. These cognitive activities are reflected in the information exchanges seen among team members. Figure 6 illustrates these types of joint cognitive activity (and information exchanges) and their role in serving the goals of management and diagnosis. The arrows that feed into management and diagnosis indicate cognitive activities that support management and diagnosis. The figure depicts past, current and future views into the process, i.e., patient's physiological state as monitored via vital signs. Updating, in which one team member informs another about relevant events that occurred during his absence, concerns past events, though it is driven by the goals of the present. Team members also talk about the past in another important way--in order to evaluate the effects of their interventions.

Practitioners establish expectations about the future of the process and about what actions should be taken; these expectations and plans[2] are incorporated into the present actions (Cook, Woods, and McDonald, 1991; Xiao, 1994). Information exchanges about the future reflect the need to anticipate problems that may arise, and to prepare courses of action to avoid them or be better equipped to deal with them (i.e., contingency plans). Some of these problems may be unlikely, but practitioners prepare for them either because the consequences are high or because coping with these problems when they tend to arise would be too resource-consuming.

---

[2] It should be noted that plans for courses of actions can be formulated at varying levels of detail and subject to the contingencies of the current situation (Suchman, 1987).

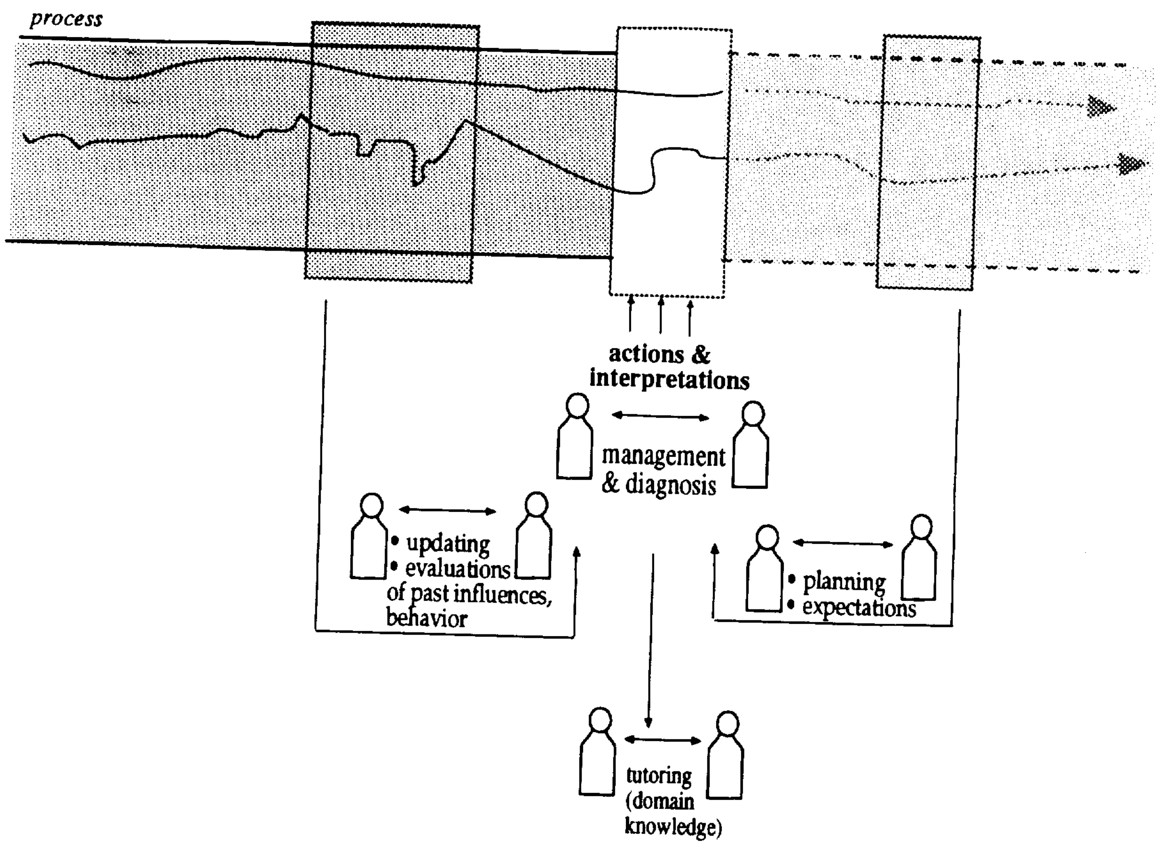

Figure 6. Different Types of Joint Activity and their Relationships.

Tutorial interactions are another type of information exchange observed in the data. These are exchanges about domain knowledge, given for the purpose of instructing a less expert practitioner (or would-be-practitioner). They are not critical to the moment-to-moment management and diagnosis.

The Field Study

Ten neurosurgery operations were observed and videotaped. The neurosurgeries involved one of the following: clipping of a cerebral aneurysm[3], removal of a brain tumor, or a laminectomy.[4] A list of these operations, along with some particulars, is given in the Appendix. The potential anesthesiology staff involved in these operations was comprised of 5 attendings and 5 residents. The practitioners were recruited by the second author who is also an anaesthesiologist working within the same organization.[5] The data collection was part of a larger study of physician

[3]An abnormal bulge in the wall of an artery in the brain, which could rupture and result in fatal hemorrhage.
[4]Removal of bone from the spinal column.
[5]Data collection procedures were approved by the university human subjects research committee; informed consent was obtained from physicians and patients prior to data gathering.

interaction with computers and physician expertise. Note that they were not told that this was a study of communication or explanation. All but one resident agreed to participate in the study.

A transcription of each was prepared from the videotapes. The verbal behavior was integrated with (a) other data about patient state, (b) physician activities, and (c) goals and intentions of the practitioners based on domain knowledge, procedures, typical practices to form a behavioral protocol (Woods, 1993b).[6]

Note that this study does not take a hypothesis testing approach. Rather, the approach involves posing guiding questions which define episodes of interest, which in turn, are analyzed given the conceptual framework.

Guiding Questions

The impetus for this study was the problem with the conventional approach to AI diagnostic assistance, in which diagnoses or recommendations are provided along with retrospective explanations for dynamic fault management. Hence, the questions addressed in this investigation are: how do human team members provide diagnostic and management support in dynamic situations? What is the nature and function of their explanations? We can further refine the latter question, by distinguishing between interpretations and explanations: How do team members communicate about interpretations and assessments? How do team members provide explanations for their assessments? How are these two kinds of explanations used in joint problem solving or in establishing situation assessment?

An important aspect of situation awareness in distributed dynamic fault management is keeping aware of the relevant actions that others have taken or will take because these may impact the process or the state of problem solving. Hence another guiding question is: How do team members keep informed about the relevant actions of others?

Data and Analysis

Data Sources

The data sources relied on in the analysis are: 1) verbal communications made by the anesthesiologists and those directed to them by others, or those that may have been overheard by them, 2) actions taken by the anesthesiologists, such as: looking at the monitor, any interactions with the machines, any adjustments to drugs or objects, any samples taken (and how taken), or drugs given, 3) actions taken by other personnel when interacting with the anesthesiologists, 4) behavior of the dynamic process as

---

[6]It was critical to the analysis that an anesthesiologist who also worked in the organization was part of the research team.

indicated by the patient record kept during the operation, and as displayed by the various monitors and machines, and a record of vital signs.

Three cameras were placed in the operating room so as to capture these data sources. One camera was focused on the various anesthesia machines and displays. Another focused on the area at the head of the operating table and in front of the machines, where the anesthesiologists spend most of their time. Finally, another camera focused on the patient, which captured close-up actions taken on the patient.

Transcription

The guiding questions drive the episodes selected for analysis, as well as the transcription process to some extent. It would be misleading to speak about data acquisition and data analysis as completely separate; during the transcribing, some data analysis occurs in the sense that the data is transcribed with certain omissions, while capturing and detailing certain behaviors. As Jordan and Henderson (1994) point out, there is no ideal transcript-- "..it is impossible to include all potentially relevant aspects of an interaction, so that, in practice, the transcript emerges as an iteratively modified document that increasingly reflects the categories the analyst has found relevant to her or his analysis."

The videotapes were transcribed in two passes. The first pass consisted of transcribing all verbalizations made by the anesthesiologists and verbalizations made by other team members to the anesthesiologists. The only verbalizations omitted were those that were obviously chatting about social activities or gossip. These were summarized as such in the transcript. Note that although we focused on verbal behavior in this study, the analysis is not a linguistic one (i.e., how language is used). Rather, it concerns how information exchanges function to support dynamic fault management. Also recorded were various activities taken by the anesthesiologists, including interactions with the machines, or other equipment, or administration of drugs or other fluids.

The next stage involved reviewing the transcript to identify particular episodes (which are explained below.) Having fairly detailed transcripts in the initial pass is useful because once an interesting episode is identified, it allows one to look back within the transcript to see what other episodes, activity or events may be related. Then these episodes were reviewed on video to verify the transcription and to fill in more detail about activities and vital signs.

The goal of field studies in cognitive engineering is to be able to generalize findings to similar situations in different domains (Hollnagel et al., 1981; Woods, 1993b). This involves taking a description of the actual performance (episode in raw transcript), which is concept-free and highly context-dependent and converting it to a description that is more concept-specific and domain independent. In order to arrive at a concept-specific and domain-independent description, one needs to have a framework or "conceptual looking glass" that guides these concepts. This framework was presented in Chapter Two. Figure 7 shows an overview of the logic of the analysis. The concept-specific description of the episodes of interest is contained within the text. This description is aided by a representation of the episode in domain-independent terms and, in some cases, in problem solving terms.

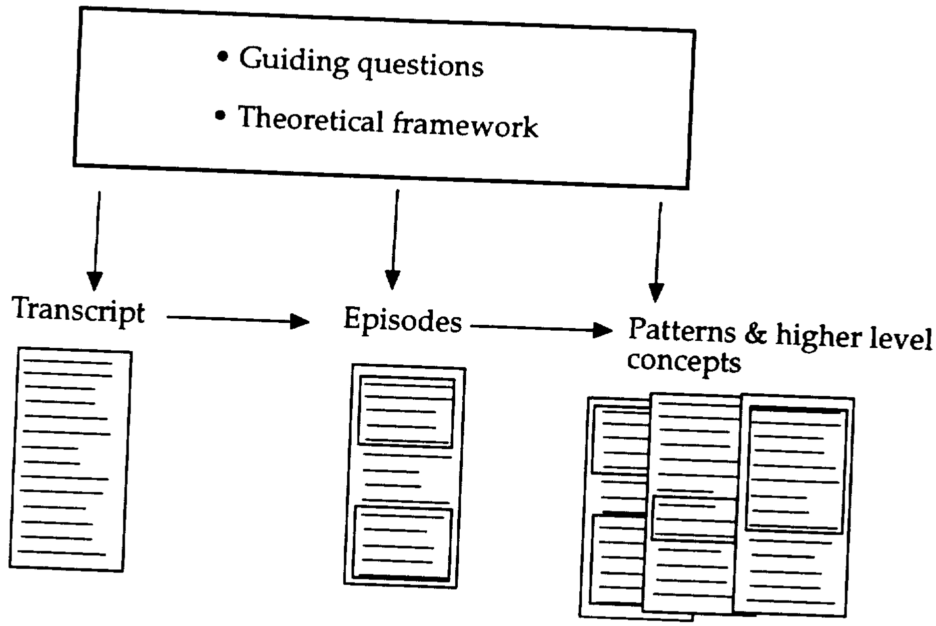

Figure 7. Basic Logic of the Analysis.

Episodes

Some of the episodes of interest are situation-driven and some are interaction-driven. The situation-driven episodes are those of *a priori* interest:

- Updates: situation in which the attending returns to the operation and is informed ("updated") of what has occurred in his absence, and
- Management/Diagnosis: situation in which team members are engaged in managing the process and/or diagnosing faults in the process.

The interaction-driven episodes are some episodes that contain information exchanges relating to situation assessment, decision making and task performance. These are any utterances in which information about activities, interpretations, explanations or any information about the process is being communicated. We call them *informs* for short. One type of inform is an *interpretation;* this can be a causal explanations or an assessment. In either case it is an interpretation of process data (e.g., relating data to one or more causes). For example, stating "pressure is low" is not a restatement of the data (e.g., blood pressure is 90) but a form of interpretation because being low depends on the context-- *low for our goals and expectations given this patient in this circumstance.*[7] *Explanations* are another type of inform; these refer to explanations of interpretations, i.e., explanations for why some interpretation is held.[8]

Updates

Updating episodes ("updates") are situations in which a returning team member (typically the attending) is updated about the state of the process and/or of problem solving. Updates are selected for study because they are found in many fields of practice in which a process must be turned over to another crew while on-line (e.g., air traffic control, nuclear power plant control room operations.)

Updates are relevant in supervisory control situations in which a supervisor periodically monitors several intelligent subordinates. The technological trend is for fewer people to monitor interacting subprocesses through the use of increased automation and intelligent subordinates. This means that supervisors may be coming into "advanced" situations. In some cases, the supervisor will have to be called into a situation which has escalated beyond the ability of the intelligent subordinate to cope with it. Effective updating is a desirable capability for such an intelligent subordinate.

Updates occur when the attending comes back into the operating room after having been away for some time.[9] Recall that the attending is present at the beginning of a case (during induction and intubation) but then may leave to oversee other cases. When the attending returns, the resident may update him on the progress of the case, and/or state of the patient. Particularly interesting are those interactions/updates after some critical event has occurred.

Management and Diagnosis

Monitoring and management occur continually. Of particular interest are episodes in which two or more team members are engaged in managing the process and/or in anomaly detection. The beginning and end points for an episode are not well defined *a priori.* However, generally speaking, this kind of episode will begin with a focus of attention on some anomaly, and end when an interpretation is arrived at and/or management action is taken and the topic is dropped, resolved or otherwise attains some closure.

---

[7] Note, however, that one cannot say that an utterance such as "blood pressure is 90" is always a statement of data. In some contexts such a statement can be intended and can communicate to someone else an interpretation or event (i.e., that it is low).

[8] These explanations may not occur with the telltale words typically associated with an explanation (i.e., "because", "the reason is..." etc.)

[9] There is another kind of update found in the data; this occurs when a team member comes to relieve the resident. This will typically be another resident or a CRNA (certified registered nurse anesthetist).

Not all of the cases yielded episodes that are discussed in the findings. Some cases, such as 3 and 6, were routine and relatively uneventful. Unlike simulator studies where the researcher can design the scenarios to address the questions of interest, field studies provide serendipitous opportunities. The virtue of this is that they offer unique conditions and situations that researchers might not have thought of ahead of time, or could not possibly devise in a simulator study.

Field studies can generate an enormous amount of "data." More precisely, what they generate is a lot of observations, notes, and transcriptions out of which very little may end up being relevant to the study. What may seem interesting at the outset, may turn out to be simply a piece of a larger concept of interest. Though it is important to have a conceptual framework and episodes of interest to guide the observation, analytic insights arise in the course of observation. The process of finding patterns in data iterates with observations; there is a cycle of observation and theory construction. As Suchman and Trigg (1991) point out, "...it is precisely in the repeated careful working through of the primary materials that theoretical insights arise. In this way analysis is something like iterative design. Articulations of themes and categories arise from familiarity with the materials and are constantly reevaluated against those materials. This, in turn, renews and extends one's familiarity. Furthermore, the identification of new themes and categories can lead one to return to the field or workplace to gather new materials."

Assumptions and Limitations

The Normative Assumption

The interactions among practitioners were studied as exemplars of good performance. However, this does not mean that performance is flawless or optimal. Rather, we assume that the patterns seen at an appropriate level of analysis can be taken as revealing the nature of effective team interaction.

Omissions

The video recording did not capture everything that may have been relevant. Video may be used in a naive attempt to "get it all," but this is not possible. In this study, for example, some exchanges among team members may have occurred off-camera, or out of line of sight). Also, not all utterances were captured on tape; some were inaudible, or incomprehensible. Both of these imply that quantitative assessments are not possible, except at a gross level.

Representations of the Findings

Understanding the information exchanges among team members and how these exchanges support dynamic fault management relies both on understanding the domain particulars (e.g., to know why mentioning blood pressure now is informative) and on understanding the context for the episode (i.e., what relevant events occurred prior to the episode). Just as an utterance may take on any of several meanings depending on its context, the meaning or significance of many episodes cannot be understood without knowing their context (e.g., what occurred previously in the case, what parameters have been of concern, what practitioner expectations are.)

The main type of representation used in the Findings section provides a segment of the transcript (corresponding to an episode) with footnotes to assist the reader in understanding some of the domain-dependent details. Also, to assist in the analysis, a domain-independent description is provided alongside the transcript. Episodes that involve diagnosis also contain a third column, indicating phase of problem solving. Some episodes are short and do not contain additional columns beyond the transcript. The conceptual level description of each episode, i.e., why the episode is significant for the purposes of the study, is contained within the text.

The utterances in the presented transcripts are not timestamped (though this data is available from the videotapes) because the dialogues in the episodes typically do not have long pauses between utterances. Where relatively longer pauses are found, these are noted.

An identifying code is used before each episode. The code indicates the case the episode occurs in and the transcript time, as follows:
[case | hour: minutes: seconds].

Key to Transcription Symbols

There are several different transcription schemes that could be used, each capturing different aspects of speech and dialogue. For my purposes, however, a rather simple scheme given below suffices. Aspects of prosody were captured in the second pass transcription, if they were deemed relevant to the analysis.

- Ellipsis indicates missing, inaudible or incomprehensible text.
- Ellipsis in parenthesis indicates approximate number of incomprehensible words represented by the dots
- Italics indicates actions
- Words in parenthesis express some uncertainty about the actual words.
- R = resident (used in cases where there is only one resident)
- RS = senior resident, RJ = junior resident
- A = Attending
- M = Medical student
- S = Surgeon, SA= assistant surgeon
- N = Nurse
- v.s. monitor = vital signs monitor, an integrated monitoring system that displays all the patient's vital signs.

CHAPTER IV
FINDINGS OF THE FIELD STUDY

The introduction pointed out the problems both with the form and timing of explanations in the standard expert system approach to diagnostic assistance. This problem defines the guiding question for the investigation: How do human team members collaborate in management and diagnosis? We begin by examining episodes in which team members are engaged in forming interpretations of the process state. Recall that these are one type of explanation relevant in dynamic fault management. Then we examine the characteristics of explanations for interpretations or actions, which are the counterparts to AI justifications. These findings lead to an investigation of important factors that play a role in the nature of diagnostic support among team members in this domain.

Explanations in Dynamic Fault Management

Joint Interpretations about the Process State

Episodes of management and diagnosis reveal instances of interactions that might be termed *joint interpretation*. This term attempts to capture three important aspects of the interaction: it is a *process* in which team members are *involved* in coming to the *mutual acceptance* of an interpretation.

Episode: Anomalous blood pressure

Consider the episode in Figure 8. In this episode, determining whether to take a management action and what it should be, depends on the exact nature of the interpretation for the anomaly. This interpretation is not simply transferred from one agent to another. Rather, both are engaged in the process of arriving at the interpretation and both contribute to its development. Though the senior resident proposes an initial interpretation ("they must've stimulated something"), it is followed by a period of investigation and verification in which both team members are involved. The junior resident is kept involved by the senior resident, in the sense that the senior resident does not simply provide directives, but also states his assessments and reasons. The junior resident's comments and actions provide evidence of attending and comprehending. For example, she implies understanding when she points out corroborating evidence.

The subsequent exchanges concern further verification and testing that relate to establishing that the cuff pressure measurement, rather than the arterial one, is the artifactual one. This phase ends with the senior resident providing a directive for corrective action, along with the statement "I think it's a true pressure" which can be considered an explanation or reason for the directive. But, in this context it functions more as a confirmation of the mutually held interpretation, because understanding the statement depends on the information that has been gained through the joint interpretation.

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE
[1014:22:38] *{R2 looks at v.s. mon}* R2: [R] R: yes?	R2 draws attention to anomaly	**anomaly identifed**
R2: his pressure's now reading 177 *{R2 hits b.p. button to start cuff measurement}* R: They must've just stimulated something.	R suggests interpretation	**hypothesis suggested**
{R2 adjusts anesthetic agent, gets syringe} R: Don't give him anything yet, see what the cuff pressure is[1]. It's a lot better waveform than we were having, so I think it's probably true, they stimulated something *{both looking at v.s. monitor}*	R tells R2 to wait to take management action she was about to take, saying they should check another parameter (because anomalous data may be artifactual). But also comments on qualitative characteristic of data suggesting accuracy of reading	**verification /testing hypothesis**
R2: yeah, his heart rate picked up 5 points *{indicates to monitor}*	R2 calls attention to a related parameter that suggests anomaly is not artifactual	**corroborating evidence pointed out**
R: yeah, you'll see that when they're doing cervical, especially anterior, posterior not so much, but the anterior, you'll definitely see, you gotta be looking for vagal stimulation, you got the vagus...you got the carotid (body), and you gotta be watching for all those things, it's just a real touchy surgery..this is not abnormal at all	R mentions conditions under which flucuations in parameter would be expected	

[continued]

[1] Blood pressure is measured by two sensors; from an arm cuff and from the arterial line. The arterial line displays blood pressure continuously as a waveform. The cuff pressure, in contrast, is a discrete value measured intermittently. When an arterial line is present, cuff pressures are measured typically every 15 to 30 minutes.

Figure 8. Episode: Anomalous blood pressure.

Figure 8 (continued),

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE
[continued]		
R2: his blood pressure cuff is making me crazy	R2 comments on second sensor value	
R: they correlated well until we started surgery and now neither one of them		
R2: yeah		
R: ..(probably cause his arterial) pressure's changing so rapidly, plus they're leaning on his arms....	R calls attention to data indicating why second sensor value is artifactual	verification/ testing
R2: *{R2 looks back}* right there	R2 corroborates latter hypothesis	
R: cuff can't read his pressure well. Give him another 50 mics of fentanyl.[1] I think it's a true pressure *{R2 administers}*	R directs R2 to take a specific management action consistent with the interpretation, provides explanation	corrective action directed and taken
R: *{hits b.p. button}* gonna drop his cuff, it's just complaining too much.	R cancels the second sensor measurement (which has been deemed artifactual)	

[1] A narcotic which blunts the response to stimulation.

Episode: Evaluating the effects of interventions

The following episode captured in Figure 9, highlights an important part of joint interpretations in this domain: keeping one another involved in the evaluation of the effects of interventions. When the senior returns from his break, the junior resident informs him of the values of some data items. The senior resident relates these results to previous interventions in an evaluative statement. He also suggests taking another intervention ("coming back down on the nitro"), which is then evaluated aloud when he returns ("turning down the nitro was a good thing....").

This episode also illustrates an interplay between data and evaluation. Notice the relationship between the junior resident's calling attention to particular data values and the senior resident's cognitive activity in response: evaluating possible management actions, developing further management actions, or verifying states (e.g., when the junior resident queries the senior resident about the low temperature, the senior resident engages in checking behavior --to see that the blood warmer is on). So, far from distracting the senior resident, the junior resident draws attention to data that assists in forming an interpretation of the process. Through the senior resident's voicing of his assessments, both team members can become calibrated about the state of the process and management, or, in the language of the conceptual framework, they add to the common ground.

Episode: The state of management

The next episode (see Figure 10) illustrates how team members keep one another aware of the state of management activities. The episode begins with the resident informing the attending of an intended action. Team members do this relatively often. Depending on context, this serves several possible functions: it allows for confirmation that the action is an appropriate action at the current time or alternatively, it allows the other team member to halt or modify the proposed action. It adds to the common ground-- the other team member knows what will be done and hence can modify expectations appropriately. The joint discussion of management strategy in this case makes explicit particular concerns and engages "two heads" in thinking about what to do.

The information exchanges among team members always occur within a dynamic physical context. Even while talking to one another, practitioners can continue to monitor the changing physical context. This physical context affects the information exchanges --interspersed with talk about evaluations or anticipations, are comments about events and data. For example, notice how the attending makes reference to the urine output level as he is talking about management activities. So even though team members are talking about management and evaluation, for example, they may also talk about perceived events occurring in the present.

This physical context also plays an important role in facilitating communication (See section on Contexts for Communication) and in the observability of interactions (see section on Information through Noticing). Figure 11b has an example of the latter; towards the end of the episode the attending sees that the resident is writing down the lab results as he gets them over the phone. The attending is able to look over the resident's shoulder and read the relevant data as soon as it comes in (i.e., to be able to take the management action as soon as possible.)

Figure 11 provides another example in which team members discuss and evaluate a course of management taken earlier in the case.

These episodes illustrate the "on-going talk" about the monitored process and problem solving state that team members engage in. This is a way team members can keep one another calibrated in the moment-to-moment interpretation and management of the case.

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION
[2 I 2:15:00]	
RJ: I did another output[1] and it was five four, something like that	RJ informs about result of test
RS: So she likes the dobutamine[2]	RS evaluates result
RJ: Her SVR [3] came down (8 point 2)	RJ informs of relevant parameter, provides reference and value
RS: So she he likes--that's--we could come back down on the nitro[4], come down about a half if you want *{RJ turns it down on infusion device}*	RS suggests management action RJ takes suggested action
RS: I'm gonna get some more gloves. I'll be right back *{RJ steps away}*	
:	
{2 minutes later RS returns}	
RJ: *{points at temperature indication on v.s. monitor}*: .. cold..hypothermia....	RJ queries about parameter value RS provides possible explanation involving an intervention he took
RS: I think the problem was I dumped all that mannitol[5] in, boom *{chuckles}* ...well it's just--her blood warmer is on right?	RS queries about relevant equipment state
RJ: yeah, I checked that	RJ responds: already checked
:	
[omitted utterances about blood warmer]	
:	
RS: Turning down the nitro was a good thing to do, she's starting to get .....I think the dobutamine's finally done its job. Took it a while.	RS evaluates previous intervention on process
RJ: ...right back where she was	RJ concurs

[1]Test of cardiac output.
[2]Generally recommended in cases of increased SVR, normal heart rate and low cardiac output.
[3]Systemic vascular resistance, which is blood pressure divided by cardiac output.
[4]Nitroglycerin, a vasodilator, for controlled hypotension, useful in patients with known or suspected coronary disease.
[5]An osmotic diuretic given intravenously.

Figure 9. Episode: Evaluating the effects of interventions.

TRANSCRIPT	Domain-Independent Description
[1 I 3:23:40]	
R: I'm gonna go ahead and send another gas[1]	R informs A of measurement he plans to take
A: yeah, let's send another gas and	A concurs
R: see where we're at. Have a feeling it's[2] still gonna be low, he's just oozing all over the place[3]	R states expectation for measurement value and reason
A: I think once we bring the temperature up we have done all we can do, you know, he's putting out urine[4] I think I see more there.	A states concern for particular parameter, summarizes state of management : "we've done all we can do" A states an observation: parameter that has been of concern seems better
R: yeah, there is a little more there. I'm gonna empty that in a couple of minutes	R verifies, mentions taking an action that will allow them to assess value
A: yeah. (..) Dr [R] R: temporarily yeah	
A: I mean, there is nothing we can do other than Bear Hugger™ now. We can get one more and put it on the lower part of the body. Another Bear Hugger™. we can get some heating lights on the field.[5] :	A summarizes management plan concerning one parameter: provides two options
[blood gas taken; episode continues]	

[1] Blood gas. Sending a blood gas refers to sending a sample of blood for analysis of: pH, partial pressure of oxygen, partial pressure of carbon dioxide, hematocrit, base excess, sodium, potassium, calcium and glucose.
[2] Hematocrit.
[3] A reference to the patient's bleeding.
[4] Low urine level has been a concern so far in this case.
[5] All of these options concern efforts to maintain adequate temperature.

Figure 10. The state of management.

Figure 10 (continued),

TRANSCRIPT	Domain Independent Description
{Blood gas taken 9 minutes ago} [continued]	
A: we'll see how the blood hematocrit comes back, if it is low then we'll add a unit of blood. I'd rather more blood than crystaloids	A informs of plan, dependent on test result value
R: I think it's gonna be low. I bet you it'll be 25, maybe 28 *{they look at surgeons}*	R states expectation
A: are you gonna be taking any grafts also or just	A requests information from different team member which could affect course of management
S: oh yeah, man we're gonna be we're gonna you name it this is a bear, yeah we gotta we're gonna give him some bone so maybe he'll fuse, heal ...	
{Phone rings, [R] answers} R: OR [#] {patient name} umhm, umhm, *{A looks over for a moment to area where R is writing, then looks up}*	A looks at incoming result
A: [N], I'll take 2 units of blood. R: 1.4, 8,4 *{pause}* alright, thank you *{hangs up}*	A directs another team member
A: I think ...should be air and 02 only *{adjusts knobs}* he's not liking nitrous very much	A takes action, informs and provides explanation
A: What was the calcium? *{looks at record on table}*	A requests parameter value
R: It was down alittle bit, 1.84	R responds, provides reference information
A: I'd give him 500... *{R gets up}*	A directs management action

TRANSCRIPT	Domain Independent Description
[5 I 1:54:43]	
A: Something I can help ya with?	A offers general assistance
RS: Nothing, he's doing okay.	
A: Did you get an output recently?	A requests information on a parameter of concern
RS: *{turning vs. knob}* yeah, 7,9, let's see	R informs of value
A: really!?	
RS: yeah. That was a combination of 3 outputs so it's pretty accurate. His index[1] is 3,4. It's still low but I'm just	R elaborates with information about accuracy of value, informs of related parameter value
A: .I would just ..LR.[2]..otherwise...	A suggests management action
RS: I'll put a little LR up on there *{indicates to left IV tree}*, I'd rather..	R indicates he will take it, elaborates
[a few omitted utterances about gas results]	
RS: yeah *{pause}* I think we did the right thing, I think things went really nice this morning.[3]	R evaluates past action/strategy
A: you can never be faulted for over-monitoring somebody, I'm sorry.	A confirms, elaborates
RS: Prone position just adds another factor that you have to think about.	futher confirming and elaborating of past action/strategy
:	

[1] Cardiac index, which is the cardiac output divided by the surface area. It allows one to adjust for the size of the patient; the cardiac output should generally be larger for larger people.
[2] Lactated Ringer's solution, a kind of IV fluid that contains a substrate that can be converted in the body for use as an energy souce.
[3] A reference to the decision to place a Swan-Ganz catheter earlier in the case. At the beginning of the case there was some uncertainty about the extent of the patient's heart disease.

Figure 11. Evaluating a course of action.

Explanations for Interpretations and Actions

In the transcripts many explanations for interpretations and actions have two basic characteristics. One is that they are unprompted that is, they are provided along with the stated interpretation or action, rather than being given in response to a query. Note that this term is not meant to imply that there is no external stimulus for the explanations in the environment, but simply that they are not explicitly requested by another team member. The other characteristic is that they are often brief and tend not to be "deep level" explanations (i.e., they do not make domain knowledge explicit). In the terminology of Toulmin (1958), team members often provide grounds (the data on which a claim is made) rather than warrants (the relationship between the grounds and the claim). The example below illustrate both characteristics.

In the following interaction, the attending has been talking to the resident about non-case related domain knowledge, then without pause he comments on something he has noticed:

A: .... He probably needs some fluid I would think, his urine looks pretty dark.
RS: Yeah, let's give him some of this.

In this example, the attending tags his assessment (which is also a directive in this context) with an unprompted explanation ("his urine's pretty dark"). Note also, that stating the data on which the assessment is based, evidently suffices as an explanation. For this to be possible, there needs to be shared mutual knowledge (i.e., such explanations would be cryptic to a layperson), in some cases it may rely on mutual knowledge established during the case. Another way to put this is that explanation can be "compact" because it relies on various contexts (see section on Contexts for Communications.)

Communications of actions are also sometimes coupled with brief explanations for these actions. For example, when the attending in Figure 10 turns down the nitrous oxide, he informs the resident of his action ("..should be air and O2 only" and also supplies the explanation ("...he's not liking nitrous very much.") As in the above example, this brief, cryptic explanation requires mutual knowledge to be completely understood.

Not all of the spontaneously offered explanations will be needed from the listener's point of view. However, the general tendency to provide unprompted explanations is useful because it is a mechanism to add to the common ground and thereby forestall future misunderstandings.

Long explanations of reasoning are the exception rather than the rule in episodes of management and diagnosis among team members. Where one sees long, retrospective explanations of reasoning are in tutorial situations and when there are disagreements about a course of action. Tutorial explanations, which are about the subject matter and practice of

TRANSCRIPT	Domain Independent Description
[10 l 0:40:00]	
R: you know we're not gonna be able to get anything bigger than a 7-0 down his nose	
A: that's fine....	
A: [R2] let's draw up some, grab a 10cc syringe, and a 16 gauge IV catheter, take the needle out of the catheter. And then draw this stuff up *{hands object to R2}* he does have tight nostrils.	
R: I just think an oral's the way to go	
A: I'll tell you why later, draw it up just from the syringe then we'll just drip that down this nostrils first...	A defers explanation (for course of action)
R: {to R2} .....we need a second IV and we need an A-line we haven't started on yet. R: I'm gonna start on the A-line. *{goes to P's left arm}*	
[omitted utterances]	
[10 l 0:44:35] R: I think topicalizing with that is the way to go, I think it's the best there is	
A: Oh I agree, but I still believe that (first medicine) was absolutely necessary, I'll explain why when(...)experiences on Thursday.	A provides partial explanation
R: Oh I heard about that, I think I walked through while you were doing that	
A: So, I'm just gonna be very conservative.	
[continued]	

Figure 12. Deferring Explanation.

Figure 12 (continued),

TRANSCRIPT	Domain Independent Description
[R has difficulty doing the nasal fiberoptic intubation; a new scope is brought in; finally A tries it]	
[10 I 1:13:15] A: Deep breath *(to P)*	
R: Let's just try an oral just once, see if we can do it	
A: Deep breath *(A inserts tube)* pressure (.) deep breath	
R: ..think he got it	
(A listens for sounds of correct placement) A: See [R]. *(then to P)* Alright you're gonna go to sleep now.....I haven't lost my touch. We're going to sleep now *(A injects P, R squeezes ventilator bag)*	Deferment of explanation (for course of action)
A: When he's asleep I'll tell you why I didn't do an oral	
(R adjusts anesthetic agent) R: Okay, is that what you tried to do the other day? you had to....	
A: We tried everything. This has to be jammed in a little tighter. *(referring to some apparatus connected to P)*	
[10 I 1:15:08] A: ..some agent?	
R: yeah	A provides explanation
A: I just think with orals you really have to sedate him a lot more...	
R: See I don't. I've done enough of them, you don't have to sedate them at all	
A: (..) that big airway... R: Absolutely. A: But you should be able to do them both	
R: No, I just think, my choice with a guy like this with very small nares, and he's this large, I rather have a bigger tube, and to do that you have to go oral *(R adjusts ventilator settings)*	
A: I think a 7 will be no problem with him, we'll	

anesthesiology,[1] are intended to help a less experienced practitioner (typically a medical student or sometimes a resident) gain domain knowledge. In the teaching setting where the data were gathered, teaching and learning are part of the work environment: residency training is an apprenticeship and medical students are allowed to attend and participate.

The following episode illustrates the important role that the on-going events and that the tempo of activity have in affecting the timing of potentially long explanations (See Figure 12). The explanation in question is expected and given because there is a disagreement about a course of action: the attending has decided on a nasal fiberoptic intubation, but the resident would rather do an oral fiberoptic intubation. The resident voices his opposition but the attending has final say. Note that at least twice the attending defers his explanation for the course of action. This deferment is based at least in part on the time pressures or on the situation not being appropriate. It is only after patient has been intubated and anesthetized that the explanation is finally completed, and the topic is dropped.

Unprompted Communications

Information Exchanges about Activities

There are numerous physical actions required to perform anesthesia. Practitioners engage in many different activities in order to get the patient ready for anesthesia and then as part of the maintenance of anesthesia. This is reflected in the high degree of coordinative talk, which includes explicit descriptions of one's own activities and queries about the activities of others. This is most obvious in tasks that requires fine synchronization, such as turning a patient on the operating table. Team members inform others of finished activity and seek information about such activity. This is important because many activities are contingent on a particular phase of the operation or on existing preconditions (i.e., other actions having been taken by other team members or not been taken by others). For example, adequate anesthesia should exist before the surgeon begins incision. If this information has not be established by some other means, the surgeon will typically ask the anesthesiologist some variant of: "is it okay to begin?" Note that this coordinative question also serves to inform the anesthesiologist that the surgeon is ready to begin.

Distributed management and diagnosis requires that team members be aware of those activities of others that may affect the monitored process. Keeping track of the various potential influences on the process is critical for diagnosis (in order to know what may be the cause(s) of an anomaly). It is also important for management because knowing what to do depends in part on knowing what has been done and what is expected.

---

[1] For example, they can be about "how something works" (e.g., the cardiovascular response to a particular drug) or about how to perform a particular technique, or about how some equipment operates.

Team members assist one another in this respect by unprompted communications about their relevant activities. The transcripts show many instances of team members spontaneously telling one another what they have done, are doing or are about to do. The following example illustrates a case in which a team member informs several other team members about an activity he is about to undertake that may affect their current or planned activities:

[5 | 0:26:35]
{RS brings a tray on a movable stand to P's left side, sets a sterile "kit" on it; someone else stands a couple of feet from tray}
RS: Okay, I'm gonna be opening up a kit here so just watch your elbows *{opens it up}*

In some cases a junior member will inform a more senior member of what he is about to do. For example, the junior resident sometimes lets the senior resident know that he is turning on the anesthetic agent or that he is going to be giving a drug. These "informs" of activity have another purpose besides letting the team member know about a new influence on the process; it allows for confirmation, prevention or modification of the planned action. The following episode, which occurs just before induction, illustrates:

[8 | 0:48:40]
R: I'm gonna turn on just a little Forane, since *{moves towards anesthetic agent}*
A: No
R: No?
A: I mean you have (syringes in your hand)...[2]

Generally though, one finds that actions are approved (see example below). This is to be expected given the informing team members are relatively knowledgeable, experienced and grounded.

[2 | 2:31:35]
RJ: I'm gonna turn the nitrous back on now
RS: yeah...

Not only do team members inform one another about their own relevant activities, they may also inform one another about a relevant action taken by a third team member. In the following episode, the information assists in preventing an undesirable situation (administration of a second dose of a drug):

[2]This appears to be a comment about preferring to use intravenous anesthetic agents instead of the potent inhaled agents, which increase cerebral blood flow and hence intracranial pressure, which would be undesirable in certain circumstances, such as in a patient with a head injury.

[2 l 0:50:33]
RS: Why don't we try a little ephedrine on her *{RJ turns knob on v.s. monitor, the left upper window goes away, gets complete waveforms}*
RJ: Yeah, he just gave some ephedrine.
RS: Did he? Okay.

Highlighting anomalies, events and parameters of concern

> When a player is a particularly high scorer it is worth noting periodically how he is doing. "
> ---*So you want to be a Sportscaster.* Coleman, Ken (1973).

Besides unprompted communication of activities, team members point out anomalies and events to one another, they voice concerns and talk about evaluations of past interventions or plans. Particular parameters may become especially important to monitor either because of some preexisting patient condition or because they become anomalous during the case. The quotation above conveys the importance of monitoring and communicating about certain parameters of concern, *even when they are not necessarily anomalous.* For example, during updates, team members will often comment on the parameter of concern whether it is abnormal or not; that a parameter of concern has become normal is also informative.

The following episode illustrates how drawing attention to a parameter of concern can lead to a discussion about management that both evaluates past management actions and attempts to develop new ones. This episode comes from a case involving a patient who has lost a lot of blood prior to this episode. His temperature became a parameter of concern early on, as did his urine output.

[1 l 1:43:44]
A: *{looking at the v.s. monitor}* temperature 35.2 eh?
R: yeah, I turned the room temperature back up but
A: but
R: I think it slipped down somehow
A: oh
R: I turned it up when I came in
A: running....and the humidifier
R: got the humidifier on maximum
A: the other thing we can do that has helped is have that aluminum foil. sometimes we can just wrap the circuit this way the loss of heat by radiation is less and it kind of just keeps it little warmer *{looking at monitor}* but I think since now they have covered the field *{indicates back}*..hopefully it should be. But ah
R: I thought these heating wires were supposed to keep it warm
A: supposed to, but you know

R: anything helps
A: anything helps just a little less heat loss here and just alittle less heat loss there

An hour and a half later the attending makes another reference to the temperature.

[1 l 3:3:40]
{*A and R looking at v.s. monitor*}
A: temperature...
R: I've been turning the room temp up, there's not much more we can do unless we get, they don't have any of those um, one time, they demonstrated a Bear Hugger that could be used interop?
A: yeah,..bring the Bear Hugger™, you know
R: {*gets on phone*} could we have one interop Bear Hugger™..

Pointing out anomalies and parameters of concern to one another may lead to a discussion about management in which new ideas may be generated, as in the above segment. It can also serve to make another team member aware of something about the process that another team member believes significant, and which they may not have noticed.

Relative References

An important point about how team members talk about anomalies and parameters of concern is that they often talk about them in a *relative* way, with reference to what the parameter value was earlier. For example, in episode 9, the resident states that "her SVR *came down*." In episode 10, the attending asks about the calcium and the resident states that it was "*down a little bit*, 1.84."

Information through Noticing

Much information about the state of problem solving can be picked up by overhearing and seeing what other team members are doing; it is not necessary for team members to always direct attention and explicitly provide this information to one another. Valuable information about team members' intentions and potential influences on the process can be picked up by being able to observe the activities and interactions of others. In studying cockpit crews, Segal (1994) notes the importance for coordination of being able to observe the activities of other team members (a form of information exchange which he terms *consequential communication*.) Some work environments foster "open" interactions, i.e., interactions that are observable and understandable by others (Huchins, 1990). Aspects of work environments also allow team members to gain information about activities that have already been taken. For example, anesthesiologists can tell what drugs other team members have given by looking at the anesthesia record, or by seeing what ampules are empty on the drug cart. Assessments and plans may also be picked up or inferred by overhearing interactions between other team members. All of this information is ultimately valuable in allowing team members to update their situation assessment and expectations of the monitored process.

Team members notice what others are doing and on relatively rare occasions may observe behavior that they don't understand, that doesn't fit with their expectations, or that suggests that the other team member could use assistance. They have a sense of how activity should be occurring and are able to pick up on discrepancies in the expected activity --when things seem "unusual." It is generally in such instances that a team member questions another about his activity. See episode below. Open interactions allow possibilities for individuals to detect actions that may be inappropriate in context and to initiate recovery before outcome failures occur.

> *[4 l 0:39:25]*
> *{RS sprays numbing medication into P's mouth, turns to get gloves, turns back, RJ is lifting P's left arm slightly, touching pressure cuff line}*
> RS: What are you looking for?
> RJ: Just to see if that was *{points towards monitor}* correlating with that.[3]
> RS: *{looks towards monitor, putting on gloves}* They were correlating yeah, very well. *{looks back to P}* She's just a little anxious with me doing this.

Another notable source of information is the "self-talk" of others. Team members occasionally talk aloud when engaged in a task or when trying to figure something out. This may serve a dual purpose. First of all, it may help the practitioner who is talking to "keep track" (e.g., of required actions or possible alternatives.) It is also a mechanism that allows other team members to notice someone's activities, plans or reasoning and to provide assistance, if necessary. This is accomplished generally without distracting or demanding attention. A common, brief form of self-talk which is found in the transcripts is saying (or sometimes sighing) "okay" or "alright" upon completion of a task or subtask. This can, in some circumstances, assist in coordinating behavior by letting someone else know that a particular stage is finished.

Heath and Luff (1991) point out that the coordination in the London underground line control room relies on "surreptitious" monitoring of self-talk by other team members. Their field study found that it was "relatively unusual" for team members to explicitly give information to one another. Overhearing and monitoring actions allows them to keep aware of what others are doing. The structures of certain work environments coupled with

---

[3] The blood pressure measurement as indicted by the arterial catheter and that measured by the sphygmomanometer (pressure cuff). A check on the arterial line blood pressure measurement is done intially by seeing if it correlates with the blood pressure cuff measurement. (The arterial line may fail or stop reading because of a blood clot at the tip of the catheter or some technical problem.).

the demands and tasks performed may make it unnecessary for team members to explicitly tell one another about their actions and assessments even though coordination is required.[4]

Shared Tools

Information about the state of the process and of problem solving can also be picked up from the tools that are publicly available to the team members, such as the various displays and the anesthesia record. In order to maintain a common frame of reference, these public tools need to afford information access in a way that is consistent with all team member expectations. Consider a shared artifact like the anesthesia record. It is used by several people both for recording actions and values (e.g., what drug was given when and how much, or what the blood pressure has been for 5 minute intervals throughout the case) as well as for retrieving that information. A representation that is used and modified by team members in different ways can create rifts in the common ground. In the following episode, the senior resident comments about the different way the junior resident has annotated the record. The junior resident has written the results of the blood gas sample which was drawn at 1:45 on the first page of the anesthesia record. However, the last time written across the top of the first page is 12:30 (see Figures 13a and b).

> [10 | 7:56:40]
> {*The senior resident has just returned from a break. The junior resident who remained in the operating room begins to update him, while the senior resident looks at the record.*}
> RS: {*looking at record*} oh, you just drew another gas
> RJ: yeah, I just sent
> RS: I usually end up putting the next gas, when you go to a new page over here so you can look down
> RJ: oh so you can follow it
> RS: it's not a big issue, that's what I usually do. No big deal.

---

[4]Humans have mechanisms that allow them to become aware of stimuli that are not directly attended to, e.g., peripheral vision, sensitivity in divided and focused attention tasks to important words, and to changes in volume or pitch. This is exploited in domains that use voice loops (e.g., in mission control or aircraft carrier operations). Rochlin et al., (1987) points out that in aircraft carrier operations checks are routinely performed on decisions via a constant conversation loop. Rochlin et al. say "At first little of this chatter seems coherent, let alone substantive, to the outside observer. With experience one discovers that seasoned personnel do not 'listen' so much as monitor for deviations, reacting almost instantaneously to anything that does not fit their expectations of the correct routine."

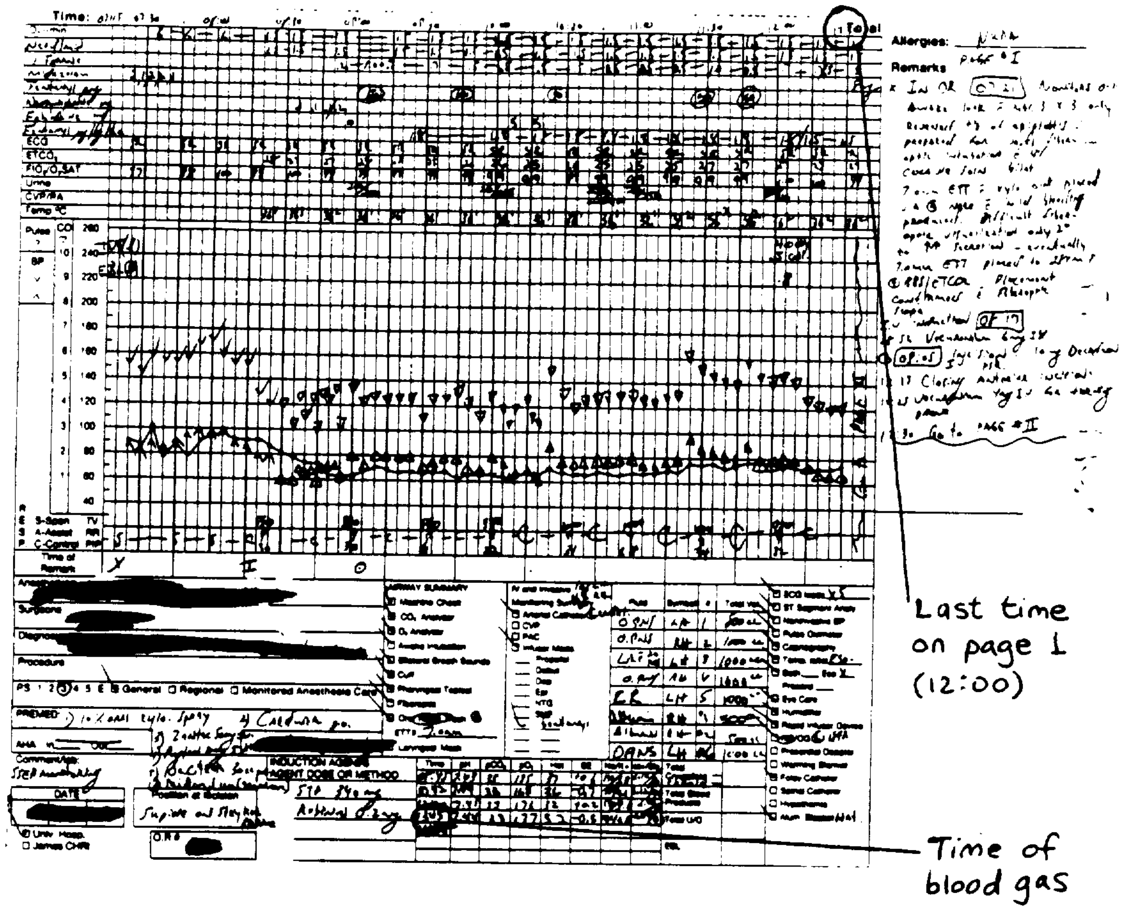

Figure 13a. Page 1 of the anesthesia record. Note time of blood gas and last time on page.

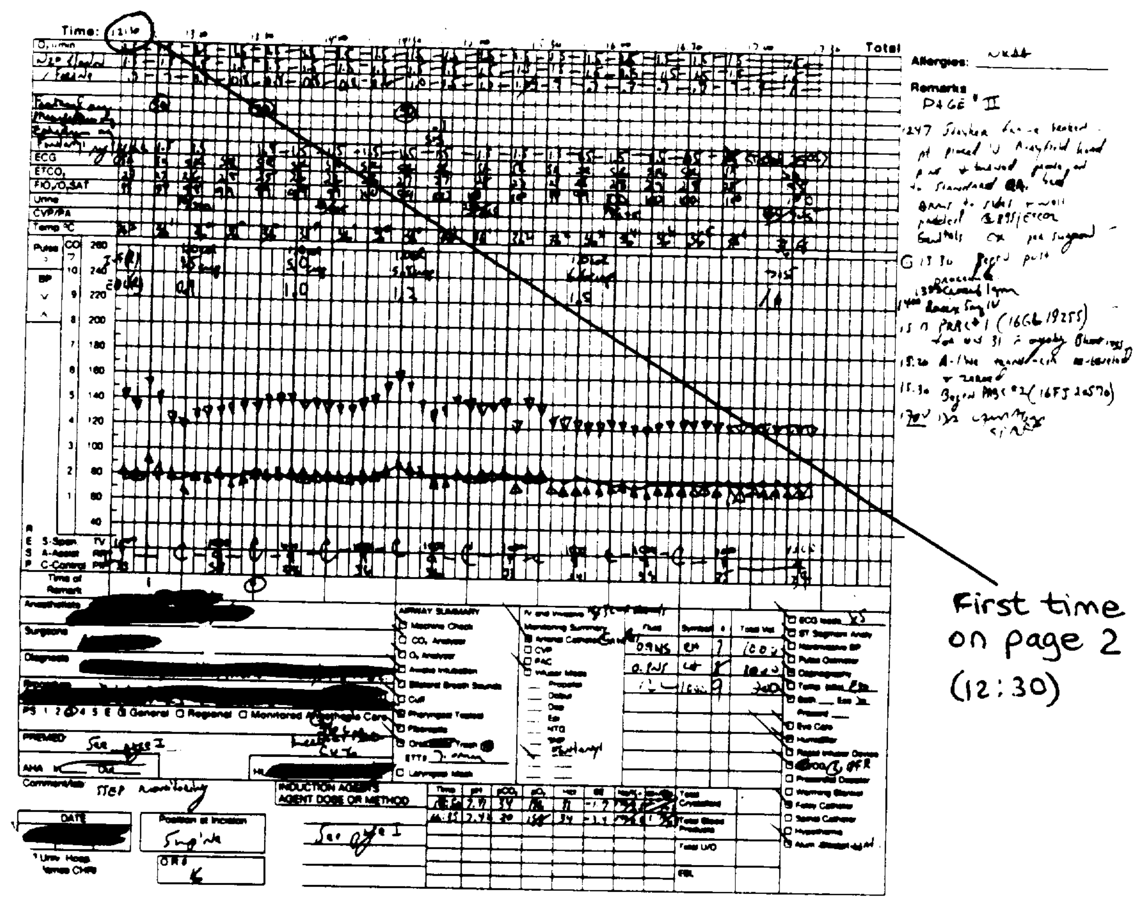

Figure 13b. Page 2 of the anesthesia record. Note time of blood gas and first time on page.

Queries and Informative Responses

Team members sometimes request specific information about events or parameters, such as "What is the blood pressure?" as a surgeon might ask during an aneurysm clipping surgery. But other questions are more open-ended, such as some variant of "What's up?" which attendings may ask upon returning to the operating room. Some queries fall somewhere in the middle. An observation statement such as "temperature 35.2, eh?" in the context of an earlier episode, besides being a comment on the temperature being low, is interpreted by the resident to mean something like "tell me what you know about this parameter being low."

The point is that team members don't have to explicitly query about all the information they may need, because team members can respond appropriately to open-ended questions. This is important in the case of tutoring, when learners lack knowledge about what information they are missing (e.g., Miyake and Norman, 1979), but it is also relevant in updating situations. Indeed, even when team members generate specific queries about the process, an assessment or action, responders often do not simply answer the explicit question posed to them. They "go beyond" the question to provide what they deem an informative response. People are sensitive to the intentions and goals that requesters have when asking for information, and they answer accordingly (e.g., Pollack, Hirschberg and Webber, 1982). People can provide relevant information even though an information seeker does not formulate the question precisely.

One way that team members go beyond a minimal response is by elaborating. For example, consider the following question and answer pair:

Q: What is the blood pressure?
A: It's 110 over 60.

The answer given here might be called a *minimal response;* it answers the explicit question posed. But, it may be informative depending in some circumstances, to provide an answer that goes beyond a minimal response. For example, it may be important for the information seeker to find out about trend information; if the blood pressure has been highly erratic recently, or it is expected to go vary due to a recent action whose effect is not yet apparent. Knowing these factors may be useful and even critical to the team member requesting the information.

Consider the following example from a surgery involving a cerebral aneurysm clipping. Before the clipping, deliberately induced hypotension is generally used in order to minimize the chances for rupture, facilitate placement of the clip and also to reduce blood loss if bleeding occurs (Barash, Cullen and Stoelting, 1991). The anesthesiologist, because he is the team member who administers the drugs, must coordinate with the surgeon concerning the start, duration and degree of hypotension. Figure 14 shows two instances in which the surgeon asked the anesthesiologist about the blood pressure: right before the clipping and shortly after the clipping. The figure also shows the desired blood pressure values around the time of clipping. In the first asking instance, the anesthesiologist provides a minimal response plus a "tag" that informs the surgeon that the value should soon drop to the expected value. After the clipping, when the pressure is to be brought up again, almost the same minimal response is provided, along with a tag that, this time, specifies how the value is related to the normal value. This example illustrates that team members provide a more "complete picture" for the information seeker than that which would be provided by simply answering their explicit question. In answering queries about process data, providing an informative response means providing information about factors that will or might affect the value within a certain horizon of the future.

The ability to provide responses that are sensitive to the goals of the information seekers is important.

Another example of a unprompted elaboration is found in Figure 11. The attending seems surprised by the parameter value told to him by the resident. The resident's response is to provide further information about the accuracy of the cardiac output (i.e., "that was a combination of 3 outputs so it's pretty accurate"). He also provides information about another parameter value that is related to the one asked about (i.e., cardiac index). Such elaborations are commonly found in the data.

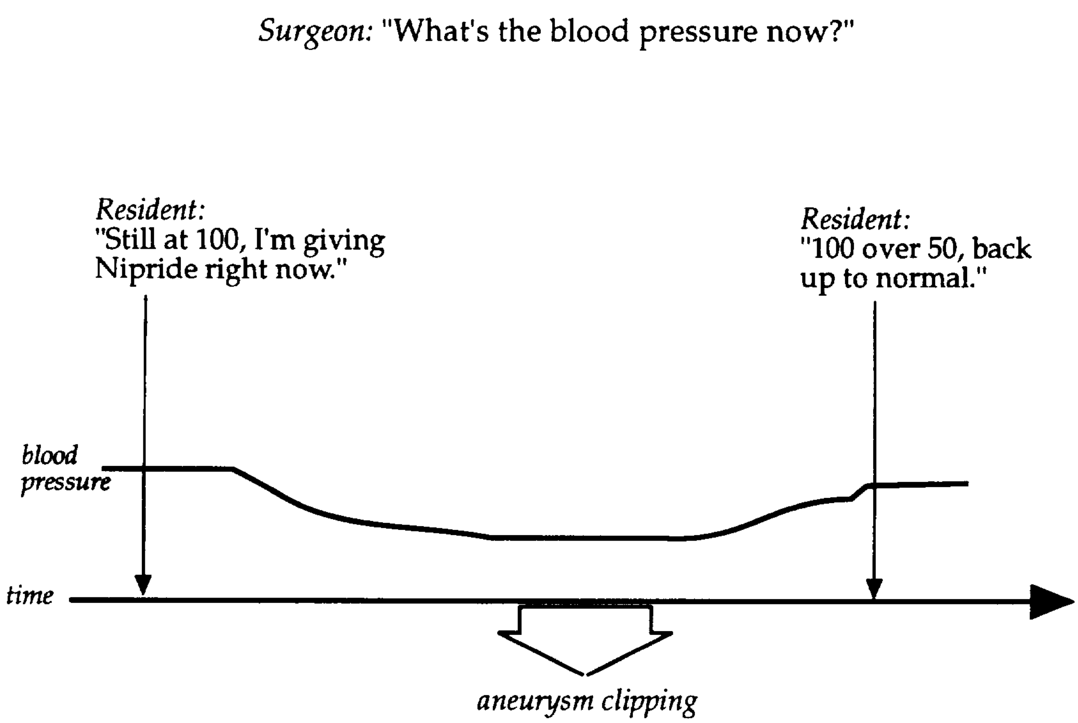

Figure 14. Different context-sensitive elaborations for the same query.

Another way that team members go beyond a minimal response is by providing responses that vary on the dimension of interpretability of process data. At one end there are statements of data values and at another end are interpretations of what they mean. Consider the following exchange:

[8 | 2:24:20]
S: Did he get a good diuresis?
R: I just emptied uh 350ccs or less, 25 minutes ago.
S: Huh, did he get a good diuresis?
R: 350 cc's (.) half hour.

In this case, the surgeon has on the surface of it asked a qualitative question. That is, the form of the question implies a judgement (e.g., was it "good" or "not"). However, the responder chooses to answer it quantitatively, in effect leaving the judgement to the surgeon.

Updating the Common Ground When a Team Member Returns

The trend in complex, dynamic fault management domains is to have fewer humans monitoring interacting subprocesses through the increased use of intelligent systems. Consequently, the human may be distanced from the process for relatively long periods of time. What is the nature of effective updates? The following episode is particularly interesting because the resident and attending engage in joint diagnosis.

Overview of Episode

The episode occurs during the maintenance phase of an operation to clip a cerebral aneurysm. The episode occurs about an hour after induction and before the surgeons have exposed the aneurysm. The senior resident is the only anesthesiologist present; the attending has been away for about half an hour and the junior resident is on break. In this episode the senior resident detects an anomaly -- bradycardia (very low heart rate). He takes corrective action by administering atropine, a drug that raises the heart rate. He has the attending paged. He mentions the event to the surgeons and enquires whether they "might have been doing anything." They answer no. The attending arrives after a few minutes and together they arrive at a diagnosis.

Anomaly Detection, Corrective Action and Investigation

To a practitioner, the bradycardia event is quite dramatic. The pulse rate as indicated by the beeping of the pulse oximeter suddenly slows down. The resident, who has bent down (apparently to check the urine output or to begin a cardiac output measurement), immediately gets up to look at the monitor. Five seconds later he injects the atropine. See Figure 15.

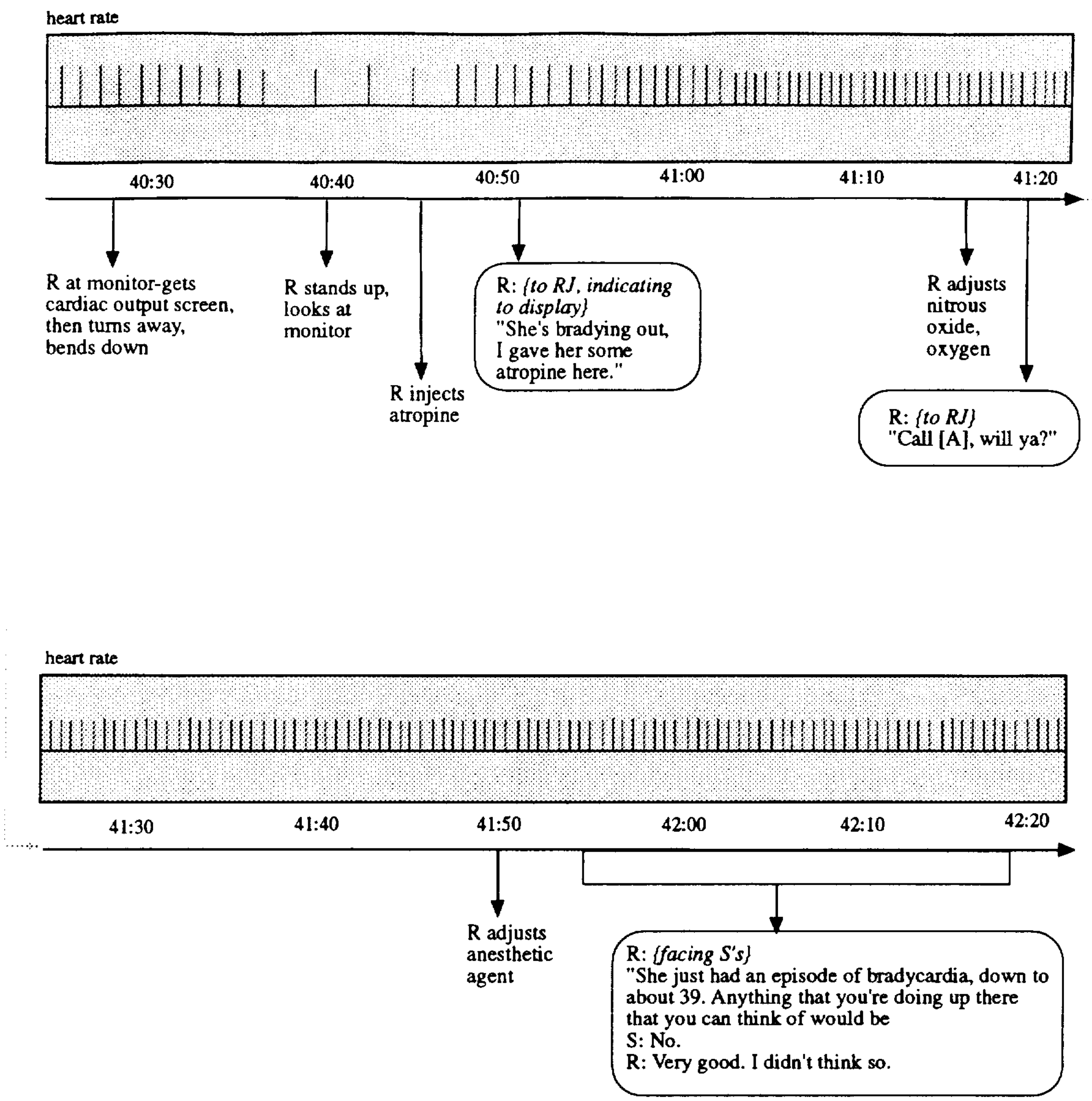

Figure 15. Context for Bradycardia Update.

Bradycardia may be expected in certain situations. For example, certain drugs given during maintenance can result in a lower than normal heart rate.[5] Also, a low heart rate indication could be expected in the case of a known artifact with monitoring equipment. However, in this case, bradycardia of such severity is unexpected. Because of its severity it is critical to treat it immediately, before its consequences begin to propagate. It is also important to understand its etiology because it could be an premonitory event, i.e., indicative of a fault that needs to be managed or corrected to prevent the condition from recurring or to prevent other possible disturbances. After investigating the surgeons' actions as a source of the event, the resident pages the attending to help him uncover the cause and also to make the attending aware of a potential premonitory event.

Update and Joint Problem Solving

Figure 16 shows what occurs when the attending arrives. Notice, first of all, that the resident answers the attending's open-ended query with a rather detailed account that includes several pieces of information. One of these is a related process event (less severe bradycardia) that occurred before the severe event. Notice that he provides information about the dynamics of the antecedent event, of the event itself, and of another relevant parameter (blood pressure).[6] He mentions what action he was taking on the process while the event occurred, the dynamics of the event, the limiting values reached and the corrective action he took and the process response to it. Finally he informs the attending about the state of problem solving, i.e., that he has no explanation. He has rejected one hypothesis (i.e., "nothing [the surgeons] were doing"), though he doesn't elaborate. At the end of the initial update, the attending queries him on this point. The resident's response is the same, unelaborated.

Notice the form of the initial update. The resident recounts it like a story, basically preserving the order of events. Such a recounting would seem to benefit causal analysis.

At this point, the state of problem solving seems to reach an impasse (i.e., when the attending says that he "can't necessarily explain that.") However, the resident continues the problem solving by telling the attending about various management influences on the process (i.e., drugs being given). He then engages in hypothesis discounting--mentioning a few potential (incomplete) hypotheses and providing reasons for discounting them.

The attending then lists what the causes of this kind of event have been in his past experience. In reaction to this, the resident seems to reevaluate the data that fed into his conclusion that it could not have been due to the surgeons. He "revisits" what, based on the attending's knowledge, seems to be an important time frame. He then point out in detail what was

[5]Halothane or large doses of morphine or fentanyl (Chung and Lam, 1990).

[6]Severe hypertension may cause bradycardia by a reflex pathway but the absence of high blood pressure rules out this mechanism.

occurring then--that it was actually when the surgeons were engaged in an activity that could have given rise to the event.

This example is used, not to suggest that this particular update is optimal in content or pattern. Rather, we use it to illustrate some characteristics of cooperative problem solving. One point is that diagnosis can be collaborative and cooperative rather than autonomous. In this episode the resident has access to the relevant data by having been present during the event, while the attending has access to more etiological knowledge. Both are essential for the appropriate diagnosis in this case.

Another characteristic illustrated by this problem solving episode is what one might call "robustness." It is robust in the sense that an initially discarded hypothesis is reintroduced and taken as the best explanation for the event. One aspect of the team interaction that seems important to its robustness is the ability of the resident to reexamine past data in the light of the attending's concerns.

Some implications for cooperative interaction in supervisor updates are suggested by this example. One is that an important characteristic of an intelligent subordinate team member is to be able to recognize that the situation is in danger of escalating beyond his or her (or its) competence, i.e., knowing when to call the supervisor. Secondly, the subordinate must be able to provide some kind of reconstruction of the event that emphasizes relevant events, actions and relationships in order to provide the supervisor with a coherent recounting of the events that led to the present state.

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE OF PROBLEM SOLVING
A: *(enters room)* Nice and tachycardic[1]	A comments on process	
R: Yeah, well, better than nice and bradycardic		
....		
A: What's going on guys?	A makes open-ended request for update	
R: *(takes end of printout, seems to show to A)* She had an episode of just kinda, all of the sudden bradying down to 50, 52 then came right back up, nothing they were doing, then all of the sudden out of the blue, I was shooting an output[2]and she dropped down to 32, 38[3] somewhere around there, pressure[4] dropped down to 60 so I gave her .5 of atropine[5] and ah, kicked her up to 6.5; she liked that, but no explanation. This is at 50 millimeters per second, twice the speed[6].	R mentions: -previous related event, including dynamics and approx values -discounting of other agents' activities as cause -action taken while event occurred - dynamics and approx values of relevant parameter during event -corrective action taken and process' response -no good candidate for diagnostic search R supplements description with artifact preserving data history	Initial update of significant event
A: They weren't in the head doing anything?	A requests specific past observation information (concerning other agents' activities) at time of event.	Hypothesis building

[continues]

[1]Tachycardia refers to rapid heart rate, while bradycardia refers to a slow heart rate.
[2] Cardiac output refers to the volume of blood per unit time that the heart moves. The measurement of cardiac output requires injection of a measured amount of IV fluid and is done infrequently.
[3] These are very low heart rate values, requiring treatment
[4] blood pressure.
[5] A drug that increases heart rate by blocking the parasympathetic system.
[6] Chart speed for EKG recording is usually 25 mm/sec. Because it's running at 50mm/sec, recorded events occupy twice the length of chart paper than they would at normal speed.

Figure 16. Bradycardia Update.

Figure 16 (continued),

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE OF PROBLEM SOLVING
R: Nothing.	R answer discounts hypothesis, but does not elaborate.	
A: Okay. Well I can't necessarily		
R: The only thing		
A: I can't necessarily explain that	A states has no candidates	
R: Yeah, neither can I. The only thing we're doing right now is just trying to open her up and fill her up. *(points to right IV tree)* She's up to a mic per kilo of nitro[7] and then she's still at the 5, started out at 3 and a half of dobutamine[8] and it did absolutely nothing, so I'm up to 5	R provides more information on current actions and previous actions	Context building
A: Okay		

[continues]

[7]Nitroglycerine. A vasodilator, for controlled hypotension.
[8] Dobutamine is generally given for low cardiac output, in order to increase contractility.

Figure 16 (continued),

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE OF PROBLEM SOLVING
R: So I don't know if she doesn't like contractility or, I can't think of anything else we're doing. The line went in perfectly normal, I can't imagine that she has a pneumo or anything that would be causing tension, her peak area pressures have not changed. Just all of the sudden -boom-out of the blue--her potassium is 3 point 3 and we're getting ready to replace that and we have been hyper-ventilating, but I don't know if low potassium can affect heart rate	R offers hypothesis but discounts based on his knowledge R offers another hypothesis but discounts it based on data Dynamics of event repeated Process variables mentioned, action to be taken mentioned R offers a third hypothesis but voices his lack of knowledge	Hypothesis discounting
A: Yeah, I don't know, I can't give you cause and effect on that. In my experience it's usually been stimulation of the trachea, it's something traction on the dura	A mentions two causes of the significant event based on his past experience	Case-based discussion
R: yeah, (absolutely)		
A: you know things		
R: yeah, it may have been dura	R remarks that one of these causes may have been cause in this case	
A: ...sort of a reflex, pressure on an eye	A provides another possible cause based on past cases	
R: *{animated}* Actually it was when they were sawing the dura open.	R remarks that event occured during a time when one of the causes mentioned by A could have occured	Discounted hypothesis reconsidered
A: well that's		

[continues]

Figure 16 (continued),

TRANSCRIPT	DOMAIN INDEPENDENT DESCRIPTION	PHASE OF PROBLEM SOLVING
R: putting tension on it		
R2: traction on the dura		
A: you touch the dura you'll get that	A states mechanism	
R: okay		
A: cause the dura is ennervated by the fifth I believe, and it somehow makes its way back to the (.) ganglion, same thing that causes oculocardiac reflex	A describes mechanism whereby hypothesized cause leads to the significant event	
R: I'd be willing to bet you're absolutely right *{RS waves pen over ventilator setting knobs, then leaves view}*	R expresses confidence for hypothesis	Hypothesis acceptance
A: is the same mechanism whereby you get (bradycardial traction) on the dura, so my guess is that's exactly what it was	A continues explanation of mechanism	
R: Okay.	R concurs (with hypothesis)	
A: you now and for future reference, if you suspect *{pause}* this lady's probably not going to mind this experience because she, we don't think she's really significantly sick, we're being a little overly cautious with her, my preference is, if you have a patient that you think has a bad heart, and you think they have a vagal problem via traction, or an eye...		
RS: so that's why		
A: It's traction on the dura		

CHAPTER V
DISCUSSION

It has been a consistent observation that advisory systems for dynamic fault management provide diagnoses and explanations that congregate at busy times, creating extra tasks and cognitive burdens for operators (Woods et al., 1991; Malin et al., 1991). This field study provides a clearer picture of why this is so: it indicates that human team members, by contrast, support one another in maintaining an updated understanding of the evolving situation.

The term "clumsy explanation" was used to refer to a form of diagnostic support that is not well adapted to the cognitive demands of dynamic fault management. This occurs when explanations are dissociated from process data both temporally and spatially. Some of the characteristics of classic explanation that impede a cooperative interaction in a diagnostic, evolving situation are its retrospective, one-shot nature, its context-independence and its expert-to-novice relationship. The cognitive implications are that the explanation occurs as an interruption to on-going lines of thought, increases workload by requiring the practitioner to read and understand an explanation at a busy time, to engage in interface management tasks and to integrate the system's information with process data.

In contrast, interpretations and explanations among team members engaged in dynamic fault management are not given in one long chunk from one agent (problem solver) to another (problem holder). Team members engage in joint interpretations in which both are involved in a process of coming to a mutually held interpretation. The strategy among human team members is to maintain a common ground as the situation evolves. They assist one another in maintaining an up-to-date interpretation in several ways. For example, they draw attention to anomalies, events and parameters of concern and they speak about them relative to expectations. Team members also provide informative responses, i.e., with elaborations tailored to the information needs in the current context. Communication among human team members, like conversation in general, reflects a sensitivity to what is informative and relevant to others (Grice, 1975). Team members also provide unprompted communication of relevant activities (i.e., their influences on the process) and assessments. They talk about strategies and evaluate the effects of past interventions. These communications provide a context in which information takes on meaning. This articulation of strategies and expectations among team members has been noted in a simulator study of coordinated activity in aircraft flight crews (Orasanu, 1990) and investigation of coordination between airline operatios center and central flow control in air traffic management (Smith, Orasanu, McCoy et al., 1994).

In both cases, the investigators find that coordinated activity requires investment to build a shared model for the situation and the perspectives of the different agents. Similar results have been obtained in other studies of coordinated activity across people (e.g., Hutchins, 1990) and in particular in studies of cockpit resource management (e.g., Foushee and Manos, 1981).

A result of the role of the common ground in communications is that it is unusual for team members to ask "why do you say that?" or "why did you do that?" These questions, which express a need for explanations, indicate a rift in the common ground. Breakdowns in cooperative interaction between pilots and cockpit automation are marked by just these questions: what is the automation doing? why is it doing that? what will it do next? (Wiener, 1989). These breakdowns in cooperation between people and automated systems have been linked to weak feedback about the current and anticipated activities of the automated system (Sarter and Woods, 1995) Such "strong but silent" systems do not function as team players (Woods, in press a). Similarly, studies of human interaction with intelligent system indicate the need for a common frame of reference to support true cooperation between the human practitioners and AI advisory system (e.g., Roth et al., 1987).

Why Invest in the Common Ground?

In general, team members invest heavily in communicating about the state of the monitored process and the state of the problem solving process. There are several good reasons for them to make this investment in the common ground. One is that diagnosis entails disentangling the various influences acting on the monitored process, some of which may be due to the interventions of other team members (human and machine agents). Hence, it is important for team members to assist in keeping one another aware of their interventions on the process.

At another level, an important reason to invest in the common ground is to help keep other team members in a state of readiness so they are able to assist in the management and diagnosis of faults in the process. The same level of effort to keep someone updated is not warranted if they are not true team members. This is reflected in an episode found in one case in which an update to a medical student was cut short in order to deal what was perceived to be a more pressing task.

Another important function of maintaining a common ground is that it can allow for more efficient communication during higher tempo periods; less needs to be said because information can be communicated relative to what is already mutually known.[1] This is consistent with Orasanu's (1990) findings concerning the temporal-sensitive nature of communication among

---

[1] For example, grounding allows references to the same item to become more concise during the evolution of a communication task (Clark & Wilkes-Gibbs, 1990; Krauss and Weinheimer, 1964).

cockpit crew members; she found that captains in high performing crews talked less than captains in low performing crews *when workload was high;* also, the captains of high performing crews requested slightly less information during abnormal phases of flight, whereas captains of poor performing crews requested more information during these phases.

Establishing common ground can make the need for retrospective explanations of assessments or actions less necessary. This is useful because such explanations would be resource-consuming at high-tempo, high-criticality times, when concentration needs to be devoted to understanding the process behavior, rather than in mending a problem in a team member's understanding (recall the episode in which the attending puts off an explanation of his decision until a more opportune time). In this purpose, maintaining common ground is similar to anesthesiologists' preparatory or anticipatory behaviors (Cook et al., 1991; Xiao, 1994), i.e., a task undertaken at the moment, so that things will be easier later on, when they can be expected to be more busy.

Implications for Human-IS Cooperative Interaction

Much research on artificial intelligence explanation assumes that explanation is linguistic. Swartout and Moore (1993) state that for an expert system to generate good explanations, it must meet the desideratum of linguistic competence, i.e., "it must be able to construct a coherent multisentential text to achieve a communicative goal." This aspect of more sophisticated explanations is still in the research stage. The danger of intelligent systems for supporting dynamic fault management that incorporate a limited degree of linguistic competence is that they would end up being "chatty" and distracting.

The study presented herein supports the notion that distributed dynamic fault management relies on maintaining a common ground. We suggest that this is a conclusion at a competence level of analysis; that is, all team players engaged in dynamic fault management, whether human or machine need to invest in the common ground. *How* this is to be done is a separate issue.[2]

Even though the metaphor of a "conversational system" is problematic for dynamic fault management, there are certain lessons that may be drawn from human-human communication. This is basically Brennan's (1990) point when she says that direct manipulation interfaces succeed because they share important features with conversation. Brennan points out that general strategies from human-human conversation apply to human-computer interaction design; these include: provide feedback (akin to the "backchannels" of conversation), have ways of establishing that

[2] The distinction I am making here is like the distinction Marr (1981) made between a computational theory (that specifies the goals to be achieved and the logic of the strategy) and the representation and algorithm level of description (that specifies how the theory can be implemented).

understanding is sufficient for current purposes, and assume errors will occur and provide ways to repair.

For dynamic fault management, some competence level principles can be listed (though these might be implemented in various ways):

Team members limit the need for others to search for information. Human team members assist one another in finding the right information at the right time: they direct attention to relevant events, and provide unprompted information about activities or assessments and context-sensitive elaborations to queries (Woods, in press b). By contrast, many support systems have characteristics that create cognitive burdens associated with retrieving relevant information, e.g., interface designs that force serial access to highly related data (Cook and Woods, 1995). The problem is that these systems provide an avalanche of data within which it is difficult to find the relevant information (recall the "man page" approach to explanations; Mastaglio and Reeves, 1992). This is particularly problematic in dynamic fault management applications because cognitive demands increase with the tempo and criticality of operations (Woods, 1994).

Team members communicate without distracting. Communication among team members generally occurs while team members are engaged in activity. The communications of team members are not a break in the flow of activity; in the usual case, it is not necessary for team members to drop what they are currently doing in order to gain information from another team member.

Team members communicate in various shared contexts. A salient characteristic of team member communications is their "compactness." By compactness we mean that a phrase, word or gesture is packed with meaning -- meaning that would generally not be extractable by a lay person, without extra information or explanation.[3] Mutual knowledge of various kinds allows for this compactness. This mutual knowledge or mutual potential knowledge can be viewed as different kinds of shared context within which communication occurs. These shared contexts of various kinds are simultaneously available. The first notable shared context is *shared domain knowledge*. The team members share domain knowledge about the subject matter and practice of anesthesiology, which allows them to understand for example, what a phrase like "taking a gas" means. It allows one to understand why the attending might say "Let's give him some dobutamine" and how to

[3]Interestingly, reference can be so compact that it involves neither words nor direct pointing. In one episode observed, a medical student elicits an explanation of the resident by "waving" towards the vital signs display. The resident turns to look at the monitor and states "cause the cuff is up. That's the pulse oximeter." Of all that is on the vital signs display, the resident picks out the flattened pulse oximeter waveform as the reference. From the resident's point of view, the flat waveform is expected because the blood pressure cuff was on the same arm as the pulse oximeter monitor; whenever the cuff inflates it squeezes off blood flow, which leads to a spurious pulse reading. However, it is the atypical item – that which would be anomalous in another context. The reference is understood partly because of the critical role of anomalies in dynamic fault management.

take this action, or what "Why don't you put the A-line in" means, why it would need to be done and how to do it.

Another context is *shared local knowledge*, that is, shared knowledge about how the team, or particular team members, tend to do things that can be done in more than one way. Often there multiple "correct" ways to do something, and the department or team may have particular ways of doing them, e.g., the default induction drug to use. Also, team members may have different "styles." For example, the data showed attendings varied in their approach to drug dosage or fluid replacement therapy; these variations are stable and are recognized by other team members[4].

Another context is the *shared temporal context.* This refers to knowledge about the history of the process, including what interventions were taken, what the evolution of the state of the process has been and of problem solving. A brief statement like "pressure's 100" in an update gains its significance (i.e., is this expected, normal, should we do anything to intervene?) depending on factors established in the past course of the case. These factors include: whether the patient is a chronic hypertensive, whether certain drugs have been given, whether certain events have occurred or are about to occur (i.e., an aneurysm clipping typically requires that hypotension be induced immediately prior to the clipping). The shared temporal context (coupled with shared domain knowledge) allows such a statement to be understood qualitatively--as a state, e.g., either low or high, depending on the mutual knowledge of the case.

A fourth context is the *physical context* which consists of both the *task environment* and the *set of available monitored process representations.* Communicating within the context of the same physical environment means that grounding is less costly because the constraints of copresence, visibility, audibility, and cotemporality are present (see Chapter 2). These constraints allow team members to ground without explicit informing; information is available about what other team members do through peripheral access--being able to see what others do, even though one is not explicitly monitoring for it. The other aspect of the physical context concerns the monitored process views. The transcripts showed that team members often talk about interpretations of the process while looking at displays and pointing. Pointing (deitic reference) makes for compact communication --pointing to some item on the display can substitute for a description or an explanation in some situations. Certain representations can provide a wealth of other information e.g., analogical, trend information, that can be had "for free" when using deitic reference.

Agents vs. Tools

[4]In case 2, the senior resident tells the junior resident "[A1] likes to fill them up, [A2] likes to keep them dry."

The need to maintain a common ground means that a "dark board" strategy, in which the intelligent system draws attention (i.e., sounds alarms) only when "something is wrong" is inappropriate model for communication between human practitioners and intelligent systems (Woods, in press b). What, then, would be an effective approach for conveying the IS's relevant assessments and actions on the process? How can common ground be maintained between human and intelligent system?

People import expectations from human-human communication in their interactions with machines (Suchman, 1987). Because of the opacity and pseudo-animacy of such systems--they seem to take actions of their own accord[5] -- human partners may be wont to take an intentional stance towards them and to imbue them with more intelligence than they deserve (Woods, in press a). Norman (1990) points out that current automation has an intermediate level of intelligence; it is smart enough to take actions and offer assessments, but not smart enough to act to handle all abnormalities and to provide the continual, appropriate feedback found among human operators. Norman provides an example of an autopilot that fails to provide feedback that it has reached the limit of its compensatory ability. He points out that for a system to be able to inform team members about this state of affairs, it would require a "higher-level of awareness, a monitoring of its own monitoring capabilities." To be able to do this in the general case, requires a degree of intelligence that is not yet been attained in the research labs.

Rather than developing the agent-like or stand-alone properties of machines, another approach for creating joint human-machine cognitive systems is to design them as tools to support practitioners in their field of activity (Roth, Bennett and Woods, 1987; Woods, 1993). In this view, information from the AI system is another form of data to be integrated with other raw data in supporting situation assessment. In this approach, understanding of the AI system's activities and assessments is supported, not by more sophisticated linguistic explanations, but by making functioning apparent.[6]

In the Context of the Monitored Process

The critical desiderata for diagnostic support is that it be "in synch" with the tempo of activity, efficient and not distracting. An important aspect of this is that assessments be integrated in the context of the monitored process views. Intelligent system assessments and activities that are dissociated from the process views can lead to extra cognitive tasks (Potter, and Woods, 1991; Remington and Shafto, 1990). Studies indicate that, when

[5]The aircraft flight management system (FMS) responds to operator inputs as well as to situational and system factors. For example, the FMS initiates a mode transition when a preprogrammed intermediate target is attained (Sarter and Woods, 1993).

[6]This is related to Suchman's (1987) point about an artifact being self-explanatory in two senses: 1) it can explain itself as a human might do when queried, or 2) its functioning or use can be easily discoverable.

communication demands with the machine agent are high, particularly during high tempo periods, practitioners will abandon cooperative strategies and switch to single-agent strategies (Woods, Johannesen, Cook and Sarter, 1994). Continuous display of AI system reasoning, if dissociated from the monitored process, is still likely to be distracting because it would draw the operator's attention away from the process.

Integrating IS assessments into the process views isn't simply about spatial contiguity; it means making the basis on which the diagnostic information is generated, apparent. One source of information that should be made apparent in the representation is what process data is being used (e.g., Roth, Butterworth and Loftus, 1985). Another important source of information used by model-based intelligent systems is the context-sensitive expected values generated for critical parameters. While this model-based capability is used for diagnosis, it could also be used to provide a context for data presentation to the human team members --so that team members know what referents, expectations, and predictions are being used in the IS's assessments. Expectations are important to convey because they set up the contrast cases for explanations. A common frame of reference is supported if these expectations are made apparent.

Common Frame of Reference

The common frame of reference concept arises from work in distributed (multi-agent) problem solving that indicates breakdowns occur when multiple agents (some of which may be machines) engaged in problem solving do not have access to the state of the problem or the problem solving approach taken by other collaborating agents (Roth, Bennett and Woods, 1987; Hutchins, 1990; Suchman, 1987). The common frame of reference is about the resources that allow for a common and accurate understanding of the state of the process and state of problem solving across team members.

As mentioned earlier, there are typically many sources of data about the process, some of which may be processed by an "intelligent agent." When there is no common frame of reference, data will be available piecemeal, relevant relationships will not be emphasized, and data will be divorced from its context. It can lead team members to form potentially diverging and inaccurate ideas about the state of the process and of the problem solving process. It can also lead to increased cognitive workload because practitioners are required to integrate various sources of data. Also, when the costs involved in coordination with a decision/problem solving support system are too high, practitioners have been known to abandon cooperative strategies, abandon use of the support system or constrain its use in severe ways (Woods, 1993a, Remington and Shafto, 1991; Cook and Woods, 1995). The common frame of reference concept expresses the need in effective distributed problem solving of integrating information into a single framework that highlights relationships among the data and places data in the context of assessments of current and expected states of the process. Figure 16 attempts to illustrate the common frame of reference idea.

Effective joint problem solving requires resources that allow for a common accessibility of the problem state and of the problem solving approach (Roth, Bennett and Woods, 1987). What this means for dynamic fault management is that it is important to have: 1) accessibility to the problem solving approach-- the capability to know what hypotheses have been considered and rejected, to know the relevant assessments and activities of others, and 2) accessibility to relevant information about the monitored process-- the capability to become aware of the relevant information at the right time concerning the process. Relevant information means information that is tailored to the interests and expectations of the observer.

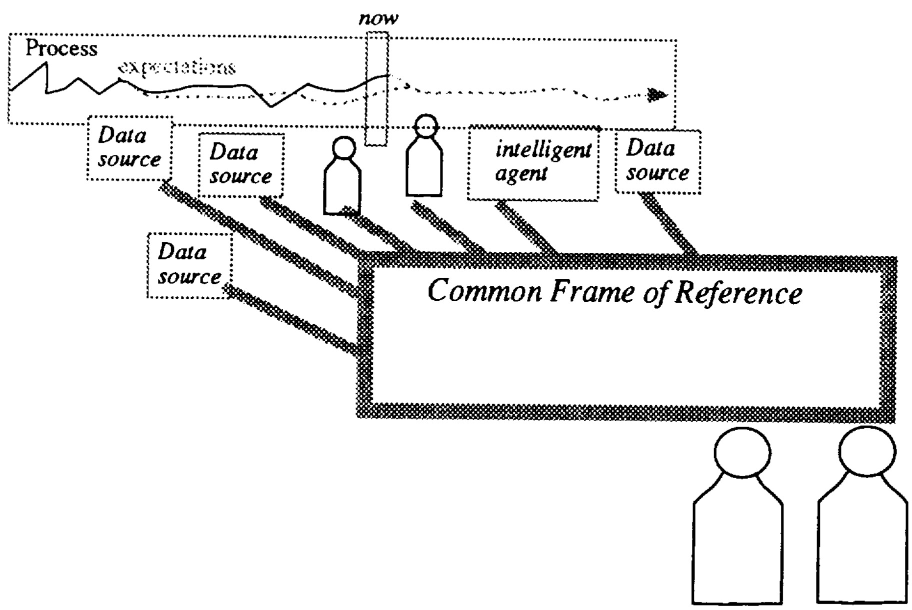

Figure 16. A Common Frame of Reference.

An important aspect of a common frame of reference is that it supports economical attention-directing reference or "joint reference." (Woods, in press b). This relies on there being some external representation of the conditions and events in the referent that is available to all agents. A shared external representation allows for "shared mindset across the cooperating agents about the background field against which the agents can all recognize interesting conditions or behaviors." (Woods, in press b).

Integrating IS and process views entails designing effective monitored process views that can be well coordinated with the IS's assessments and information about its actions. Potter et al. (1994) indicate how a function-based display can be designed to use the IS's computational power to help the operator visualize the behavior of the monitored process. The function-based display integrates the results of the IS's computations into the display of the monitored process state, and so creates a shared frame of reference (Woods, and Roth, 1988). Another representation that creates a shared frame of reference is an event-driven timeline display (Potter and Woods, 1991). This representation spatially segregates messages in terms of monitored process events and intelligent system assessments, while highlighting relationships among the events and the intelligent system assessments and placing them in the context of a timeline, so that temporal relationships are apparent.

Integrating IS assessments and activities into the context of the monitored process is a promising way, given the capabilities of current technology, to avoid retrospective explanations. In this way, when the IS presents assessments, they make sense to the human supervisor because of the previous context, rather than appearing "out of the blue," as a surprise to be investigated.

Future Directions for Research

This research has been a first step towards understanding the nature of team member support in dynamic fault management and the role of explanation. To further understanding of how team members establish common ground, it would be useful to refine the conditions under which they tell one another about assessments and activities. For example, in updates it would be useful to be able to predict what parameters team members will mention and how they will inform others about them given the case's history. We can form some preliminary hypotheses in this regard, based on some of the findings from the study. For example, one hypothesis is that updaters will call attention to certain parameters that have become anomalous, that have continued to be anomalous and parameters that have gone from anomalous to normal during the team member's absence.

A factor to investigate more deeply is the relationship between the nature of team member assistance and team member roles and relationships. What patterns of diagnostic support are found, for example, among NASA flight controllers in mission control? The roles and relationships among team members are different from those among the attending and the residents; in mission control there are several flight controllers, each with a highly specialized area of expertise, who support a flight director (supervisor) in making high level decisions.

The findings of this study indicate several ways in which team members maintain a common ground. Another important issue to investigate is how team members detect gaps and and repair "rifts" in the common ground. A simulator study in which particular scenarios can be created may offer a useful approach at this stage.

APPENDIX

LIST OF CASES

CASE	TYPE	LENGTH	RESIDENTS
1	Laminectomy	10 hrs.	1
2	Cerebral aneurysm clipping	5 hrs	2
3	Brain tumor	5 hrs	1
4	Laminectomy	4 1/2 hrs	2
5	Laminectomy	4 hrs	2
6	Laminectomy	3 hrs	2
7	Cerebral aneurysm clipping	6 hrs	2
8	Cerebral wound infection	5 1/2 hrs	1
9	Laminectomy	2 3/4 hrs	2
10	Laminectomy	8 hrs	2

BIBLIOGRAPHY

Aaronson, A. and Carroll, J. M. (1987). The answer is in the question: A protocol study of intelligent help. *Behavior and Information Technology, 6 (4),* 393-402.

Bainbridge, L. (1987) Ironies of automation. In J. Rasmussen, K. Duncan and J. Leplat. [Eds.]*New technology and human error.* Chichester: John Wiley & Sons.

Barash, P. G., Cullen, B. F. and Stoelting, R. K. (1993). *Clinical anesthesia.* Philadelphia, PA: Lippincott.

Bentley, R., Hughes, J. A., Randall, D., Rodden,T. Sawyer, P. Sommerville, I., and Shapiro, D. (1992). Ethnographically-informed systems design for air traffic control. *Computer-Supported Cooperative Work Proceedings.* November.

Billings, C. (1991). *Human-centered aircraft automation: A concept and guidelines.* NASA Technical Memo 103885.

Brennan, S. (1990). Conversation as direct manipulation: An iconoclastic view. In B. Laurel [Ed.]*The art of human-computer interface design.* Boston: Addison Wessley.

Cawsey, A. (1991). Generating interactive explanations. *Proceedings of the Ninth National Conference on AI.* AAAI-91.

Cawsey, A. (1992). *Explanation and interaction.* Cambridge, MA: MIT Press.

Cawsey, A., Galliers, J., Reece, S. and Jones, K. S. (1992). The role of explanation in collaborative problem solving. P. Brezillon. [Ed.] *Proceedings of the ECAI-92 Workshop.* "Improving the Use of Knowledge-based Systems with Explanations."

Chandrasekaran, B., Tanner, M. C. and Josephson, J. (1989). Explaining control strategies in problem solving. *IEEE Expert, 4 (1).*

Chapanis, A. (1978). Interactive human communication: A few research answers for a technological explosion. Paper presented at the *Annual*

Meeting of the American Psychological Association. Toronto, Canada. (August).

Clark, H. H. and Brennan, S. E. (1991). Grounding in communication. In L. Resnick, J. Levine, and S. Teasley. [Eds.] *Socially Shared Cognition.* Washington DC: American Psychological Association.

Clark, H. H. and Schaefer, E. F. (1989) Contributing to discourse. *Cognitive Science 13,* 259-294.

Clark, H. H. and Wilkes-Gibbs, D. (1986) Referring as a collaborative process. *Cognition, 22,* 1-39.

Cohen, P. (1984). The pragmatics of referring and the modality of communication. *Computational Linguistics, 10(2),* 97-146.

Cook, R. I. and Woods, D. D. (1995) Adapting to new technology in the operating room. Manuscript submitted for publication.

Cook, R. I., Woods, D. D. and McDonald, J. S. (1991). Human performance in anesthesia: A corpus of cases. Cognitive Systems Engineering Laboratory, Technical Report 91-TR-03, Columbus, OH: Ohio State University.

Coombs, M. J. and Alty, J. L. (1980). Face-to-face guidance of university computer users II: Characterizing advisory interactions. *International Journal of Man-Machine Studies, 12,* 407-429.

Feiner, S. and McKeown, K. (1990). Coordinating text and graphics in explanation generation. *AAAI.*

Foushee, H. C. and Manos, K. L. (1981). Information transfer within the cockpit: Problems in intracockpit communications. In C. E. Billings and E. S. Cheaney [Eds.], *Information transfer problems in the aviation system.* (NASA Technical Paper 1875). Moffett Field, CA: NASA Ames Research Center.

Galegher, J., Kraut, R. E. and Egido, C. [Eds.] (1990). *Intellectual teamwork social and technological foundations of cooperative work.* Hillsdale, NJ: Erlbaum.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan [Eds.], *Syntax and Semantics, 3,* New York: Academic Press.

Grosz, B. J. (1981). Focusing and description in natural language dialogues. In A. K. Joshi, B. L. Webber, and I. A. Sag. [Eds.], *Elements of discourse understanding*. Cambridge: Cambridge University Press.

Hasling, D. W., Clancey, W. J. and Rennels, G. (1984). Strategic explanation for a diagnostic consultation system. *International Journal of Man-Machine Studies 20*, 3-19.

Haviland, S. and Clark, H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512-521.

Heath, C. and Luff, P. (1992) Collaboration and control: Crisis management and multimedia technology in London underground line control rooms. *Computer Supported Cooperative Work (CSCW) 1*, 69-94.

Heritage, J. (1988). Explanations as accounts: a conversation analytic perspective. In Antaki, C. [Ed.], *Analyzing everyday explanation. A casebook of methods.*, London: Sage.

Hilton, D. (1990) Conversational processes and causal explanation. *Psychological Bulletin, 197 (1).*

Hollnagel, E., Pederson, O.M. and Rasmussen, J. (1981). Notes on human performance analysis. RISO National Laboratory, Roskilde, Denmark. Electronics Department.

Hughes, J., Randall, D. and Shapiro, D. (1992) Faltering from ethnography to design. *Computer Supported Cooperative Work.* November.

Hutchins, E. (1994). *Cognition in the wild.* MIT Press, Cambridge MA.

Hutchins, E. (1990) The technology of team navigation. In J. Galegher, R. Kraut, and C. Egido [Eds]. *Intellectual teamwork: Social and technical bases of cooperative work,* Hillsdale, NJ: Erlbaum.

Johns, G. (1990) Graphic interfaces to intelligent fault management systems: issues and guidelines. Mitre Technical Report. MTR-90W00103. NASA.

Jordan, B. and Henderson, A. (1994). Interaction analysis: foundations and practice. Institute for Research on Learning. IRL Report 94-0027.

Karsenty, L. and Falzon, P. (1992). Spontaneous explanations in cooperative validation dialogues. In P. Brezillon. [Ed.], *Proceedings of the ECAI-92*

Workshop. "Improving the Use of Knowledge-based Systems with Explanations."

Krauss, R. M. and Fussell, S. (1991). Constructing shared communicative environments. In L. Resnick, J. Levine and S. Teasley [Eds.], *Socially Shared Cognition.* Washington, DC: American Psychological Association.

Krauss, R. M. and Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction. A preliminary study. *Pyschonomic Society, 4,* 343-346.

Layton, C., Smith, P.J., McCoy, E. and Bihari, T. (1994). Design Concepts for the Development of Cooperative Problem-Solving Systems. *Human Factors, 35 (1),* 94-119.

Leake, D. (1991) Goal-based explanation evaluation. *Cognitive Science, 15,* 509-545.

Lewis, D. (1969). *Convention.* Cambridge, Mass: Harvard University Press.

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly. Vol 2 (4).*

Malin, J. and Schreckenghost, D. (1993). User as intelligent agent: Interface design for intelligent monitoring systems. *Conference on Artificial Intelligence Applications.*

Malin, J., Schreckenghost, D., Woods, D., Potter, S., Johannesen, L., Holloway, M., and Forbus, K. (1991) Making Intelligent Systems Team Players: Case Studies And Design Issues. NASA Tech Memo 104738.

Marr, D. (1982). *Vision: A compuational investigation into the human representation and processing of visual information.* San Francisco: W.H. Freeman.

Mastaglio, T. and Reeves, B. (1990). Explanations in cooperative problem solving systems. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society.* July 25-28, Cambridge, MA.

McCarthy, J. C., Miles, V. C. and Monk, A. F. (1991). An experimental study of common ground in text-based communication. *CHI '91.* ACM Press.

Miyake, N. and Norman, D. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18*, 357-364.

Moore, J. and Swartout, W. (1989). A reactive approach to explanation. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, MI, August 20-25.

Moore, J. and Swartout, W. (1990). Pointing: A way towards explanation dialogue. *Proceedings of the Eighth National Conference on AI.* AAAI-90. Boston.

Murray, C. and Cox, C. B. (1989). *Apollo: The race to the moon.* New York: Simon and Schuster.

Ng, H. and Mooney, R. (1990). On the role of coherence in abductive explanation. *Proceedings of the Eighth National Conference on AI.* AAAI-90. Boston.

Norman, D.A. (1990). The 'problem' of automation: Inappropriate feedback and interaction, not 'over-automation.' *Philosophical Transactions of the Royal Society of London*, B 327, 585-593.

Orasanu, J. M. (1990). Shared mental models and crew decision making. Princeton University Technical Report CSL 46. (October).

Orasanu, J. and Salas, E. (1991) Team decision making in complex environments. In G. Klein, J. Orasanu, R. Calderwood, and C. Zsambok [Eds.],*Decision making in action: Models and methods.* New Jersey: Ablex.

Paris, C. (1988). Tailoring object descriptions to a user's level of expertise. *Computational Linguistics (Special Issue on User Modeling), 14 (3)*, 64-78.

Pavard, B., Salembier, P., de Medeiros, E. and Benchekroun, T. (1989). Simulation cognitive de la conduite de proces: Evaluation de differentes approaches. Laboratoire d'Ergonomie, Technical Report ARD-AID T5403T, Conservatoire National des Arts et Metiers, Paris, France.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies.* New York: Basic Books.

Pollack, M., Hirschberg, J. and Webber, B. (1982) User participation in the reasoning processes of expert systems. *Proc. AAI '82*, 358-361.

Potter, S. and Woods, D. (1991). Event-driven timeline displays: Beyond message lists in human-intelligent system interaction. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics.*

Potter, S. and Woods, D. (1994). Breaking down barriers in cooperative fault management: Temporal and functional information displays. Cognitive Systems Engineering Laboratory, Technical Report, 94-TR-02.

Remington, R. W. and Shafto, M. G. (1990). Building human interfaces to fault diagnostic expert systems I: Designing the human interface to support cooperative fault diagnosis. In *CHI '90 Workshop on Computer-Human Interaction in Aerospace Systems.* Seattle, WA., April 1-2.

Roth, E. M., Bennett, K. B., & Woods, D. D. (1987). Human interaction with an "intelligent" machine. *International Journal of Man-Machine Studies, 27,* 479-525.

Roth, E. M., Butterworth, G. and Loftus, M. (1985). The problem of explanation: Placing computer generated answers in context. *Proceedings of the Human Factors Society*--29th Annual Meeting.

Sacks, H. and Schegloff, E. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas [Ed.], *Everyday language: Studies in ethnomethodology.* New York: Irvington.

Sarter, N. and Woods, D. D. (1994). Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the flight management system. *International Journal of Aviation Psychology, 4(1),* 1-28.

Sarter, N. and Woods, D. D. (1995). "How in the world did we get into that mode?" Mode error and awareness in supervisory dontrol. *Human Factors.* 37(1), in press.

Searle, J. (1992). *(On) Searle on Conversation.* Philadelphia: John Benjamins Publishing Co.

Segal, L. (1994). Actions speak louder than words: How pilots use nonverbal information for crew communications. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting.* Oct. 24-28. Nashville, TN.

Sheridan, T. (1988). Task allocation and supervisory control. In M. Helander, [Ed.], *Handbook of human-computer interaction.* North Holland: Elsevier Science.

Smith, P., McCoy, E., Orasanu, J.,et al. (1994). Cooperative Problem-Solving in the Interactions of Airline Dispatchers with ATCSCC. Cognitive Systems Engineering Laboratory Technical Report, 94-TR-04, Ohio State University, Columbus OH.

Sperber, D. and Wilson, D. (1986). *Relevance. Communication and cognition.* Cambridge, Mass: Harvard University Press.

Stalnaker, R. C. (1978). Assertion. In P. Cole [Ed.], *Syntax and semantics: Pragmatics, 9.* New York: Academic.

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

Suchman, L. (1990). What is human-machine interaction. In S. P. Robertson, W. Zachary and J. B. Black [Eds.],*Cognition, computing and cooperation.* Norwood, NJ: Ablex.

Suchman, L. and Trigg, R. (1993). Understanding practice: video as a medium for reflection and design. In Baecker, R. [Ed.], *Readings in groupware and computer-supported cooperative work. Assisting human-human collaboration.* San Mateo, CA: Morgan Kaufmann.

Sutcliffe, A. G., Maiden, N.A.M., Cooper, M. A. (1992). Information seeking and discourse structure in explanation. *Samos Workshop on Task-Based Explanation.* June 28-July 1. Research Lab of Samos, University of the Aegean.

Suthers, D., Woolf, B. and Cornell, M (1992) Steps from explanation planning to model construction dialogues. *Proceedings of the 10th National Conference on AI.*

Swartout, W. R. and Moore, J. D. (1993). Explanation in second generation expert systems. In J.-M. David, J.-P. Krivine and R. Simmons. [Eds.] *Second generation expert systems.* New York: Springer-Verlag.

Toulmin, S. (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Turnball, W. (1986). Everyday explanation: The pragmatics of puzzle resolution. *Journal for the Theory of Social Behaviour, 16,* 141-160.

Turnball, W. and Slugoski, B. R. (1988). Conversational and linguistic processes in causal attribution. In D. J. Hilton [Ed.], *Contemporary Science and Natural Explanation.* New York: New York University Press.

Weiner, E. (1989) Human Factors of advanced technology ("Glass cockpit") transport aircraft. Technical Report 117528. Moffett Field, CA: NASA Ames Research Center.

Woods, D. D. and Roth, E. M. (1988). Aiding human performance II: From cognitive analysis to support systems. *Le Travail Humain, 51 (2),* 139-171.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L.P. Goodstein, H.B. Andersen, and S.E. Olsen, [Eds.], *Mental models, tasks and errors.* London: Taylor & Francis.

Woods, D. D. (1989). The effects of automation on the human's role: Experience from non-aviation industries. In S. Norman and H. Orlady, [Eds.] *Flightdeck automation: Promises and realities,* NASA Ames Research Center, Moffett Field CA.

Woods, D. D., Potter, S., Johannesen, L., and Holloway (1991). Human interaction with intelligent systems. Vol. 1. Trends, problems and new directions. CSEL Report 1991-001. The Ohio State University.

Woods, D. D. (1993a). The price of flexibility in intelligent interfaces. *Knowledge-Based Systems,* 6, 1-8.

Woods, D. D. (1993b). Process tracing methods for the study of cognition outside of the experimental laboratory. In G. A. Klein, J. Orasanu and R. Calderwood [Eds]. *Decision making in action: Models and methods.* New Jersey: Ablex.

Woods, D. D. (1994). Cognitive demands and activities in dynamic fault management: Abductive reasoning and disturbance management. In N. Stanton (Ed.), *Human factors of alarm design.* London:Taylor & Francis.

Woods, D. D., Johannesen, L., Cook, R. I. and Sarter, N. (1994). *Behind human error: Cognitive systems, computers and hindsight.* CSERIAC (Crew System Ergonomics Information Analysis Center), Dayton OH.

Woods, D. D. (in press a). Decomposing Automation: Apparent Simplicity, Real Complexity. In R. Parasuraman and M. Mouloua, editors, *Automation and Human Performance,* Erlbaum.

Woods, D. D. (in press b). The alarm problem and directed attention in dynamic fault management. *Ergonomics*.

Xiao, Y. (1994). Interacting with complex work environments: A field study and a planning model. Unpublished Dissertation. Dept. of Industrial Engineering, University of Toronto.