To Be Published In J. Cannon-Bowers & E. Salas (Eds***.), Decision making under stress: Implications for training and simulation***. Washington, DC: American Psychological Association.

COGNITIVE TASK ANALYSIS AND MODELING OF DECISION MAKING IN COMPLEX ENVIRONMENTS

by

Wayne W. Zachary, Joan M. Ryder, and James H. Hicinbothom

CHI Systems, Incorporated

Lower Gwynedd, PA 19002

submitted to

Janis Cannon-Bowers and Eduardo Salas (Eds.)

Decision Making Under Stress: Implications for Training and Simulation

COGNITIVE TASK ANALYSIS AND MODELING OF DECISION MAKING IN COMPLEX ENVIRONMENTS[1]

Wayne W. Zachary, Joan M. Ryder, and James H. Hicinbothom

CHI Systems, Incorporated

Lower Gwynedd, PA 19002

Introduction

Decision theory has been characterized, for much of its history, by a debate on whether human decision processes are inherently flawed. The remarkable part of this debate is that for virtually its entire duration, it has been conducted without reference to detailed data on how people actually make decisions in everyday settings. In recent years, this issue has come to the forefront in the work of Cohen (1981), Barwise and Perry (1983), Klein et al. (1993), and others, who have pointed out the fundamental differences between decision making as it has been studied using traditional decision theory, and as it occurs in socially-situated naturalistic settings. The resulting *naturalistic decision theory* has emphasized highly detailed, almost ethnographic, studies of decision processes in specific domains. This had resulted in dense data but primarily prose representations and analyses.

In parallel to the rise of naturalistic decision theory, cognitive science and human-computer interaction (HCI) researchers were developing increasingly powerful analysis methods that collectively were called *cognitive task analysis* techniques. The purpose of these techniques was to analyze and model the cognitive processes that gave rise to human task performance in specific domains, as the basis for design and evaluation of computer-based systems and their user-interfaces. The TADMUS project provided a unique opportunity for these two avenues of inquiry to come together. This paper describes research that combined the highly formal methods and tools of the HCI community with the theoretical orientation of naturalistic decision theory. The aim of the research reported here was to create a detailed and domain specific model of decision making in the Anti-Air Warfare domain which was also sufficiently formal that it could be used to drive the design and development of systems to support and train decision making in that domain.

Cognitive task analysis and modeling were undertaken in the TADMUS program for three reasons (see also Cannon-Bowers and Salas, this volume). The first reason was theoretical. Theories of situated cognition and naturalistic decision making formed a major theoretical underpinning of the larger TADMUS program. These theories posit that the form and content of decision processes are highly determined by the organization of the domain in which the decision maker is operating. From this theoretical position, one cannot begin to develop decision making interventions, such as decision training tools or decision support systems, without a detailed understanding and formal representation of the relationship between the decision making expertise and knowledge that is unique to experts in that domain.

The second reason was historical. It is a virtual truism, based on decades of collective experience in human factors engineering, that the design of new or modified systems that include human operators should begin with a detailed mapping of what the human beings are (or should be) doing. This notion of "task analysis" is so strong as to be perhaps the single most unifying principle of human factors. More importantly, is the repeated observation that systems that have been built or redesigned with a sound task analysis at the onset are much more usable, lead to higher human performance, and require less training. Thus, there is every reason to suspect that a task analysis, albeit one that includes cognitive as well as observable acts, would be necessary in the TADMUS case as well.

The third reason is technological. The larger TADMUS program sought ultimately to develop actual systems, specifically decision support systems and team decision-training tools, that would improve empirical decision making. Many advanced technologies that could be incorporated into such systems required detailed models and analyses of the decision strategies of the human operators. For example, embedded user models could be used to create intelligent or adaptive user interfaces to the decision support system (e.g., Rouse et al., 1987) and the support system itself could incorporate or be designed from models of user strategies (see Kieras, 1997). The development of a detailed and accurate cognitive analysis would thus be an enabling condition for application of a broad range of potentially useful technologies for improving decision performance.

This paper discusses both the method for, and the results of, detailed cognitive analysis and modeling in complex, tactical domains of the kind considered by the TADMUS program. The first part of the paper is theoretical. It begins with a set of logical requirements for the cognitive analysis process, and then describes a framework used to meet those requirements. This framework, called COGNET, is discussed in terms of its theoretical underpinnings, its description, and data collection and analysis methods. The remainder of the paper presents the cognitive analysis of tactical decision processes that was conducted in the TADMUS research, in the form of the COGNET model that resulted from the analysis. The applications of the model are discussed in the conclusions.

Logical Requirements for Cognitive Analysis and Modeling

Cognitive analysis and modeling is a relatively new subject, and it means many things to many people. Many techniques for analyzing human cognitive processes and decision making have been developed (see Meyer and Kieras, 1996; Essens et al., 1995), and there is no clear 'standard' method that is appropriate for all situations and domains. Rather, the method used must be selected (or developed) to meet the specific needs of the analysis. The major needs for a cognitive analysis in TADMUS were cited above. From these, the requirements of the analytic method were defined. Specifically, the cognitive analysis had to be able to represent four major aspects of tactical decision making:

- *real-time* -- in tactical domains, data arrive and must be processed in real-time, so decisions have temporal constraints. Making the right decision too late is as bad (or worse!) than making the wrong decision in a timely manner. The cognitive analysis had to make clear how decisions were temporally organized and related to the flow of external events in the problem environment.

- *opportunistic and uncertain* -- while the tactical decision maker will have clear goals, the external events to be faced will typically be unpredictable. This means that it will be unclear exactly what decisions may be required until the situation unfolds. Moreover, the results of actions taken by the person are uncertain (i.e., they may or may not have the desired result). The decision maker thus must adapt both to the unfolding situation and to the results of actions taken. The cognitive analysis had to be able to capture this opportunistic aspect of the decision process.
- *multi-tasking* -- the pace of events and the uncertain nature of the process require the decision maker to be prepared to interrupt any cognitive activity to address a more critical decision at any time. This will typically result in a weakly-concurrent multi-tasking, in which the decision maker may have several decision processes underway at a time (with one processing and the others suspended). Managing competing demands for (limited) attention is a critical part of the decision process requiring the cognitive analysis to address the ways in which decision makers managed and shared attention.
- *situated in computer-based and verbal interactions* -- the majority of information available to the tactical decision maker comes not from direct sensation of the problem environment, but rather through information displayed at computer-based workstations and verbal messages from teammates. Similarly, decisions are implemented not through direct action, but as interactions with the computer workstation or verbal messages to other persons. The cognitive analysis had to be able to capture decision processes that were based on these types of input/output stimuli.

In addition to these requirements and constraints imposed by the tactical decision process itself, the cognitive analysis also needed two other properties:

- *integrate behavior and cognitive processes* -- the cognitive analysis must feed both into training and decision support interventions, both of which function by relating decision maker *actions* to decision maker cognitive states. Thus the analysis to be undertaken had to decompose and describe not just cognitive processes, but also the way in which those cognitive processes were linked to observable behavior.

- *generic (predictive) rather than situation-specific* -- many forms of conventional and cognitive task analysis attempt only to describe and decompose human behavior in the context of a specific exemplar situation, called a scenario. However, the analysis required here clearly needed to be at a more general level. Because it needed to be used in training and decision support interventions that could apply to *any* scenario, the analysis and resulting cognitive model also had to be able to predict the decision processes and their associated observable actions in any scenario.

Finally, of course, the analysis had to be undertaken with a technique that was able to deal with the complexity of a difficult naturalistic setting such as Naval command and control.

The COGNET Framework

To address the above requirements, the cognitive analysis was undertaken with an adaptation of a cognitive analysis and modeling method developed by the authors and colleagues in prior research. This framework, called COGNET (for COGnition as a NEtwork of Tasks), is a theoretically based set of tools and techniques for performing cognitive task analyses and building models of human-computer interaction in real-time, multi-tasking environments (Zachary, Ryder, Ross, and Weiland, 1992). COGNET had been developed, applied, and refined in a series of earlier studies. The original development and application was to a vehicle tracking domain (Zubritzky, Zachary and Ryder, 1989; Zachary et al., 1992). COGNET's ability to represent and predict attention-switching performance was empirically demonstrated in a validation study based on the vehicle tracking model (Ryder and Zachary, 1991). The COGNET vehicle-tracking model was also successfully translated into an embedded user model for an intelligent, adaptive human-computer interface (Zachary and Ross, 1991). The framework was subsequently applied to several other complex domains, including en-route air traffic control (Seamster, Redding, Cannon, Ryder, and Purcell, 1993), and telephone operator services (Ryder, Weiland, Szczepkowski, and Zachary, in press). The COGNET analysis of air traffic control was used as the basis for redesigning a training curriculum (Australian Civil Aviation Authority, 1994), while the analysis of operator services was used to design new interfaces and decision support tools.

The COGNET framework is summarized below, in terms of its theoretical base, its description language (in which the knowledge is actually represented), and its data collection, knowledge elicitation, analysis, and knowledge representation methods.

Theoretical Underpinnings

The theoretical underpinnings of COGNET research lie in cognitive science research, particularly the symbolic computation branch which views cognitive processes as the operation of a specific computational mechanism on a set of symbols, which are themselves a representation of sensation, experience, and its abstraction (see, for example, Pylyshyn, 1984, and Newell (1980) for theoretical discussions of this viewpoint). Thus COGNET presumes:

- an underlying mechanism of a specific structure with clear principles of operation (henceforth termed the cognitive architecture) , and
- a set of underlying symbols on which it operates (henceforth termed the internal knowledge), and which are organized in a specific representational scheme (henceforth termed the knowledge representation).

Both are largely developed from the work of Newell (see Newell and Simon, 1972; Card, Moran and Newell, 1983; Newell, 1990), which in its simplest form breaks human information processing into three parallel macro-level mechanisms -- perception, cognition, and motor activity -- shown as the ovals in Figure 1. Perception (which, in COGNET, includes the physical process of sensation) receives information from the external world and internalizes it into the symbolic or semantic information store that is accessed by both the perceptual and cognitive mechanisms through an information store that is shared by both. As used in COGNET, this symbol store corresponds to what has come to be called extended working memory (see Ericsson & Kintsch, 1995). This shared store is depicted in Figure 1 because it is shared by both mechanisms, but it is not the only information store in the COGNET architecture. Both the cognitive and sensory/perceptual mechanisms incorporate other information stores that are accessed by each mechanism (i.e., long-term memory accessed by the cognitive mechanism, and acoustic/visual information stores accessed by the perceptual mechanism).

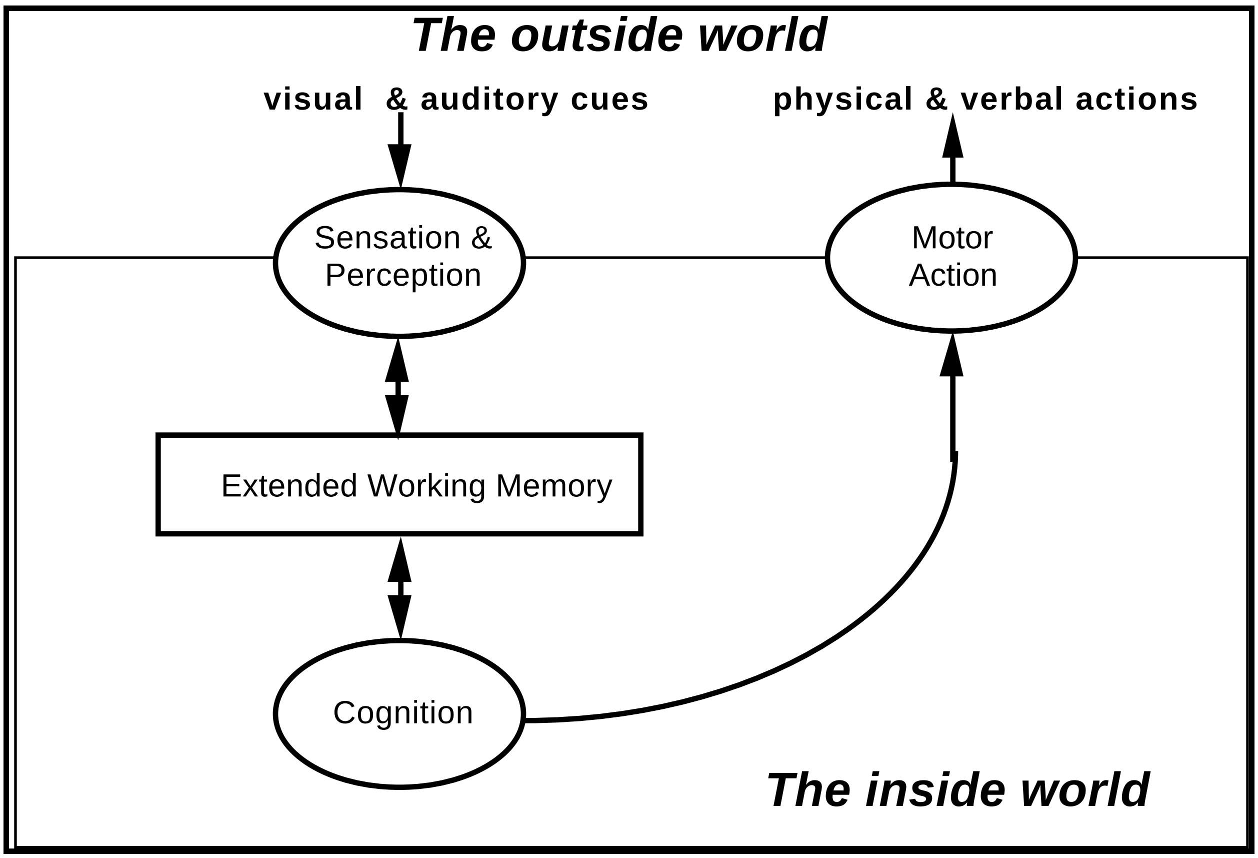

Figure 1. Conceptual View of COGNET Cognitive Architecture

A completely parallel cognitive process manipulates this internal symbolic representation of the external world, using previously acquired procedural knowledge. The cognitive process thus operates on an internal 'mental model' of the world, not on direct sensations of the world. The cognitive process also modifies the mental model, as a result of cognitive reasoning processes (induction, deduction, abduction). The problem representation thus is affected both by the perceptual processes and the cognitive processes. The cognitive process, in addition to being able to modify the problem representation, can also invoke actions through commands or instructions to the motor system. This system operates outside the scope of the problem representation (i.e., does not have access to and does not depend on the contents of the current extended working memory). The motor system provides manipulation of physical instrumentalities that in turn manipulate the environment (Card, Moran and Newell, 1983, provide a detailed empirically-based argument for this underlying structure).

Each person is presumed to possess and use the same mechanisms described above (subject to individual variations in parameters of the mechanism). Thus, mechanism by itself does not help differentiate the novice from the expert or the ability to make decisions in one domain from another. Given this observation, it must be the other component of the system -- the internal knowledge -- that differentiates performance among domains and expertise levels. Thus, the goal of cognitive task analysis must be to understand and represent the internal knowledge of experts in the domain of interest. The COGNET framework has been developed to provide a practical tool which can be used to pursue this goal in complex, real-world domains.

COGNET presumes a certain organization and representation of internal knowledge, based on the architecture in Figure 1, and the emerging cognitive theory of expertise, as discussed generally in Chi, Glaser, and Farr (1988); Ericsson and Smith (1991); Hoffman (1992); Ryder, Zachary, Zaklad, and Purcell (1994); VanLehn (1996); or Zachary and Ryder (1997). In real-time, multi-tasking, HCI-based decision domains, the person interacts with the external problem environment through the medium of the machine system (and specifically, through the person-machine interface). The person is implicitly assumed to be in a work-setting, and therefore to be pursuing some high level mission or goal with regard to the external environment. Within this overall goal, the activities of the expert human operator of the person-machine system appear as a set of tasks with complex inter-relationships. These tasks represent chunks of knowledge that the expert has compiled from lower-level procedures and rules for use in a broad range of situations or cases. They are analogous to the various 'case strategies' that are the basis for the case-based reasoning theory of highly expert decision-making and planning (e.g., Kolodner, 1988). Some of these tasks may compete to be performed simultaneously while others may be complementary; still others may need to occur essentially in sequence. Each task represents a specific ‘local’ goal which the operator may pursue to achieve or maintain some aspect of the overall mission/goal.

The way in which a task is accomplished depends heavily on the evolution of the current problem-instance to that point and the current problem situation at each instantiation of the task. The knowledge in the task contains the expert's understanding of how the task's goal can be achieved in different contexts. The knowledge that makes up the task also includes the knowledge needed to recognize the situations or contexts in which the task goal is relevant. In this regard, the COGNET tasks are activated by a recognitional process, analogous to that used in Klein's recognition primed decision-making (Klein, 1989). In a real-time domain, however, multiple cognitive tasks may be recognized as needing to be done without there being enough time to actually carry out each task. Thus, the tasks must compete with each other for attention, with each task that has been recognized as relevant 'shrieking' for the focus of the person's attention. This shrieking process is analogous to the Pandemonium process for attention originally postulated by Selfridge (1959). Even when a task gains the focus of the attention and begins to be executed by the cognitive processor, the process of tasks competing for attention continues unabated. The result is that tasks may (and often do) interrupt one another, and a given task may be interrupted and resumed several times as the operator copes with the evolving sequence of events.

What unites these separate chunks of procedural knowledge or tasks into a more global problem solving strategy is a common declarative representation of the overall situation and its evolution (Hayes-Roth; 1979, 1985). This common problem representation is highly interactive with the individual tasks. As a given task is performed, the person gains knowledge about the situation and incorporates it into the current problem representation; similarly, as the problem representation evolves, it can change the relative priority among tasks and lead one task to ‘come to the front’ and require immediate attention. At the same time, much of the information in the current problem representation is obtained from perceptual processes, for example, by scanning and noting information from displays, external scenes, or auditory cues, encoding it symbolically, and adding it onto the declarative problem knowledge. The procedural knowledge in each task includes knowledge about when and how to activate specific actions at the workstation or in the environment. These action activations are passed to the motor system where they are translated into specific motor activity (e.g., button presses).

This conceptual view of the types and organization of knowledge is pictured in Figure 2. It gives COGNET the structures necessary to link sensation/perception, reasoning and decision-making, and action into a common framework. The conceptual structure shown in Figure 2 was created to deal with individual-level decision making. In Anti-Air Warfare or AAW, the specific domain of interest in TADMUS, decision making is highly distributed throughout a team. Thus, this pre-existing COGNET framework had be enhanced to address the team nature of command and control decision-making. The concept of perceptual knowledge was broadened to include verbal communication among team members. In addition, the concept of declarative knowledge was broadened to include a representation of other team members and their roles within the team.

COGNET Description Language for Cognitive Task Analysis

One main reason for developing a theoretical framework for human information processing and decision making is that the framework can provide a means of decomposing empirical phenomena in a way that permits their more formal description. The process of constructing a formal description for a specific set of human activities in a specific domain constitutes a form of that mainstay of human factors, the task analysis. In particular, it is a cognitive form of task analysis, because it relates cognitive constructs and mechanisms to the observed behaviors. In order to conduct a cognitive task analysis (with COGNET or any other framework), it is necessary to have a set of constructs that are to be identified and described, and notation with which to describe them. The knowledge framework pictured above in Figure 2 identifies the set of constructs that are necessary for real-time, multi-tasking performance. To construct a cognitive task analysis using this framework, specific notations are used to describe each of the four major types of knowledge included in Figure 2 -- perceptual knowledge, declarative knowledge, procedural knowledge, and action knowledge. These notations, summarized below, have been derived from existing notations within cognitive science and knowledge engineering wherever possible. The attention process, which is really an epiphenomenon of the knowledge description process, is also summarized.

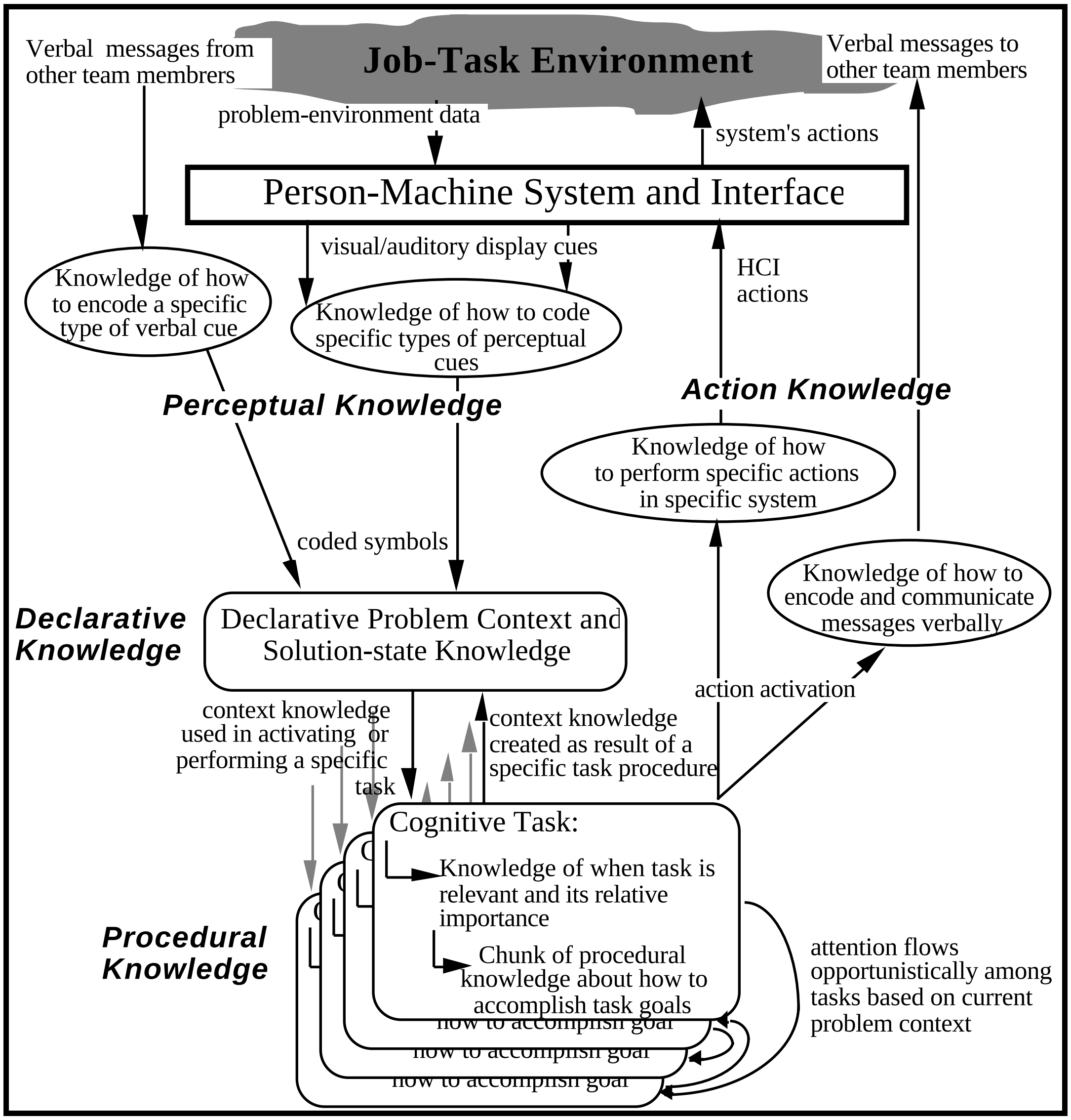

Figure 2. COGNET Knowledge Framework

Declarative knowledge, in COGNET, refers specifically to the person's internal representation of the current problem, including its history or evolution to the current point, all declarative information related to the solution strategy for that current problem, and all long-term knowledge about things in the environment (e.g., system characteristics). Solution strategy information includes, for example, the features of a plan that is being developed, and/or expectations about future events. Representing this kind of knowledge has been widely studied under the rubric of blackboard systems (e.g., Nii, 1986a,b; Englemore and Morgan, 1988; Carver and Lesser, 1994), and the COGNET framework uses blackboard notation to describe declarative knowledge. A declarative knowledge blackboard is a collection of individual hierarchical structures, each of which is called a panel. Each panel consists of a hierarchy of knowledge elements, called levels, that are conceptually related. Often, but not always, the levels represent different degrees of abstraction in the conceptual space defined by the panel, although they may also represent a simple partitioning of the conceptual space into different aspects. Each level represents a dynamic collection of different individual concepts that provide specific instances of information at that level of the panel's overall conceptual space. For example, a panel may represent a tactical situation, and different levels may partition that conceptual space into air tracks, surface tracks, and subsurface tracks. The air track level, in this example, would then contain individual concepts that correspond to the individual air tracks of which the person is aware in the current situation. Each individual concept will have a number of attributes that are common to concepts to the panel/level where it is located (also termed where it is 'posted'). The attributes of air tracks, for example, might be different from the attributes of surface tracks (e.g., the latter not having an altitude attribute). The attribute values differentiate the individual concepts from one another at a given panel/level. Additionally, concepts on the blackboard can be semantically associated with other concepts. For example, an air track can have the relationship 'took off from' a surface track. However, every concept at a given level may not have a relationship of each kind (e.g., some air tracks may not have taken off from any surface track, but may have taken off from a land location, or from another air track).

Two additional pieces of terminology are applied to the declarative knowledge blackboard in COGNET. First, all the messages on the blackboard at any one time constitute the *context* for cognitive processes at that time. This momentary context drives the way in which procedural knowledge is activated and applied. Second, the structure in which the declarative knowledge in the blackboard is organized is sometimes termed the person's *mental model* of the domain. This usage is primarily metaphorical, helping explain the construct to domain experts from whom knowledge must be acquired, and who must help validate the model. It is not intended as a strong theory of mental models in the cognitive science sense.

In COGNET, the information processing activity is presumed to occur through the activation and execution of chunks of *procedural knowledge*, each of which represents an integration or compilation of multiple lower-level information processing operations around a domain-specific high-level goal. This combination of the high level goal and the procedural knowledge needed to fulfill it are referred to as a *cognitive task*. All the knowledge compiled into each task is activated whenever the high-level goal defining that task is activated. Each task-level goal includes metacognitive knowledge that defines the contexts in which that task is relevant. This metacognitve knowledge is simply a description of the contexts (as defined above) under which the goal should be activated. Thus, the high-level goals are activated according to the current problem context, as defined by this metacognitive 'Trigger.' In addition to this Trigger, another piece of metacognitive knowledge defines the relative priority of that goal in the current context, so that attention can be allocated to the goal with the highest priority given the current context. This second piece of metacognitive knowledge is called the Priority expression. These common features provide the structure for describing a cognitive task in COGNET. Each task has two main parts, the Task definition, and the Task body. The Task definition identifies the high-level goal involved, and a specification of when that goal is activated and the priority it will have when activated. A cognitive task is defined in the following form:

TASK <task-goal-name> … *activation condition /Priority (formula)*

<task body>

The task body is a hierarchy of lower-level information processing operators, based strongly on the GOMS (Goals-Operators-Methods-Selection Rules) notation of Card, Moran and Newell (1983), but with customizations to allow for:

- manipulation of concepts on the blackboard;
- evaluating GOAL conditions on the basis of the blackboard context; and

- interrupting and suspending the current task.

As in GOMS, the GOALs can be either sequential or subordinated to one-another, forming a hierarchical structure that defines branch points in the procedural logic. The lowest level GOALs have only non-goal operators as 'children'; which, when performed, accomplish that goal. Other COGNET Operators fall into three groups: action, cognitive, and metacognitive.

- Action operators involve interactions with the workstation, both generic (e.g. Point, Enter, etc.) and workstation-specific (Perform <FUNCTION>). Verbal actions are denoted with the "Communicate" operator. Action operators comprise *action knowledge*, in that they define how actions are to be performed in a specific job-task environment.
- Cognitive operators create (i.e., POST), delete (i.e., UNPOST), or manipulate (i.e., TRANSFORM) declarative knowledge on the blackboard, or encapsulate lower-level cognitive processes that are included in a task only by reference (e.g., Determine ["Is_track_ flying_commercial_air_ route?"]).
- Metacognitive operators actually affect the execution of the procedure via conditional suspensions (i.e., Suspend).

As in GOMS, frequently-used goal-operator subhierarchies can be subsumed into separate units called methods, that can be invoked directly by name. It should be noted that not all of the notation is necessarily needed to describe any specific domain. Domains with a great deal of explicit human-computer interaction may require more use of the action operators, while those in which the cognitive processes are not closely coupled with machine controls may make little or no use of these action-level operators.

The final element of COGNET is a notation for describing the person's *perceptual knowledge*. In Figure 2, perceptual processes are assumed to operate in parallel with the cognitive processes, and the information registered by perceptual processes are passed to the cognitive subsystem by being entered directly onto the problem representation blackboard. Thus, information enters the purview of the cognitive processes via spontaneous events or activities of the perceptual systems which POST objects onto the blackboard. These events are modeled in COGNET simply as production rules, of the form:

IF <environmental event> Then POST <panel:level:attributes>

Each of these rules is termed a *perceptual demon*, or sometimes a *perceptual monitor*. A demon is spontaneously activated and executed whenever the corresponding sensory event, typically a verbal or visual cue, is sensed.

In COGNET, the person's *attention* resides, at any given moment, in some specific cognitive task which is being performed. Attention can shift in one of two ways. First, attention can remain at one task until it is *captured* by another task. This results when a change in the problem representation causes some second cognitive task to be activated (i.e., causes its Trigger condition to be satisfied), and results in it having higher priority (as defined by its Priority expression) than the currently-executing task. When this occurs, the second task will capture the focus of attention from the first task, which will remain suspended until it, once again, has the highest priority or until its activation TRIGGER condition is no longer satisfied.[2] Second, attention can be suspended. A given procedure within a task can involve events or actions which involve expected delays (e.g., giving a navigational direction and then waiting for it to be carried out). Attention in the current task can then be deliberately suspended. This then allows whatever other task has highest current priority to gain the focus of attention. When the expected event occurs, the suspended tasks must then again compete to 'recapture' attention from whatever task then has the focus of attention.

Data Collection and Analysis Methodology

The main purpose of the COGNET framework is to facilitate the cognitive task analysis and description of specific work domains[3]. The COGNET analysis notation described above is supported by a methodology for collecting, analyzing, and reducing empirical data on behavioral and cognitive processes so that they can be represented in the COGNET notation. The general approach is naturalistic, in that the behavior of an expert decision-maker in a realistic problem solving context is the data used for the analysis. In more concrete terms, experts in the domain are typically asked to make a series of decisions, normally representative problems embedded in the form of scenarios, in the domain of interest. While this can be accomplished in either the actual environment or a simulated equivalent, the latter is usually chosen because it affords more experimental control. The selection of both the scenarios and the operational personnel must reflect the range of problem solving challenges posed by the actual operational environment and the range of strategies to meet those challenges. This scope insures that the diversity and complexity of the environment will be captured by the COGNET model.

For each real or simulated scenario, the activities of each subject expert are observed and recorded for subsequent analysis in conjunction with verbal introspective data collected using knowledge elicitation methods. The verbal data, in the form of thinking-aloud protocols and question-answering protocols (obtained while reviewing the recorded behavior), are taken immediately after the problem or simulation has been completed. This is done because in high workload, time-constrained domains of the kinds studied here, taking the verbal protocol *during* the problem-solving process is too intrusive to be practical. Experience has shown, however, that high quality protocols can be obtained in response to recordings of actual behavior, particularly if they are taken shortly afterwards when the problem is still fresh in the subject's mind. In these verbal protocols, subjects are asked to introspect and recount their internal decision process. Specific verbal probes are often made to clarify these accounts (at which time the replay of the problem is temporarily halted). These primary verbal data are supported by unstructured debriefs by participants, and interviews and critiques by subject matter experts (SMEs) from the domain (especially instructors), particularly during the data-analysis process.

This method places COGNET somewhere along a continuum between purely experimental methods and purely naturalistic ones. Experimental paradigms often use data gathering techniques that employ subjects unfamiliar with the problem domain and artificially-constructed experimental problems to generate experimental data. This affords great experimental control as well as convenient statistical analysis, but excludes most or all features of domain-specific context. Naturalistic paradigms, on the other hand, typically rely on observation of behavior in its usual (or ‘natural’) context, and/or verbal accounts of behavior in that context. This affords a great ability to understand the role of context in decision making and other cognitive processes, but usually excludes very detailed data needed for quantitative, computational, or statistical analysis.

Once the problem-solving and verbal data have been collected, the analysis of this data proceeds. Identification of available motor Actions (e.g., buttons to press, messages to speak, displays to observe) and required Perceptual demons (e.g., information content of visual displays, messages overheard, and other perceived events) proceeds as early as possible. These model elements are usually identified from system documentation, job descriptions, and information environment audits drawn from the recorded data and interviews. The initial stages of the analysis decomposes the decision processes in the problem domain into a set of cognitive tasks that organize the decision maker's procedural knowledge and an initial problem representation structure (i.e., blackboard panels and levels). This is done by reviewing the sequences of problem solving behavior (either through video/audio recordings or use of computer-generated problem 'replays') in conjunction with verbal protocols from the subject experts. For the task decomposition, the principal focus is on recreating expert-level context-sensitive attention shifts among competing cognitive tasks. With regard to the blackboard structure, the focus is on identification of the primary principles for structuring domain knowledge to support decision making.

The detailed modeling of each task is undertaken, in general, as a traditional GOMS analysis. Unlike GOMS analysis, however, a COGNET analysis involves extensive definition of specific aspects of the task representations based on the declarative knowledge incorporated into the blackboard structure. Thus, the blackboard structure is defined iteratively as its elements (i.e., concepts, their attributes, and semantic links relating them to each other) are needed to specify the Trigger condition, Priority expression, and GOAL conditions in the various cognitive Tasks. As this is done, the necessary cognitive operators to POST, UNPOST, and TRANSFORM the concepts on the blackboard must be inserted into the Tasks and Perceptual demons. The final phase of the analysis involves reviewing, validating, and revising the model to increase its completeness and quality. A major means of doing this is through a series of walk-throughs using domain experts and (ideally) a set of problem scenarios not used in the initial data collection. The model is walked through against the scenario step-by-step, assessing with the domain expert whether the actions predicted by the model are valid and complete. The model is then revised as specific problems are identified, with the walk-through process continuing until an acceptable level of model validity is reached.

The remainder of this paper describes the COGNET analysis undertaken for the Anti-Air Warfare domain that was the focus of the TADMUS program.[4]

Applying COGNET to the AAW Domain

The abstract methodology described above needed to be operationalized before it was possible to conduct a COGNET analysis of the Anti-Air Warfare domain studied in the TADMUS program. In particular, it was necessary to:

- identify a specific decision maker or set of decision makers to analyze,
- define a setting in which baseline decision performance could be observed and recorded,
- collect and analyze an appropriate body of data on AAW decision performance using the COGNET method and description language.

This customization of the general methodology to the AAW problem is discussed below.

Selecting an AAW Role to Analyze and Model

Anti-Air Warfare decisions are made at multiple organizations by multiple individuals in a Naval Battle Group. Even with the TADMUS focus on ship self-defense, many individual roles are involved in the AAW process. The Commanding Officer (CO) and Tactical Action Officer (TAO) have overall tactical responsibility for all ship activities, and so make AAW decisions. Within the Combat Information Center several teams focus on different tactical areas, one of which is AAW. The AAW team consists of many roles that deal with specific parts of the AAW problem, such as identifying tracks, observing and interpreting electronic emissions from tracks, controlling

Combat Air Patrol (CAP), which are fighter aircraft assigned to an AAW role, and so on. The activities of all the individual roles are integrated and organized by the Anti-Air Warfare Coordinator or AAWC.

The COGNET analysis was focused on the AAWC, for several reasons. The main function of the individual standing watch as AAWC is the coordination of own-ship AAW resources to defend own ship (and to defend the battle group as well, if so directed). The AAWC works under the command of the own-ship CO/TAO, and under control of a battle group AAW commander in battle group AAW operations. The AAWC is the highest level individual on the AEGIS ship who is concerned with just AAW. The AAWC performs this job mainly through interactions with the computer-based workstation, hooking (i.e., selecting) and interrogating (i.e., requesting data on hooked) track symbols on the display, moving attention among different part of the tactical scene, trying (on the one hand) to actively push the process through the detection-to-engage cycle for individual tracks, and (on the other) to coordinate the use of resources for AAW, usually via voice traffic. At any given time, the AAWC must choose which of many tracks/tasks to attend to, internalize implications of decisions/actions of others on the AAW team, and also infer the intents and implications of the enemy and the overall tactical situation as shown on screen.

Defining a Data Collection Setting

A major problem for naturalistic methodologies such as COGNET is the difficulty in gathering data on realistic, *in-situ* behavior. Purely introspective approaches such as the critical decision method used by Klein and colleagues (e.g., Klein, Calderwood, and MacGregor, 1989) avoid this problem by relying only on recollections of past decision situations. COGNET analysis, however, relies on the long-standing protocol analysis approach of integrating verbal introspective accounts with physical task performance. Thus the analysis needs to work with, as primary data, expert level individuals solving realistic problems in the setting in which those problems are usually encountered. In domains such as Anti-Air Warfare, this can create some problems in data acquisition.

There were several ways in which data on experienced AAW teams working realistic AAW problems could be gathered. One was on-board Navy ships. The logistics involved with this and the intrusiveness of the data-recording process to shipboard life ruled this option out almost immediately. Other options all involved simulated problem solving. One simulation possibility was the DEFTT simulation system that was developed for the TADMUS program (see Johnston, Poirier, & Smith-Jentsch, this volume). However, DEFTT had drawbacks, most notably the fact that it was still under development at the time the data for this analysis had to be collected. Even if it had been available, though, two problems still remained. First, it was believed that DEFTT posed too low a fidelity in mission- and (particularly) workstation-simulation. A model build with DEFTT might be complete, but would be difficult to apply back to the more complex ship-board systems for which the analysis was ultimately directed. And second, DEFTT contained no communication networks (internal to the CIC or external to the ship), thus removing one major source of perceptual cues for the operators.

Another simulation source was the embedded training simulations used for on-board training of ship crews. This option, though, raised the same logistical problems as the original on-board data collection option. The one remaining option -- shore-based team-training simulators -- fortunately provided a workable solution. These simulators used high-fidelity mission simulations and the same workstations as used on shipboard, contained both internal and external communication networks, ran realistic scenarios of varying complexity, and had powerful built-in facilities for recording data on operator performance. There were still difficulties, of course. Team training occurred at irregular and often widely-spaced intervals, and access to the facility required complex scheduling arrangements. When a team was in the facility, their main purpose was training, and other needs (such as COGNET data collection!) had to be 'piggy-backed' onto the schedule of the team and the data collection plans of the training-facility staff. In general, though, the team-training simulator provided ample opportunities to observe and record data on experienced teams solving realistic AAW problems in their 'natural' setting.

Data Collection and Analysis

The COGNET analysis of the AAWC role was based upon the following specific data sources and types:

- observations and recordings of AAWC actions during parts of four, three-week, team-training classes at a high-fidelity team-training simulator;[5]
- debriefings by participants in observed team training sessions;
- verbal protocols collected in the context of replays of the recordings of observed AAWC behavior; and
- reviews and walkthroughs of the evolving analysis by a variety of experts in AAW, Navy tactical experts, members of several ships' crews, and civilian AAW trainers.

The observed behavior was recorded on videotape, which captured:

- the AAWC video displays' contents, including the spatial display of tracks from the radar display and most of the textual data from the Character Read-Out display;
- voice communications, including the ship's internal command network communications and the cross-ship AAW network[6] plus much of the AAWC's direct voice communications with the other AAW team members seated nearby; and
- physical interaction with console and environment, including a wide-angle view of AAWC seated at the console.

The analysis was conducted using the methodology described above, with one modification. The verbal introspective data were collected not from the original subjects who performed the tasks, but rather from a second group of experts. This was done primarily because the original subjects were generally unavailable for the verbal data collection process. The observed behavior that was recorded was collected as part of operational team training at a working Navy training facility. The team members whose actions were observed simply had to return to other duties shortly after the training simulations were completed and did not have time to participate in reviews and verbal protocol procedures. Although this posed a problem, the extreme difficulty of collecting recordings of observed behavior with experienced teams using realistic equipment and scenarios in any other setting outweighed it. In lieu of the original subjects, a set of mastery-level experts were obtained from another source[7] , and used to collect the verbal data. As experienced trainers, these individuals were cognitively familiar with the task of observing, inferring, and commenting on the decision processes of individuals standing watch in the AAWC (and other AAW) role. The only modification of the methodology discussed above was that each expert saw each recording twice, the first time without comment to familiarize himself with the problem and what the subject did, and the second time immediately afterward to provide the verbal protocol.

The verbal data were then analyzed to identify cognitive tasks and the procedural bodies of these tasks, a blackboard structure, and a set of perceptual demons which linked the cognitive processes to the external audio and visual cues. The results of this analysis are given in the next section, in the form of the resulting COGNET model.

Analysis Results: The AAWC COGNET Model

The result of a COGNET cognitive task analysis, as discussed earlier, is a COGNET model that represents the underlying knowledge that an expert maintains and uses to make and implement decisions in a specific domain. Following the structure of the conceptual framework and description language, a COGNET model has three key parts:

- the blackboard;
- the set of cognitive tasks, each of which includes a definition of its associated triggers, priority expression, and procedural bodies; and
- the set of perceptual demons which filter relevant information from the environments and post it on the blackboard.

Descriptions of each of these three components are given below. Following that is an example of a realistic AAW situation describing how that situation would be modeled dynamically.

Blackboard

Figure 3 shows the panel/level organization of the blackboard portion of the AAWC COGNET model. The structure of this blackboard implies a conceptual framework used by the AAWC for organizing domain knowledge and implies a strategy for applying the knowledge in job conduct. The contents of the blackboard at any point in time represent the AAWC's dynamic understanding of the tactical situation and his plan for handling it. This blackboard is composed of six panels where different categories of concepts get posted. The blackboard organization corresponds to the two primary aspects of the AAWC's job of monitoring and assessing the tactical situation and evaluating and responding to threats. Threat response management includes evaluation of air targets for threat status, determining those to engage, monitoring the progress of engagements, as well as plans for engaging threatening targets and determining that they have been destroyed. The two panels dealing with threat response management are:

- *Threat Status* -- information on tracks that are potentially or actually threatening, must be monitored or acted upon or used in an engagement.
- *Plans* -- strategies and plans for responding to anticipated or actual threats.

Situation assessment includes maintaining an understanding of the tactical picture; its geo-political context; the status of resources needed to obtain information or conduct an engagement; and ownership and force command structure, coordination agreements and communication links. Four panels represent the AAWC's understanding of the tactical situation, as follows:

- *Geo-Political Picture* -- information on the geo-political context in which the battle/watch takes place.
- *Tactical Picture* -- snapshot of the current status of an evolving battle/watch.
- *Resource Status* -- status of all resources needed or under control.
- *Team Relationships* -- AAWC's relationships with other players with whom he must coordinate, including methods of communication and coordination.

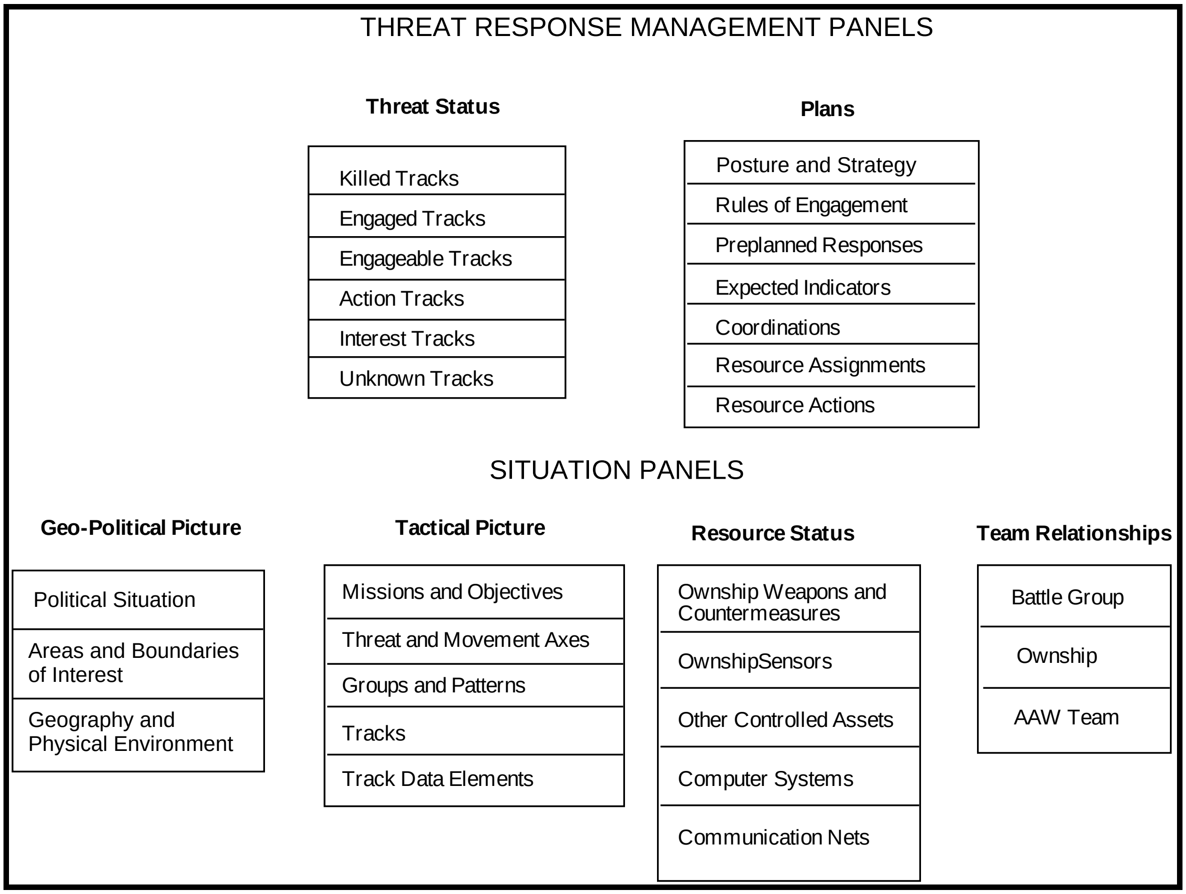

Figure 3. AAWC COGNET Model Blackboard

Each of the six panels are composed of a number of levels on which specific concepts are posted. The individual concepts posted in each level are structured by specific attributes. The detailed definition of one blackboard panel is given below, to provide a flavor of the blackboard contents. Each description takes the form of a specification, for each level on that panel, of the structure of concepts posted at that level, as:

[main attribute, modifying attribute1, modifying attribute2, ... modifying attribute n]

optional parameters are given as <parameter name>

The Threat Status Panel contains information on tracks that are potentially or actually threatening, and must be monitored or acted upon because of their relationship to a threat (e.g., a friendly track whose position must be considered in threat planning). A separate panel, the Tactical Picture panel contains all tracks in the area of operation. The Threat Status panel contains only a subset of the tracks contained on the Tactical Picture panel. Thus, it is a mental construct for reducing the amount of information that must be attended to by tagging those tracks that are of interest, in that they must be monitored or some action must be taken regarding them. The ordering of the levels within a given panel often can provide further organization of the knowledge situated on that panel. For example, the ordering of the levels often represents a *constructive* process whereby a solution to some significant decision problem or aspect of the overall situation is constructed by building the solution knowledge. In this panel, the level hierarchy represents a progression, from bottom to top, in understanding what, if any, threat the track poses, and progress in eliminating the track as a threat.

The six levels of the Threat Status panel are shown in Figure 4. New or unevaluated tracks are posted on the bottom level by a perceptual demon. The tasks of “Manage Battlespace” and “Evaluate Track” include cognitive operations to evaluate tracks that are or could be threats. Those that the AAWC determines must be monitored are promoted to the Interest Track level by UNPOSTing them on the Unknown tracks level and POSTing them on the Interest Track level with the appropriate attributes added. If at any time a response is required to a track behavior or threatening intention, the track is promoted to the Action Track level. If unknown tracks are identified as commercial air or if a potentially hostile track turns away, these tracks are deleted from the Threat Status panel, although they would remain as part of the Tactical Picture panel. When a track meets the Rules of Engagement (ROE) and is classified as hostile, it is promoted to the Engageable Tracks level, which triggers the task “Take Track.” Once the engagement has begun, the track is promoted to the Engaged Tracks level, and if it is killed, it is promoted to the Killed Tracks level. Messages for killed tracks would only remain on the blackboard until the evaluation is complete and the track is deleted from the system.

Killed Tracks -- tracks that have been destroyed
[track #, location, weapon]
Engaged Tracks -- tracks that have been engaged
[track #, location, characteristic(s) of interest, weapon engaged with]
Engageable Tracks -- tracks that have met the Rules of Engagement (ROE)
[track #, location, characteristic(s) of interest, weapons capable of engaging]
Action Tracks -- tracks requiring some action to be taken (e.g., warnings, reports)
[track #, location, category (air, surface, sub-surface), classification (hostile, neutral, friendly, etc.), characteristic(s) of interest, action needed]
Interest Tracks -- tracks that must be monitored because they are potential threats or are friendly tracks that the AAWC must be aware of for coordination in an engagement
[track #, location, category, classification, characteristic(s) of interest]
Unknown Tracks -- unevaluated tracks
[track #, category, location, status (new, unknown)]

Figure 4. Threat Status Panel

Cognitive Tasks

The overall role of the AAWC is to monitor and evaluate air targets for threat value and engage and destroy all air threats (under direction of CO/TAO and battle group anti-air warfare commander). In a COGNET analysis, tasks are the primary units of cognitive activity, and are defined as a single unit of goal-directed activity that would be performed to completion if uninterrupted. Thus, each cognitive task encapsulates a logically self-contained procedure, which is formalized as a set of subgoals that are sequentially pursued to attain the overall task goal. Ten tasks resulted from the COGNET analysis of AAWC. They are as follows:

1. MANAGE BATTLE SPACE -- Scan tracks in larger context of evolving scenario.
2. EVALUATE TRACK -- Identify/classify an individual track in terms of its tactical significance.
3. PLAN SPECIFIC THREAT RESPONSE -- Plan a response to a specific track that is a potential or actual threat. The plan may include specific actions to take if the track becomes hostile at any of various points along its projected path.
4. PLAN POSTURE FOR EXPECTED THREAT -- Plan posture and strategies for handling expected classes of threats. Includes determining assets needed for expected

threats, establishing preplanned responses in accordance with battle orders, and understanding of the geo-political situation.

5. NEUTRALIZE/CONTROL POTENTIAL THREAT -- Get a potentially hostile track to conform to your needs/wishes.
6. COVER TRACK -- Mapping a specific tactical response/targeting solution to a track of interest.
7. TAKE (ENGAGE) TRACK -- Engaging a hostile track.
8. POSITION AAW ASSETS TO MAINTAIN DESIRED POSTURE -- Positioning assets in accordance with plan.
9. MANAGE CAP STATUS -- Monitoring and maintaining CAP (Combat Air Patrol) in readiness for expected or actual threats. Includes maintaining adequate fuel and weapon load.
10. MANAGE RESOURCES -- Monitoring and maintaining AAW resources, including sensor, weapon/countermeasure, computer, and communication systems.

A task description includes the task name, which defines the goal associated with the task, and the trigger which defines the conditions under which the task is activated for performance. The body of the task is described in the COGNET description language reviewed above. The format for representing references to the blackboard in the triggers, goal/subgoal conditions, and cognitive operators (POST/UNPOST/TRANSFORM) is:

[<object> posted on PANEL: *Level*]

The full blackboard panel names (from Figure 4) are abbreviated in these PANEL:*level* references as follows:

- "Threat Status" abbreviated as THREAT
- "Tactical Picture" abbreviated as TAC PIC

The analysis only decomposes the tasks into subgoals, and in some cases, sub-subgoals, thus giving the primary skeletal structure of the model.[8] Where appropriate, cognitive and communicative operators are included to illustrate how changes are made to the contents of the blackboard and how task subrogation occurs. Part of one of the ten individual task models, Evaluate Track, is as follows:

TASK: Evaluate Track...*IF <new track> posted on THREAT: Interest Track OR query/command regarding specific track or locational fix OR (<new track data> posted on TAC PIC: Track Data Elements AND track on THREAT: any level*) OR lost track

GOAL: Locate track

GOAL: Review track data

Operator: Determine how much time available to evaluate track (distance from ownership)

GOAL: Assess intentions and threat capability of track... *if time available*

Operator: TRANSFORM <track data> on TAC PIC: *Track Data Elements* to <tracks> on TAC PIC: *Tracks*

GOAL: Determine if track is part of group or pattern... *if time available*

Operator: TRANSFORM <tracks> on TAC PIC: *Tracks* to <track groups> on TAC PIC: *Groups and Patterns*>

GOAL: Determine composition of group ...*if part of group or pattern*

• • •

It has a number of conditions under which it is appropriate, one of which is that a new track has been posted on the "Interest Track" level of the THREAT STATUS panel of the blackboard. The body of the task is composed of goals including "locate the track [on the display]" and "review the track data [shown on the display]." Some cognitive operators are included which indicate decision processes (e.g., "TRANSFORM <track data> on TAC PIC: *Track Data Elements* to

<tracks> on TAC PIC: *Tracks"* indicates that the AAWC would at that point in the task evaluate data about a track from multiple potential sources and form an integrated picture of the track, posting the result on a higher level.

Perceptual Demons

Perceptual processing involves translating sensed information to symbolic terms and making it accessible to the cognitive system by posting it on the problem representation blackboard. Each perceptual demon describes *how* a specific class of cue is processed by the perceptual system, indicating what information is posted or transformed on the blackboard as a result of the processing of a specific cue. Once the information is in the blackboard, it may affect the flow of attention because the task triggers are based on patterns of information on the blackboard. This provides the mechanism for situational changes to affect selection and sequencing of tasks (put differently, this provides the mechanism for data-driven cognitive processes, while the task models provide for goal-driven processes).

The COGNET analysis of the perceptual processes of the AAWC identified 19 key perceptual demons, which respond to either visual cues (display events at the workstation) or auditory cues (voice communication from team members), as listed in Figure 5. In the COGNET notation, each demon is modeled as a production rule, in which the sensing of a specific cue is the antecedent condition, and a (possibly conditional) blackboard operation forms the consequent perceptual process. For example, when the event is a radar's acquisition of a new air track, the resulting demon would be formalized as:

IF air track

POST : "New air track [track number] at time [mission-time] held by Radar with Bearing [bearing], Course [course], Range [range], Speed [speed], Altitude [altitude], and Track Quality [track quality]" on TAC PIC:*Tracks* and THREAT:Unknown *Tracks*

• Radar acquires new track	• Radar loses track
• New track acquired through datalink	• Datalinked track lost by reporting unit
• Combat Air Patrol reports of track behavior	• AAW readiness reports
• State reports	• Doctrine setup
• Link reports	• Electronic Warfare reports (from Electronic Warfare Supervisor)
• Identification reports (from Identification Supervisor)	• Reports from Air Intercept Controller, Air Control Supervisor
• Missile System Supervisor reports	• Battle Group AAW commander orders
• Tactical Action Officer orders	• Commercial air track coming out of commercial air corridor
• Course change report (from Tactical Action Officer)	• Query about track, track group, or engagement
• Change in weapon/warning status	

Figure 5. Perceptual Demons in AAWC COGNET Model

Model Dynamics

The structure and content of blackboard, cognitive tasks, and perceptual demons components define the knowledge that the person needs to perform the AAWC job. However, a COGNET model is intended to be used dynamically as well, and can be used to analyze how this knowledge would be applied dynamically in the context of a specific problem situation.[9] The analyst can trace arbitrarily long threads of information processing, up to and including an entire problem, through the model. For example, a set of display events would result in the corresponding perceptual demons being fired, which would result in certain information being internalized and POSTed on the blackboard. As soon as there is information on the blackboard, the analyst can check (each time the blackboard changes) to see if any cognitive task has been triggered. If it has, the procedural knowledge in that task can be stepped through, one goal, method, or operator at a time.

These procedural traces will proceed through the task, executing the behavioral operators (e.g. PERFORM function), and thus indicating which observable actions should be taken as a consequence of the internal decision process.   Each time the blackboard changes, either through the firing of a new demon or through execution of a POST, UNPOST, or TRANSFORM operator, all the cognitive task triggers will have to be re-examined to see if another task has been activated for execution. If and when this happens, the priority formulae of the newly triggered tasks will have to be evaluated and compared to the priority formula of the currently-active task, to see if one of the newly-activated tasks would have more priority in that context, and therefore capture attention away from the currently active one.

A simplified example of how this model works might start with a perceptual demon posting a message (in the “Unknown Tracks” level of the “Threat Status” panel) that a new unknown track has been spotted on the AAWC’s display. This message might then trigger the “Evaluate Track” task, if it had a high enough priority at that point. This task would use other messages from other panels and levels to decide that this track was interesting and post a message to the “Interest Tracks” level of the “Threat Status” panel. Then, the task might determine to take some action on the track (e.g., investigate with CAP) and post a new message to the “Action Tracks” level (removing the relevant “Interest Tracks” message) and post a CAP investigation plan message to the “Resource Assignments” level of the “Plans” panel. The task would then likely release (or lose) AAWC’s attention, waiting for the plan to be fleshed out and implemented. Once a visual identification from the CAP is received and perceived (triggering another perceptual demon to post a message about it), the “Evaluate Track” task might be triggered again (assuming it has high enough priority). The fact that the track is a commercial airliner would cause removal of the message about this track from the “Threat Status” panel and posting of an updated message (including the fact that it is an airliner) to the “Tracks” level of the “Tactical Picture” panel as a track that no longer had to be monitored.

Lessons Learned

1. While the work done by the AAWC and other personnel in the ship's combat information center certainly focused on the computer-based watchstation, we learned that that work is even more strongly tied to the dynamics of the larger decision-making team. Even though there is always a specific individual in command of the team, all decision making is not performed by that individual. The complex nature of the information, assets, and tasks inherent in AAW require that multiple individuals be involved in the decision-making and command and control process.[10] These interdependencies also requrie that information must flow to many places, and that many individuals must coordinate their activities to avoid chaotic or ineffective operations. The Navy doctrine of ‘command by negation’ also plays a key role in distributing decision-making across a team. Under the command by negation philosophy, subordinates are responsible for making tentative decisions and announcing them as intentions which are passed upward on the command and control hierarchy; the superior individual may then accept the decision by saying nothing or countervail it by explicit negation (which is often followed by an alternative decision/intention).

Implications for Theory. Analyzing this type of team-situated decision making required advancing the COGNET theoretical base beyond a traditional human-computer interaction framework to one that dealt with team-based interactions as well. Two specific conceptual modifications were made to the COGNET framework to accommodate team-situated decision-making. The first was a broadening of the concept of perceptual knowledge beyond simple workstation-based cues (displays and auditory alerts) to include communications with other individuals in the decision-making team. This expansion required the cognitive task analysis notation to deal with loosely structured linguistic information as well as the highly-structured display and alert information. The second broadening of the theoretical framework was to expand the notion of the mental model from a representation of just the problem to a representation of the problem and of the team. In team settings, people maintain declarative knowledge about the team and its structure, organization and roles, along with knowledge about the problem being solved, and we found that this knowledge is just as important in organizing the decision process as is knowledge of the tracks, rules of engagement, etc.

2. Although there is a strong tradition of employing behavioral task analysis in the design of systems and (manual and automated) operator-aids, we found that cognitive task analyses provided different kinds of data that have to be applied in different ways . We learned that cognitive analysis results were able to affect design at a much deeper level than traditional behavioral analyses, which are most easily applied to surface features of the system design, particulary person-machine interface layout and functionaltiy. During the course of this research, various means were explored for using the cognitive task analysis data to direct and inform the design of the decision-support system and human-computer interface for AAW decision-makers.

Implications for Design. The ways of using cognitive models in DSS and HCI design were codified as a set of design principles (see Zaklad & Zachary, 1992, and Zachary et al., 1993). These principles, which concerned both functionality of the DSS and its structure, were articulated at an abstract level for general use but also were tailored to TADMUS application at a more specific level by more detailed principles derived from the content of the COGNET AAW model. For example, one general principle was that the DSS should support coordination among team members. At this level, the principle could apply to any team-oriented DSS, but additional, more detailed, principles relating it to the TADMUS/AAW case were derived. Detailed specifics of this principle suggested that the coordination support specifically focus on:

- the sharing of mental models of team activities (derived from the blackboard representation),
- the transmission and acknowledgment of intentions across the AAW team, because of the prominent role that such communications play in the timely triggering of cognitive tasks; and

• helping each operator in the team know when it would be appropriate to notify or ask for permission from higher authorities.

3. The depth of understanding of the decision-making process that was provided by the AAW COGNET model was also useful in support of training. Because the COGNET model provided a model of an expert-level decision strategy, it could be used analytically to derive observable characteristics of the decision process that could in turn be codified and incorporated into performance assessment and measurement instruments. A team outcome instrument, the Anti-Air Team Performance Index (ATPI), identified preliminary outcomes from the COGNET cognitive task goals and subgoals (Dwyer, 1992). Preliminary individual outcomes contributing to the team outcomes were also determined and codified into an individual outcome measurement instrument, the sequenced actions and latencies index (SALI), drawing on the same COGNET components for the AAWC portion (Johnston et al., in press).

Implications for Training. Because COGNET models of expertise can be used as a benchmark of expertise at a particular position, they can be used to derive measurement instruments for use in performance assessment. Specifically, the cognitive tasks indicate actions that should be performed and the contexts in which they should be performed. Thus, performance predictions can be derived for a given training scenario.

4. In the course of applying the pencil-and-paper cognitive task analysis in the ways described immediately above, we learned that there were many potential applications of a cognitive model that couldn't be accomplished with a paper analysis alone, but that instead required an *executable* version of the model. An executable model refers to a software program that can simulate the cognitive, perceptual and motor activities of a person. Such an executable model could be embedded inside the software of a training system or decision support system to provide many new capabilities, such as:

- generating expectations of desired trainee performance, against which actual performance is measured;

- identifying the possible cognitive bases for trainee performance which does not achieve the expected levels;
- determining possible content for training interventions to improve training performance in the future; and
- acting as surrogate operators in team training settings when all roles can not be staffed by human operators.

During the TADMUS program, such a capability to execute COGNET models did not exist. However, through the efforts to apply it to team training and DSS design, we learned the importance of having such a capability for future use.

Implications for System Development. Executable cognitive models embedded in training or decision support systems promise to enhance the capabilities of such systems, as described above. Thus, development of executable cognitive architectures capable of generating executable models may enhance future systems. The research described here stimulated the development of an executable software version of the underlying cognitive architecture of COGNET as shown in Figure 2 (see Zachary, Le Mentec & Ryder, 1996, for details). The executable architecture provides a software environment in which a COGNET model can be executed, given appropriate sensory stimuli and a (digital) means to implement its actions. When an executable model of the AAWC is provided with the visual and auditory stimuli available to a person at the AAWC watchstation, and is given the ability to implement is actions at the workstation digitally, the model is able to emulate (expert-level) human decision-making performance in that problem.

Conclusions

As a final point in this paper, two observations are offered about the development of executable cognitive models. The first is simply that the required data collection, model-building, and model testing processes become much more complex when the result of a cognitive task analysis becomes a piece of executable software. The original COGENT analysis of the AAWC yielded a pencil-and-paper model that was intended to be read, understood, and applied by a human system or training designer. The model builders could therefore rely on the abilities of the human 'reader' of the model to deal with minor inconsistencies, points that were implied but not explicit, and varying levels of detail. This is not so when the 'reader' of a model is a piece of software that emulates an underlying cognitive architecture. In that case, the model-building process becomes as painstaking as any other software development effort. In particular, each goal within each task has to be fully decomposed to the point that every keystroke needed to work an AAW problem at the AAWC watchstation is included. All ambiguity must be removed, everything must be made explicit, and all details must be specified to the same level, or the model will not be able to execute as intended. The result is that the cognitive task analysis effort expands greatly (perhaps by as much as a factor of two!), but also blends into the software development sphere much more so than when the result was (only) a pencil-and-paper analysis.

This leads to the second observation. With the advent of executable cognitive architectures (which include not only COGNET but others such as SOAR (Laird, Rosenbloom, & Newell, 1987), EPIC [Meyer and Kieras, 1996], and ACT-R (Anderson, 1993), cognitive task analysis may be poised to assume a much more prominent role in the development of future systems for training and decision support. The executable architecture allows the (more extensive) cognitive task analysis to be rapidly and directly transitioned into working software that can be incorporated into the application itself, in addition to its existing use as an analytical support for design and evaluation. Such a transition does not eliminate the current uses of cognitive analyses and cognitive models in system design, but rather enhances these roles by integrating cognitive analysis and modeling into most phases of the system life cycle, including the three largest -- implementation, maintenance and support. The changes that this enhanced role could bring are potentially enormous, although the full implications are perhaps difficult to imagine.

REFERENCES

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Australian Civil Aviation Authority (1994). *Air Traffic Control Abinitio Training Program*. Canberra, Australia: Civil Aviation Authority, Air Traffic Services, ATS Human Resources.

Barwise, J., & Perry, J. (1983). *Situations and Attitudes.* Cambridge, MA: MIT Press.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1990). Cognitive Psychology and Team Training: Training Shared Mental Models of Complex Systems. *Human Factors Society Bulletin*, 33, 12.

Cannon-Bowers, J., & Salas, E. (1997) Decision Making Under Stress: Theoretical Underpinnings. This volume.

Card, S., Moran, T., & Newell, A. (1983). *The psychology of computer-human interaction.* Hillsdale, NJ: Lawrence Erlbaum Press.

Carver, N., & Lesser, V. (1994). Evolution of blackboard control architectures. *Expert Systems with Applications, 7(1*), 1-30.

Chi, M.T.H., Glaser, R., & Farr, M.J. (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioral and Brain Sciences, 4,* 317-370.

Dwyer, D. J. (1992). An index for measuring naval team performance. *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1356-1360). Santa Monica, CA: The Human Factors Society, Inc.

Englemore, J., & Morgan, T. (1988). *Blackboard systems.* Reading, MA: Addison-Wesley.

Ericsson, K. A., & Kinsch, W. (1995). Long-term working memory. *Psychological Review, 102 (2),* 211-245.

Ericcson, K. A., & Smith, J. (Ed.) (1991). *Toward a general theory of expertise: Prospects and limits.* Cambridge: Cambridge University Press.

Essens, P., Fallesen, J., McCann, C., Cannon-Bowers, J., & Dorfel, G. (1995). *COADE - A Framework for Cognitive Analysis, Design, and Evaluation* [Technical Report AC/243 (Panel 8)TR/17]. Brussels, Belgium: NATO Defence Research Group.

Hayes-Roth, B. (1985). A blackboard architecture for control. *Artificial Intelligence, 26,* 251-321.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A Cognitive Model of Planning. *Cognitive Science, 3,* 275-310.

Hoffman R. R. (1992). *The Psychology of Expertise: Cognitive Research and Empirical AI.* New York: Springer-Verlag.

Johnston, J.H., Poirier, J., & Smith-Jentsch, K. A. (this volume). Decision making under stress: Creating a research methodology. In J. A. Cannon-Bowers & E. Salas (Eds.), *Decision making under stress: Implications for individual and team training.* Washington, D.C.: APA.

Kieras, D., & Meyer, D. E. (1995). *An overview of the EPIC architecture for cognition and performance with application to human-computer interaction.* (EPIC Tech. Rep. No. 5 (TR-95/ONR-EPIC-5). Ann Arbor, MI: University of Michigan, Electrical Engineering and Computer Science Department.

Kieras, D. (forthcoming in 1997). Task analysis and the design of functionality. *CRC Handbook of Computer Science and Engineering.* CRC Press.

Klein, G. (1989). Recognition-primed decisions. In W. Rouse (Ed.), *Advances in man-machine systems research, 5,* 47-92. Greenwich, CT: JAI Press, Inc.

Klein, G., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics, 19(1),* 462-472.

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. (Eds.) (1993). *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Kolodner, J. (1988). *Proceedings of the Case-Based Reasoning Workshop.* San Mateo, CA: Morgan Kofffman.

Laird, J. E., Newell, A., & Rosenbloom P. S. (1987). SOAR: An architecture for general Intelligence, *Artificial Intelligence*, *33(1),* 1-63.

Meyer, D. E., & Kieras, D.E. (1996). *A Computational Theory of Executive Cognitive Processes and Human Multiple-Task Performance: Part 1. Basic Mechanisms* [EPIC Report No. 6 (TR-96/ONR-EPIC-06)]. Ann Arbor, MI: University of Michigan (to appear in *Psychological Review*, 1997).

Newell, A. (1980) Physical symbol systems. *Cognitive Science, 4.* 135-183.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard Univ. Press.

Newell, A., & Simon, H. (1972). *Human problem solving.* Englewood Cliffs, N.J.: Prentice Hall.

Nii, P. H. (1986a). Blackboard systems: The blackboard model of problem solving and the evolution of blackboard architectures PART ONE, *AI Magazine , 7,(2),* 38-53.

Nii, P. H. (1986b). Blackboard systems: The blackboard model of problem solving and the evolution of blackboard architectures PART TWO, *AI Magazine*, *7(3),* 82-106.

Pylyshyn, Z. (1984). *Computation and cognition.* Cambridge, MA: The MIT Press.

Rouse, W., Geddes, N. D., & Curry, R. E. (1987). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. *Human-Computer Interaction, 3*, 87-122.

Ryder, J.M., & Zachary, W. (1991). Experimental validation of the attention switching component of the COGNET framework. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 72-76), Santa Monica, CA: Human Factors Society.

Ryder, J.M., Weiland, M.Z., Szczepkowski, M.A., & Zachary, W.W. (in press). Cognitive engineering of a new telephone operator workstation using COGNET. *International Journal of Industrial Ergonomics*.

Ryder, J.M., Zachary, W.W., Zaklad, A.L., & Purcell, J.A. (1994). *A cognitive model for Integrated Decision Aiding/Training Embedded Systems (IDATES)* (Technical Report NTSC-92-010). Orlando, FL: Naval Training Systems Center.

Seamster, T. L., Redding, R. E., Cannon, J. R., Ryder, J. M., & Purcell, J. A. (1993). Cognitive task analysis of expertise in air traffic control. *International Journal of Aviation Psychology,. 3(4)*, 257-283.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning, in *Proceedings of the Symposium on the Mechanization of Thought Processes*, H.M.S.O.pp. 511-529.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47*, 513-539.

Zachary, W. & Ryder, J. (1997) Decision support systems: Integrating decision aiding and decision training. In Helander, M., Landauer, T., & Prabhu, P. (Eds.) *Handbook of Human-Computer Interaction, 2nd Edition.* Amsterdam, The Netherlands: Elsevier Science.

Zachary, W. & Ross, L. (1991). Enhancing human-computer interaction through use of embedded COGNET models. In *Proceedings of the 35rd Annual Meeting of the Human Factors Society* (pp. 425-429). Santa Monica, CA: Human Factors Society.

Zachary, W., Le Mentec, J-C., & Ryder, J. (1996). Interface agents in complex systems. In Ntuen, C. and Park, E. H. (Eds.), *Human interaction with complex systems: Conceptual Principles and Design Practice.* Norwell, MA: Kluwer Academic Publishers.

Zachary, W., Zaklad, A., Hicinbothom, J., Ryder, J., Purcell, J., & Wherry, Jr. (1992). COGNET Representation of Tactical Decision-making in Ship-Based Anti-Air Warfare. CHI Systems Technical Report 920211.9009. Springhouse, PA: CHI Systems, Inc.

Zaklad, A. and Zachary, W. (1992). *Decision Support Design Principles for Tactical Decision-Making in Ship-Based Anti-Air Warfare.* CHI Systems Technical Report 920930.9000. Springhouse, PA: CHI Systems, Inc.

Zaklad, A., Hicinbothom, J., & Zachary, W. (1993) Decision-Support System Human-Computer Interface Design Concept. CHI Systems Technical Report 930930.9000. Springhouse, PA: CHI Systems, Inc.

Zubritzky, M., Zachary, W., & Ryder, J. (1989). Constructing and applying cognitive models to mission management problems in air anti-submarine warfare. In *Proceedings of the Human*

Factors Society 33rd Annual Meeting (pp. 129-134). Santa Monica, CA: Human Factors Society.

[1]The authors wish to acknowledge the contributions made by Janine Purcell to the research reported here, as well as the cooperation and effort of the many indiviuals who acted as subjects in the data collection effort. The efforts of Don MacConkey and John Pollen in supporing the data collection and analysis effort on the Navy side are also gratefully acknowledged.

[2]This is the main attention process in COGNET, and is based on the Pandemonium model originally developed by Selfridge (1959).

[3]This perhaps distinguishes it from other frameworks such as SOAR (Laird, Newell and Rosenbloom, 1987; Newell, 1990) or EPIC (Kieras and Meyer, 1995) which have been principally developed as vehicles to develolp and test psychological theory.

[4]This analysis is reported in greater detail in Zachary, Zaklad, Hicinbothom, Ryder, Purcell and Wherry (1992).

[5]Specifically, the simulator at the Combat Systems Engineering Development Site in Moorestown, New Jersey, was used.

[6]The AAWC typically listens to *both* of these networks simultaneously via the headset, which can receive a different communication channel on each earpiece!

[7] The experts were from the Navy's Surface Warfare Development Group or SWDG.

[8]This level of detail is sufficient to indicate the content of the cognitive processing within each activity, without providing an overwhelming level of detail. It also allows the cognitive organization of the task to be examined, without reference to the (sometimes sensitive) details of the particular combat system involved.

[9]At the time of this research, the underlying mechanism had not yet been reduced to a fully executable architecture. Subsequently, however, an executable architecture was created (Zachary, LeMentec and Ryder, 1996), and is being used to create fully executable versions of the model described here, as discussed in the conclusions to this paper.

[10] This is further complicated by the intertwining of AAW and other aspects of ship and battle group operations, such as anti-surface warfare, anti-submarine warfare, or air resources coordination.