Cognitive
Engineering
Laboratory

Development of an Analytical Framework and Measurement Tools to Assess Adaptive Performance of Individual and Teams in Military Work Domains

John R. Hajdukiewicz, Kim J. Vicente, Robert G. Eggleston

CEL 99-01

VELUT ARBOR ÆVO

Cognitive Engineering Laboratory Department of Mechanical & Industrial Engineering University of Toronto
5 King's College Rd. Toronto, Ontario, Canada M5S 3G8
Phone: +1 (416) 978-7399 Fax: +1 (416) 978-3453
Email: benfica@mie.utoronto.ca URL: www.mie.utoronto.ca/labs/cel/

Cognitive Engineering Laboratory

Director: Kim J. Vicente, B.A.Sc., M.S., Ph.D.

The Cognitive Engineering Laboratory (CEL) at the University of Toronto (U of T) is located in the Department of Mechanical & Industrial Engineering, and is one of three laboratories that comprise the U of T Human Factors Research Group. CEL began in 1992 and is primarily concerned with conducting basic and applied research on how to introduce information technology into complex work environments, with a particular emphasis on power plant control rooms. Professor Vicente's areas of expertise include advanced interface design principles, the study of expertise, and cognitive work analysis. Thus, the general mission of CEL is to conduct principled investigations of the impact of information technology on human work so as to develop research findings that are both relevant and useful to industries in which such issues arise.

Current CEL Research Topics

CEL has been funded by Atomic Energy Control Board of Canada, AECL Research, Alias|Wavefront, Asea Brown Boveri Corporate Research - Heidelberg, Defense and Civil Institute for Environmental Medicine, Honeywell Technology Center, Japan Atomic Energy Research Institute, Natural Sciences and Engineering Research Council of Canada, Nova Chemicals, Rotoflex International, and Westinghouse Science & Technology Center. CEL also has collaborations and close contacts with the Mitsubishi Heavy Industries and Toshiba Nuclear Energy Laboratory. Recent CEL projects include:

- Studying the interaction between interface design and adaptation in process control systems.
- Understanding control strategy differences between people of various levels of expertise within the context of process control systems.
- Developing safer and more efficient interfaces for computer-based medical devices.
- Designing novel computer interfaces to display the status of aircraft engineering systems.
- Developing and evaluating advanced user interfaces (in particular, transparent UI tools) for 3-D modeling, animation and painting systems.

CEL Technical Reports

For more information about CEL, CEL technical reports, or graduate school at the University of Toronto, please contact Dr. Kim J. Vicente at the address printed on the front of this technical report.

Abstract

Military work domains are undergoing a paradigm shift characterized by increasing complexity, integration, diversity, dynamic nature, distributed control and coordination, and need for knowledge and information. As a result, there will be an increasing necessity for flexible, adaptive behavior of individuals and teams. Accordingly, it is important that empirical measures of adaptation be developed to assess the effectiveness of different types of systems design interventions in these new environments. In this report, we have adopted Rasmussen's cognitive work analysis (CWA) framework as a basis for addressing this problem. CWA is a constraint-based approach that is tailored to the need for worker adaptation. We show how the layers in CWA can be used to develop and organize empirical measures of adaptation. Several example measures are discussed based on previous research conducted for a process control microworld simulation (DURESS II). In addition, new measures are proposed based on dynamical systems and stability concepts, and higher-order moments of response distributions. Next, case studies showing how our approach can be applied to command and control type problems are presented. Finally, we conclude by briefly outlining an agenda for future research.

1 Introduction

Like so many other work domains, military operations is undergoing a paradigm shift characterized by a number of changes (Morefield, 1995):

1. There is a greater emphasis on acquisition and communication of knowledge and information. As a result, the demands experienced by people will be increasingly cognitive in nature. Cognitive demands can usually be satisfied in many different ways, thereby requiring people to engage in discretionary decision making.
2. Operations are becoming increasingly dynamic, changing at a faster pace than in the past. Thus, the need to react quickly to change and novelty will become the norm rather than the exception.
3. The level of complexity and integration is increasing. Small changes in one part of the work domain can propagate to other, seemingly distant parts. As a result, the need to deal with changes that originated outside of one's local domain of control will increase. Demands for global coordination are becoming more severe.
4. Human and machine actors are becoming increasingly distributed. Rather than having a centralized mechanism that is solely responsible for control, there will be more of a need for coordination to occur locally. Personnel will need to make sure that their actions are compatible with those of others.
5. The diversity of military operations is also increasing. There is a greater need to perform many different types of tasks or missions on a global basis. Consequently, personnel will have to become increasingly flexible, becoming ready to configure to the unique demands of a very wide range of problems.

These trends, albeit different in nature, have a common effect. They signify a transition from a more closed system, where demands can be anticipated and planned-for up front, to a more open system, where demands cannot all be anticipated beforehand and therefore must frequently be dealt with on-line in real-time (Vicente, in press). This increasing need to deal with unanticipated events increases the requirement for adaptation on the part of both individuals and organizations.

This paradigm shift has important consequences for the measurement of human performance. Traditional approaches to measurement were developed within, and are most appropriate for, the previous paradigm. For example, counting the number of errors someone makes is feasible in the case where the situation can be analyzed ahead of time, and a “golden standard” identified. Any deviation from this standard can be considered an error. In contrast, under the new paradigm, it is very difficult to develop a golden standard because the task demands are so diverse, dynamic, and unpredictable. Without a referent by which to assess deviations, it is very difficult to measure errors. The general lesson seems to be that a new approach to human performance measurement is needed for turbulent environments teeming with change and novelty.

The purpose of this report is to propose a framework for the measurement of adaptation that can be applied to complex sociotechnical systems, particularly military operations. This framework will provide direction in establishing appropriate measures that capture adaptive performance in changing circumstances. First, we will describe the need for such a framework. Second, we will identify a cognitive engineering approach to this problem, based on the cognitive work analysis (CWA) framework developed by Rasmussen, Pejtersen, & Goodstein (1994) (see also Vicente, in press). Third, we will show how CWA can be used for the more specific purpose of the measurement of adaptation. Two example measures based on previous work will be described. In addition, several novel measures will be proposed. Fourth, we will describe an example of how the framework and measures may be used for intentional systems (e.g., emergency ambulance dispatch management and military command and control). Finally, a plan for future empirical research will be put forth based on the ideas developed in this report.

2 Adaptation Requires a Constraint-based Approach

Complex work situations are created in work environments where individuals and teams solve complex, open-ended problems under high stress, time pressure, and uncertainty. Consequently, workers must be highly flexible to solve problems, make timely and critical decisions, and coordinate actions in often chaotic and uncertain circumstances. Individual and team performance in these types of situations are generally characterized as being highly adaptive and nonlinear, reflecting the ability to take advantage of opportunities as they arise, as well as to compensate for shortages and disruptions to the nominal work processes. To study such situations, we need to choose evaluation measures that are well suited to these characteristics.

2.1 Structuring the Problem

A useful method for narrowing the possibilities is represented in the hierarchy shown in Figure 1. The three conceptual levels -- theoretical approach, analysis framework, and measurement tools -- provide a basis for strategically selecting, in a top-down manner, methods that are appropriate for the purposes of analysis and systems considered. The choice of theoretical approach is based on the phenomena of interest or the characteristics of what needs to be modeled and studied. The choice of analysis framework is based on what needs to be evaluated and how the evaluation can be structured, in the context of the chosen theoretical approach. The choice of measurement tools is based on the methods that are appropriate for evaluation and how to perform the evaluation, in the context of the chosen analysis framework. Note that the levels are nested in the sense that higher levels constrain the feasible alternatives at lower levels.

An example of how this method can be used is shown in Figure 1. The theoretical approaches illustrated are the constraint-based and behavior-based approaches. Each of these approaches can be instantiated in various analysis frameworks. For example, cognitive work analysis (CWA) is an instance of a constraint-based approach, while traditional task analysis is an instance of a behavior-based approach. Conceivably, many other analysis frameworks could potentially fit into these theoretical approaches. Moving onto the next level in Figure 1, each analysis framework can be instantiated in various measurement tools. For example, the CWA framework can be realized in variability measures at multiple levels in an abstraction hierarchy representation (this measure will be discussed later in the report). The traditional task analysis framework can be realized in measures of error type and frequency. Many other measurement tools could potentially be fit into either analysis framework. Some additional measures for the CWA framework are discussed in Section 4.

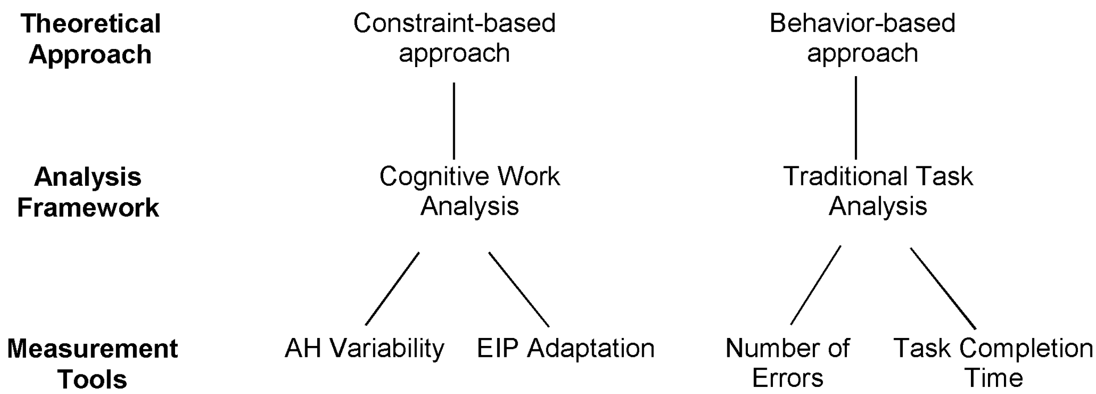

Figure 1: Method for systematically narrowing approaches, frameworks, and measurement tools for evaluating complex work situations.

It is important to consider this top-down method for narrowing the approach for studying work situations. If not, there is a danger that the methods chosen will not be appropriate for the situation considered or the purposes of analysis. This may result in differing and misleading conclusions.

2.2 Constraint- vs. Behavior-based Approaches

Which theoretical approach is most appropriate for the complex work situations that are of interest in this report? To answer this question, we need to better understand the differences between the constraint- and behavior-based approaches. To anticipate, we will conclude that a constraint-based approach is more appropriate for the assessment of adaptive performance of individuals and teams in military work domains.

The difference between the constraint- and behavior-based approaches can be discussed by looking at the differences between fields and trajectories (Vicente, in press). As shown in Figure 2, a negatively charged particle in a magnetic field can be used as an example to illustrate the difference between fields and trajectories (cf. Kugler, Shaw, Vicente, & Kinsella-Shaw, 1990). The upper part of the figure shows the hypothetical magnetic field created by two identical, positively charged particles. The bottom of the figure shows the cross-sectional view of the magnetic field, with the horizontal axis representing spatial position and the vertical axis representing the field strength.

If a negatively charged particle is placed in the field, its trajectory will depend on its initial conditions and the characteristics of the field. For example, if the negatively charged particle is placed to the far right, the field strength is zero and there would be no net movement. If the negatively charged particle is positioned closer to the positively charged particles, the field strength begins to affect its behavior or trajectory. For a particular initial condition, the trajectory is the same for a system that has predictable properties (e.g., unchanging field) and is stable. However, if the system has properties that change and are unpredictable (e.g., changing field due to an unanticipated disturbance), trajectories will differ for the same initial conditions. In addition, if the characteristics of the changing field are not known, trajectories can appear to be random or chaotic because the underlying forces and constraints are not visible.

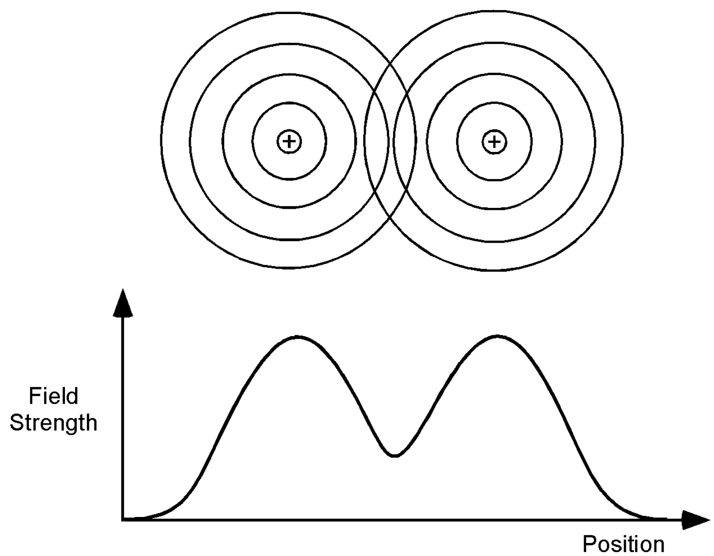

Figure 2: A hypothetical magnetic field created by two positively charged particles (Vicente, in press).

Figure 3 extends the contrast between trajectories and fields to a hypothetical model of a work environment. The constraint-based approach tries to model the work situation in terms of a field description (e.g., constraint boundaries), with no explicit reference to trajectories within the field. Human operators adapt to the constraint boundaries, so trajectories to reach the goal may be determined dynamically and implicitly from any starting position in the action space. In contrast, the behavior-based approach generally specifies trajectories or overt behaviors beforehand, incorporating past experience and expert behavior in the work environment. There is less emphasis on the field description, resulting in limitations in diverting from this prescribed path because the constraint boundaries and other action possibilities may not be apparent to the human operator.

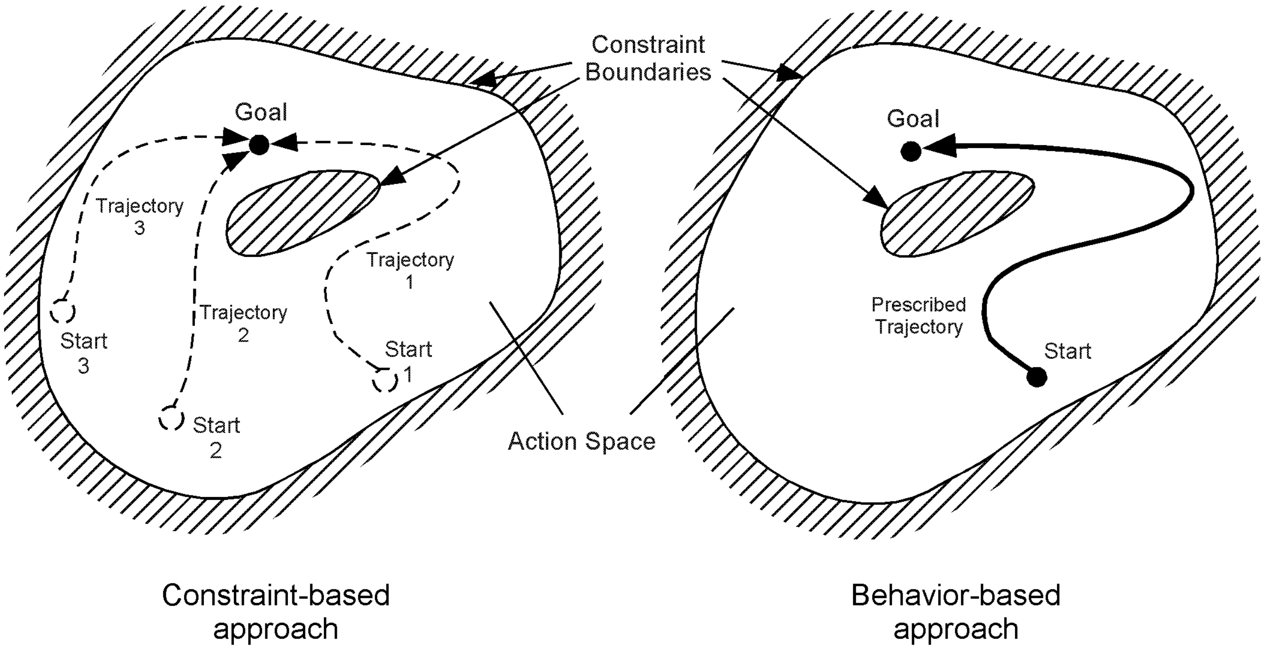

Figure 3: Difference between constraint- and behavior-based approaches.

If the field is relatively dynamic, uncertain, or situation-dependent, then the constraint-based approach may be more appropriate because it provides the flexible support that the human operator needs to adapt to changing conditions. However, if the field is predictable and stable, the behavior-based approach may be more appropriate because there may be no need for the human operator to implicitly generate trajectories dynamically in real-time. Therefore, it is important to consider the characteristics of the system that is analyzed to determine whether the constraint- or behavior-based approach is appropriate.

2.3 A Simple Example

The relevance of the distinction between trajectories and field for the measurement of adaptive performance can be illustrated by the hypothetical automobile driving example in Figure 4. The left side of the figure shows hypothetical trajectories of the car as it passes through a lane. These paths are variable, and thus, may seem random. If we compare the trajectories to the golden standard that we are all taught in driving school (i.e., “keep your car centered in the lane”), then some of the trajectories may deviate substantially from the idealized vision of driving behavior. These are the kinds of insights that might be obtained from a behavior-based approach.

A constraint-based approach would differ because it would look at the extent to which behavior is adapted to goal-relevant constraints, rather than the extent to which behavior trajectories deviate from some idealized, referent trajectory. For example, if the action frequency distribution (i.e., the number of actions the driver puts on the steering wheel) is plotted against the position of the car with respect to the lane boundaries, we might start to develop an idea of the field and how the driver adapts to these constraints with experience.

As shown on the right side of Figure 4, for complete novices, the distribution might be roughly uniform across lane position. There is no sensitivity to the goal-relevant constraints (i.e., the lane boundaries). At an intermediate level of expertise, the driver might become more sensitive to these constraints. This change in adaptation would be indicated by a change in the action frequency distribution. Now, most steering actions are performed when the position of the car is closer to the lane boundary. At an expert level, we might even see that the driver does nothing most of the time. It is only when the car gets very close to the constraint boundaries that a compensatory action is performed.

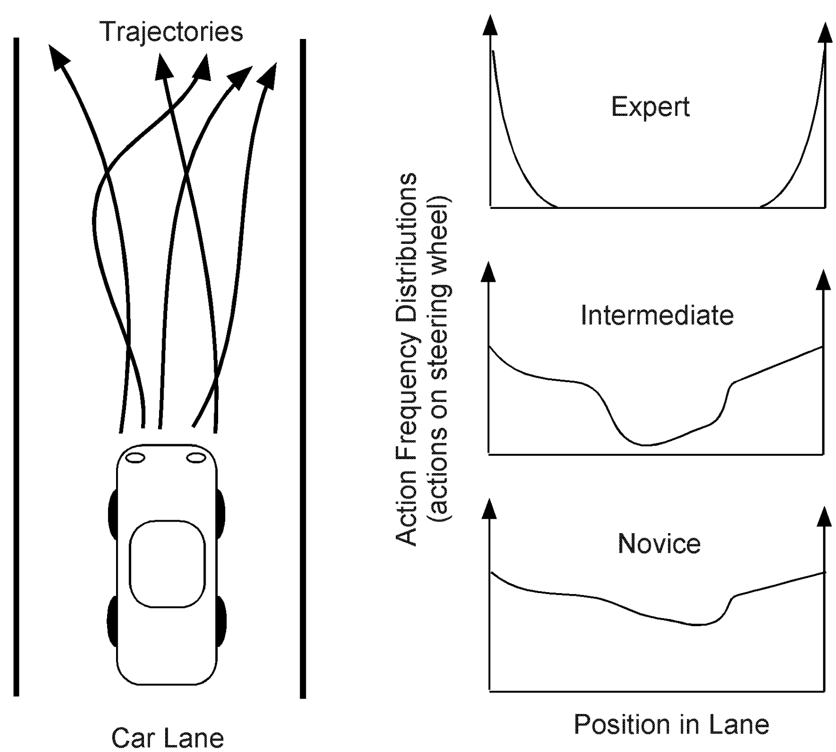

Figure 4: Hypothetical example of car driving.

The insights obtained from these two approaches are quite different. For example, the trajectory- or behavior-based approach might judge the expert's behavior as being eccentric because the car is not always kept right in the middle of the lane (i.e., the referent trajectory defined by the golden standard). In contrast, the field- or constraint-based approach would judge the expert's behavior as being highly adaptive because it is exquisitely tuned to the goal-relevant constraints of the task. Actions are only performed when necessary, thereby achieving economy of movements while still preserving task goals in a flexible manner.

2.4 Constraint-based or Behavior-based? A Resolution

Now that the difference between the constraint-based and behavior-based approaches has been illustrated, we can address the issue of which is more appropriate for the types of problems we are concerned with in this report. As mentioned in Section 1, military work domains are complex sociotechnical systems. Such systems are characterized as having large problem spaces, social coordination and cooperation, heterogeneous perspectives, distributed personnel and components, dynamic behavior, potentially high hazards, many coupled subsystems, automated systems, uncertain data, mediated interaction via computers, and disturbances (Vicente, in press). These characteristics result in a field that is highly dynamic. As a result, human operators need to adapt to change and novelty.

Support for this level of flexibility is not possible with a trajectory-based approach that identifies a golden standard for behavior. The goal of this approach is to minimize variability in behavior, not accommodate it. In contrast, the constraint-based approach does not identify the one best way of performing a task, but instead identifies a constraint envelope that explicitly represents the multiple possibilities for goal-directed behavior. As a result, this approach deliberately accommodates flexible, adaptive behavior. Consequently, a constraint-based approach is tailored to the adaptation required by the new paradigm of military operations.

In the next section of the report, we discuss CWA, a framework that adopts a constraint-based approach to analyzing adaptive performance in complex work environments.

3 A Framework for Cognitive Work Analysis

In developing a method for capturing and analyzing adaptive performance in complex work environments, a framework that assists in structuring the problem would be beneficial. Rasmussen (1986) developed an analysis framework using the constraint-based approach for analyzing complex sociotechnical systems: Cognitive Work Analysis (CWA). This framework identifies five conceptual distinctions or phases of analysis, representing different aspects of work for complex sociotechnical systems: work domain, control tasks, strategies, social organization and cooperation, and competencies. The following discussion of CWA is based on that provided by Vicente (in press).

The first concept, work domain, results in a representation of the system being controlled, independent of any particular worker, automation, event, task, goal, or interface. A work domain analysis is like a map in that it shows the "lay of the land" independently of any particular activity on that land. That is, it shows the possibilities for action. For example, in the nuclear domain, the plant itself could be the work domain.

The second concept, control tasks, are the goals that need to be achieved, independently of how they are to be achieved or by whom. This analysis identifies the product constraints that govern activity on the work domain (as opposed to the constraints that govern the work domain itself). In other words, the focus is on identifying what needs to done, independently of the strategy (how) or actor (who). Continuing with the same example, control tasks would include the activities associated with starting up the nuclear power plant.

The third concept, strategies, are the generative mechanisms by which particular control tasks can be achieved, independently of who is executing them. Whereas control tasks are product representations (descriptions of what is to be accomplished), strategies are process representations (descriptions of how it can be accomplished). They describe how control task goals can be effectively achieved, independently of any particular actors. For example, there may be different processes that could be used to perform the activities associated with starting up the plant.

The fourth concept, social organization and cooperation, deals with the relationships between actors, whether they be workers or automation. This representation describes how responsibility for different areas of the work domain may be allocated among actors, how control tasks may be allocated among actors, and how strategies may be distributed across actors. Thus, a social-organizational analysis describes how actors may be organized into groups or teams, how they may communicate and cooperate with each other, and what authority relationships may govern their cooperation. For example, the startup of a plant could be conducted in a supervisory control mode with a crew of workers monitoring the steps being carried out by the automation.

Finally, the fifth concept, worker competencies, represents the set of constraints associated with the workers themselves. In addition to considering generic human capabilities and limitations, this analysis also identifies the particular competencies that various workers should exhibit if they are to function effectively. Different jobs require different competencies. Thus, it is important to identify the knowledge, rules, and skills that workers should have to fulfill particular roles in the organization effectively. For example, the competencies that the crew of workers would need to function effectively in a supervisory control mode should be identified.

The five concepts can be roughly categorized into two groups. The first group comprises the work domain, control tasks, and strategies. It describes the characteristics of the problem demands that must be satisfied. The second group of concepts comprises social organization and cooperation and worker competencies. It describes the characteristics of the organization and actors who will be responsible for satisfying those problem demands. An ideal sociotechnical system is one that provides a seamless fit between these two groups of concepts.

CWA is based on an ecological approach that emphasizes the importance of the environment in shaping human behavior (Flach, Hancock, Caird, & Vicente, 1995; Hancock, Flach, Caird, & Vicente, 1995). The analysis should begin with the environment because those constraints must be managed to achieve the possibility of reliable and effective performance. When ecological constraints have been dealt with, then it is useful to consider cognitive constraints. For this reason, the five phases of CWA are performed in a particular order, shown in Figure 5, with a gradual transition from ecological to cognitive considerations. Each successive phase in CWA further constrains the space of action possibilities for the human operator. The first layer of constraint to be analyzed is the work domain, which is a representation of the environment to be acted on. The second layer to be analyzed is the set of control tasks, representing what needs to be done to the work domain. The third layer is the set of strategies, representing the processes by which tasks can be done. The fourth layer is the social-organization and cooperative structure, representing the allocation of functions and mechanisms for coordination across many actors. The fifth layer is the set of worker competencies, representing the capabilities and limitations that are required (Vicente, in press).

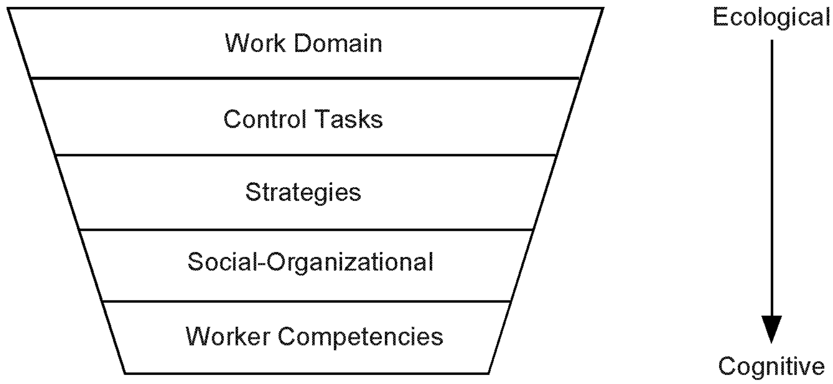

Figure 5: The transition of constraint layers in CWA from ecological to cognitive considerations (from Vicente, in press)

Another way to look at the layers of constraint is shown in Figure 6. This figure shows the dynamic reduction of degrees of freedom and possibilities of action for the actor as each layer of constraint is nested within the previous layers. The map of action possibilities is dynamic because the relevant set of constraints will change based on the particular work situation.

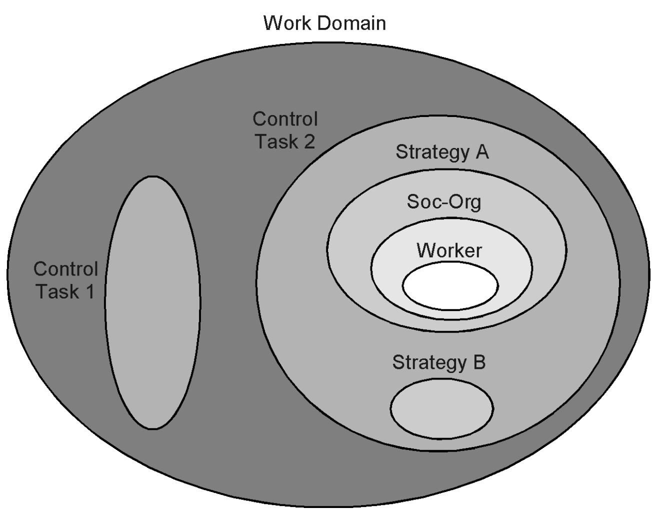

Figure 6: Dynamic reduction of degrees of freedom associated with the nested relationship between layers of constraint of CWA (from Vicente, in press).

The different phases of analysis provide the basis for identifying constraint boundaries in the work environment that shape behavior. Each distinction has a set of modeling tools to assist in identifying these constraints (refer to Table 1). Some of the modeling tools will be discussed in the next section. Vicente (in press) and Ramussen et al. (1994) provide detailed discussions of how these modeling tools can be incorporated in the analysis.

Analysis Phase in CWA	**Modeling Tools**
Work Domain	Abstraction hierarchy
Control Tasks	Decision ladder
Strategies	Information flow map
Social Organization and Cooperation	All the above
Competencies	Skills, rules, knowledge

Table 1: Overview of analysis phases and modeling tools for CWA.

Human behavior or trajectories result from the interaction of the five layers of behavior-shaping constraints. Figure 7 shows that each phase in CWA identifies a particular layer of constraint that effects human behavior. The combined result is behavior that is guided by the interaction between the human operator and the work environment. As we will see in the next section, using this type of analysis framework can lead to developing novel measures of adaptive performance in complex work systems, by considering behavior with respect to the field (i.e., layers of constraint).

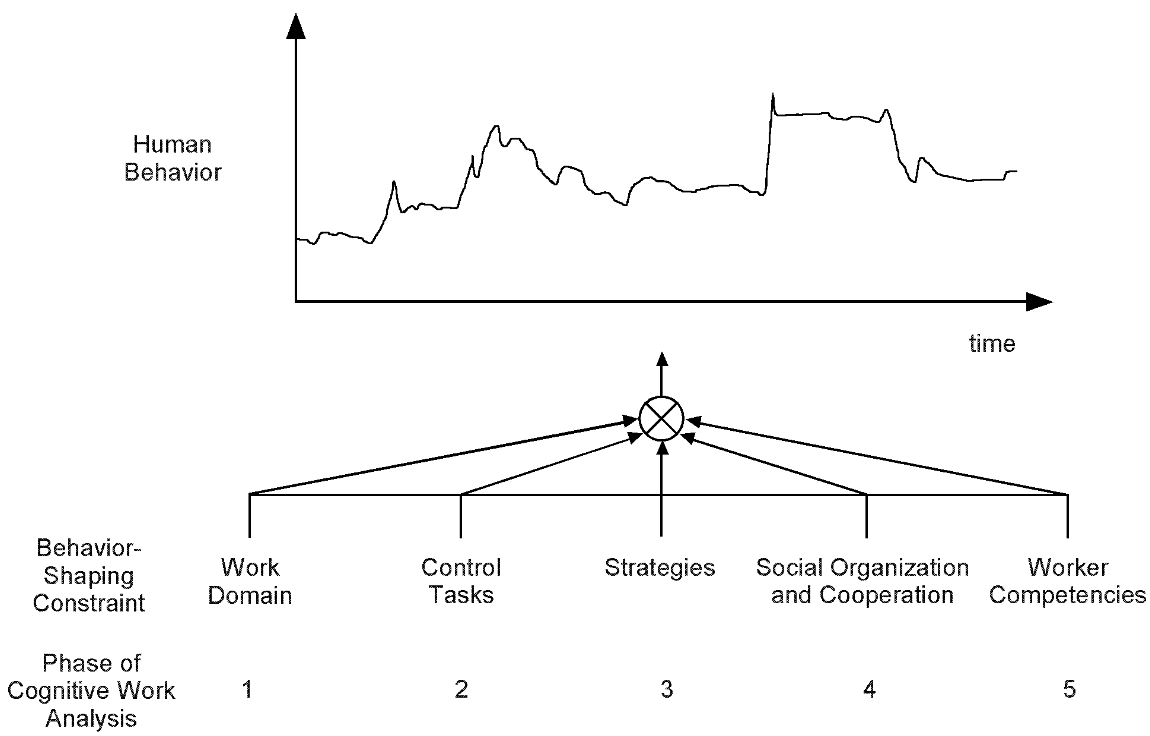

Figure 7: Human behavior and CWA (from Vicente, in press).

4 Measures of Adaptive Performance

There is a very large number of measures that can potentially be used as depending variables for evaluation purposes in complex work environments. The question then becomes which measures should be adopted for a particular study? Furthermore, if a number of different measures are adopted, how can these dependent variables be organized in a meaningful and concise way?

The CWA framework described above provides a basis for structuring measures of adaptive performance by contextualizing behavior with respect to the different layers of behavior-shaping constraint. This section first outlines how the measures may be structured using the framework. Then, two measures of adaptive performance that have been developed from previous research conducted in our laboratory will be discussed. Third, a number of new measures of adaptation are proposed.

4.1 Structuring Measures using the CWA Framework

The CWA framework is useful for structuring measurement tools for complex work environments. The framework can be used to discover and develop measures of adaptive performance of individuals and teams, by partitioning different behavior-shaping constraints. Each layer of constraint defines a distinct class of potential dependent variables. Novel methods for analyzing these variables can provide indications of adaptive performance that are sensitive to the selected behavior-shaping constraints.

Measures can be classified under three general categories with respect to the CWA framework: (1) <u>within-layer</u> measures, (2) <u>nested</u> measures, and (3) <u>generic</u> measures. Within-layer measures are specific to a particular layer of constraint. Nested measures are specific to a group of nested layers of constraint. Generic measures can be applied to any layer of constraint. Examples for each measure category will be provided later in this section. Next, an overview of the types of dependent variables that are specific to each layer of constraint is presented. The discussion is based on that provided by Vicente (in press).

The first layer of CWA, <u>work domain analysis</u>, defines variables that identify the state of the work domain (i.e., state variables). However, sometimes other variables are also useful. For example, higher-order constructs that are derivable from state variables can frequently provide information about the state of a work domain that is perhaps more meaningful than that provided by state variables alone. An example would be a measure of how close the work domain is to exceeding one or more of its safety boundaries.

The second layer of CWA, control task analysis, defines product measures of performance that describe what subjects do. Examples would include: task completion time, number of control actions, number of worker verbalizations, and nonverbal behaviors (note that not all of these measures are based on an adaptation perspective). All of these data can be used to determine what is being done, but by themselves (i.e., without further analysis) they do not indicate how the tasks are being done and by whom because the exact same sequence of overt behavior could have been generated by a number of different strategies and actors. Note that this level of measurement is quite distinct from the previous one; measures of work domain state and product measures of subject performance should not be confused because they are conceptually different.

The third layer of CWA, strategies analysis, defines process measures of performance that describe how subjects do what they do. Such measures go beyond merely describing what workers are doing, and try to capture and describe the strategies that are generating the patterns of behavior identified by the product measures described above. Examples of such process tracing measures include: eye movement analysis, verbal protocol analysis, state-space diagrams, and action transition graphs (Moray, Lootsteen, & Pajak, 1986; Sanderson, Verhage, & Fuld, 1989; Howie & Vicente, 1998).

Note that, in some cases, the same raw data (e.g., worker verbalizations) can be relevant to both this level and the previous one. The difference depends on the extent to which the data are interpreted by the experimenter. For example, if we merely classify verbalizations according to which control tasks they are referring to, then we are in the realm of product measures because all we can reliably infer is what tasks are being done, not how they are being done. On the other hand, if we take those same raw data and process them more extensively (e.g., by looking at sequential dependencies, and cross-correlating multiple measures), we would be in the realm of process measures because we would be developing an understanding of the strategies that are behind the verbalizations. Nevertheless, it is important to keep in mind that the question of what tasks are done and how tasks are done are conceptually distinct because the very same task can be done in multiple ways and the very same sequence of overt behavior could have been generated by different strategies. Therefore, the two types of dependent variables should not be confused because the inferences we can make from each are different.

The fourth layer of CWA, social organization and cooperation analysis, defines measures of team or group communication and cooperation that can be used to understand the degree and quality of coordination across multiple actors. Examples include: the direction and frequency of communication between different actors, measures of nonverbal cooperative behavior, speech act analyses, and the form of communications (e.g., orders, advice, instructions or neutral facts). This level of measurement is important because it investigates the way in which teams/groups/organizations work within the constraints defined by the demands of the work domain, the tasks that are required in that domain, and the different strategies that each actor might use. Thus, while building upon previous levels, this level of measurement contributes towards an understanding of a conceptually distinct set of issues, namely the form, structure, and content of team coordination.

Finally, the fifth level, worker competencies analysis, defines measures that describe subjects' level of expertise. It is useful to distinguish between two different subclasses at this level, one situation-dependent and the other relatively situation-independent. Regarding the former, several different types of on-line measures of competency can be adopted including measures of situation awareness (Endsley, 1995) and mental workload (O'Donnell & Eggemeier, 1986). These dependent variables assess operators' competencies in specific situations. In contrast, situation-independent measures assess competence in a broader fashion. For example, knowledge elicitation measures developed in the psychology and cognitive science literatures (e.g., categorization tasks, transfer of training, control recipes, memory recall tasks) can be used to assess a subject's knowledge of a particular work domain (Christoffersen, Hunter, & Vicente, 1994; Cooke, 1994).

Again, note that this level is distinct from the previous ones because it is possible for workers to perform the exact same task in the same way, or to perform a task using the same strategy, but still exhibit different competencies. As an example of the former, two workers could exhibit the same overt sequence of behaviors but one could be using spatial resources whereas another could be using verbal resources (cf. Wickens, 1992). Similarly, two workers could be performing fault diagnosis using a hypothesis and test strategy, but they could be using very different psychological resources (e.g., if the workers were of different levels of expertise, or if one of them was using an interface that off-loaded part of the demands of that strategy to a computer aid).

The remainder of this section outlines a number of measures using some of the dependent variables discussed, which have been incorporated in previous research. In addition, proposed measures and directions for future measures are included.

4.2 Measures of Adaptive Performance from DURESS II Microworld

A number of measures for adaptive performance have been developed and tested using the DURESS (DUal REservoir System Simulation) II computer microworld simulation environment (Yu, Lau, Vicente, & Carter, 1998; Yu, Chow, Jamieson, Khayat, Lau, Torenvilet, Vicente, & Carter, 1997). DURESS II was designed to be representative (cf. Brunswik, 1956) of industrial process control systems, albeit on a much smaller scale, thereby enhancing the generalizability of research results to practical design problems (Vicente, 1996).

The physical structure of DURESS II is shown in Figure 8. DURESS II consists of two feedwater systems, which supply water to two reservoirs. The operator has control over eight valves (six input valves: VA, VB, VA1, VA2, VB1, VB2, and two output valves: VO1 and VO2), two pumps (PA and PB), and two heaters (HTR1 and HTR2). The operator is required to achieve the dual goals of satisfying external, dynamic demands for water (MO1 and MO2) while also maintaining each of the reservoirs at their respective temperature setpoints (T1 and T2). Two different interfaces for the same microworld were developed for the purpose of conducting laboratory experiments (Vicente & Rasmussen, 1990; Pawlak & Vicente, 1996). The P interface presented only physical information about the work domain. In contrast, the P+F interface presented both physical and functional information about the work domain based on an abstraction hierarchy representation.

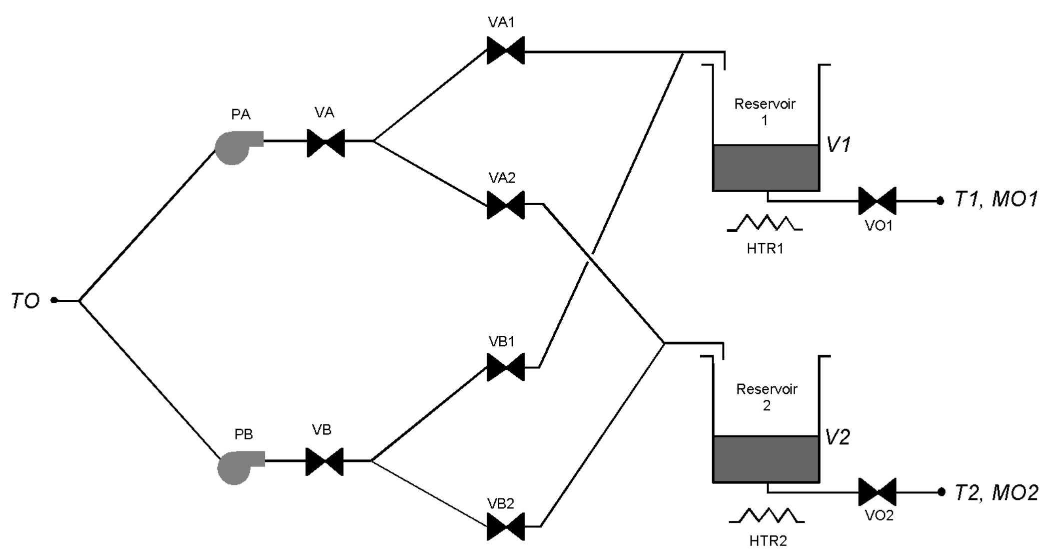

Figure 8: A schematic diagram of the DURESS II process control microworld (Vicente, in press).

Two examples of particular measures that were developed for DURESS II for assessing adaptive performance are discussed next. This discussion is based on the research on long-term adaptation found in Yu et al. (1998) and Yu et al. (1997). Their research included six subjects using different interfaces during the experimental investigation: AS, AV, and IS used the P interface and TL, ML, and WL used the P+F interface.

4.2.1 Abstraction Hierarchy (AH) Variability

The AH variability measure is a novel usage of a work domain analysis of DURESS II using the abstraction hierarchy representation obtained from a work domain analysis as a measurement instrument (i.e., within-layer measure). This can be appreciated by examining Figure 9, which provides an outline of the work domain representation that was developed for DURESS II (Vicente & Rasmussen, 1990; Bisantz & Vicente, 1994). There are three levels of resolution in this space connected by part-whole links (system, subsystem, and component). Also, there are five levels of abstraction connected by structural means-ends links (physical form, physical function, generalized function, abstract function, and functional purpose). The bottom level of physical form will not be used here since it refers to the physical location and appearance of the work domain, features that are not particularly meaningful in a microworld simulation like DURESS II.

Aggregation - Decomposition

Functional ↔ Physical		Whole System (DURESS)	Subsystem (Reservoir)	Component
Functional	Functional Purpose	Outputs to Environment		
	Abstract Function		Mass/Energy Topology	
	Generalized Function			Liquid Flow & Heat Transfer
Physical	Physical Function			Component States
	Physical Form			Appearance & Location

Figure 9: Work domain representation of DURESS II (from Bisantz & Vicente, 1994).

Figure 9 shows that the abstraction and part-whole dimensions, while conceptually orthogonal, are coupled in practice. At higher levels of abstraction (e.g., functional purpose), operators tend to think of the work domain at a coarse level of resolution (e.g., system), whereas at lower levels of abstraction (e.g., physical function), operators tend to think of the work domain at a detailed level of resolution (e.g., component). Therefore, certain cells in the space are not very meaningful (e.g., functional purpose/component). In the specific case of DURESS II shown in Figure 9, four cells were identified as being useful for the purposes of measures of adaptive performance:

1. Functional Purpose/System – outputs to the environment
2. Abstract Function/Subsystem – mass and energy topologies
3. Generalized Function/Component – liquid flow and heat transfer rates
4. Physical Function/Component – component settings

Each of these cells contains a different representation of the very same work domain.

The basic insight behind the analyses was that each of these four representations of DURESS II provided a different frame of reference for measurement. Each frame of reference could then be adopted for the measurement of adaptation. For example, at the top level of Functional Purpose/System, the entire system could be described in a four dimensional space defined by the outputs to the environment (i.e., the goal variables: T1, MO1, T2, MO2). The behavior of the process during one trial could be plotted as a function of time as a trajectory in this space. For a successfully completed startup trial, this trajectory would start at the origin of the space (because the system is shutdown) and would end at the small area defined by the particular goal values for that trial. For this reason, this level was referred to as a goal space. The next level (Abstract Function/Subsystem) described DURESS II in terms of a 12 dimensional space consisting of mass and energy inputs, levels, and outputs: MI1, M1, MO1, EI1, E1, EO1, MI2, M2, MO2, EI2, E2, and EO2. Again, the behavior of the process for one trial could be plotted as a function of time as a trajectory in this space. Because the frame of reference is different at this level of abstraction, we would expect to get different insights into participants' adaptation behavior. Similarly, the third representation (Generalized Function/Component) described DURESS II in terms of a different 10 dimensional space consisting of the liquid flowrates and heat transfer rates being produced by each of the components: FA, FA1, FA2, FO1, FH1, FB, FB1, FB2, FO2, and FH2 (the flows through the pumps were the same as those through valves VA and VB and thus are not included). The fourth and final representation (Physical Function/Component) was somewhat different from the previous ones because it described the state of variables that participants could control directly (i.e., the settings of the components). This space consisted of 12 dimensions, one for each of the controllable components in DURESS II: PA, VA, VA1, VA2, VO1, HTR1, PB, VB, VB1, VB2, VO2, and HTR2 (see Figure 8). In this case, a trajectory in the space was a record of the participant's actions for a particular trial. Thus, whereas the previous trajectories evolved as a function of process dynamics, at this level the trajectory did not evolve until a participant changed the setting of one of the components. It is for this reason that this frame of reference was referred to as an action space.

In summary, the AH from a work domain analysis provided a systematic basis for identifying four frames of references for adaptation measurement. Each cell in the abstraction/decomposition space provided a different representation of the same work domain, and thus a different frame of reference for measurement. Each of these frames of reference provided complementary information, and as a whole, they provided a great deal of insight about operator adaptation that had not been obtained using any other measures. Note, however, that the frames of reference described above were not mutually exclusive. For example, the variable MO1 appeared in more than one level because it was a goal variable as well as a mass variable. Note also that some discretion was involved in choosing the specific variables to be included at each level. For example, we decided to omit the reservoir volumes from the Physical Function/Component description. Although a reservoir is a component, operators have no direct control over the state of the reservoir. By omitting the volumes, this frame of reference was preserved as an action space.

An example of how these different frames of reference and work domain variables were used is now discussed for the AH variability measure. The variability in the trajectories for each participant was calculated at each level of the AH described above, by trial block (see Yu et al., 1997 for a detailed account of the mathematics). The results from the goal space level, Functional Purpose/System, are illustrated in Figure 10. Note that these trajectories were normalized with respect to the setpoint values for each trial, thereby allowing us to meaningfully compare trajectories across trials. The graphs in Figure 10 show the variability in trajectories for each participant over the course of the entire experiment, as a function of 11 blocks of approximately 20 trials each. The data for the best participant in the P and P+F interfaces, TL and AV, respectively, were similar. After the initial part of the experiment, both of these participants exhibited very consistent trajectories at the goal space level. This frame of reference does not allow us to differentiate TL and AV. According to this measure, they were behaving in the same fashion. Figure 10 also shows that, with the exception of ML, all of the other participants were also able to achieve consistently low goal space variances by the latter part of the study. However, we know from previous analyses (see Yu et al., 1997) that there were substantial differences in skill between these subjects, particularly AS and AV. Thus, the goal space variability measure is not a very diagnostic one.

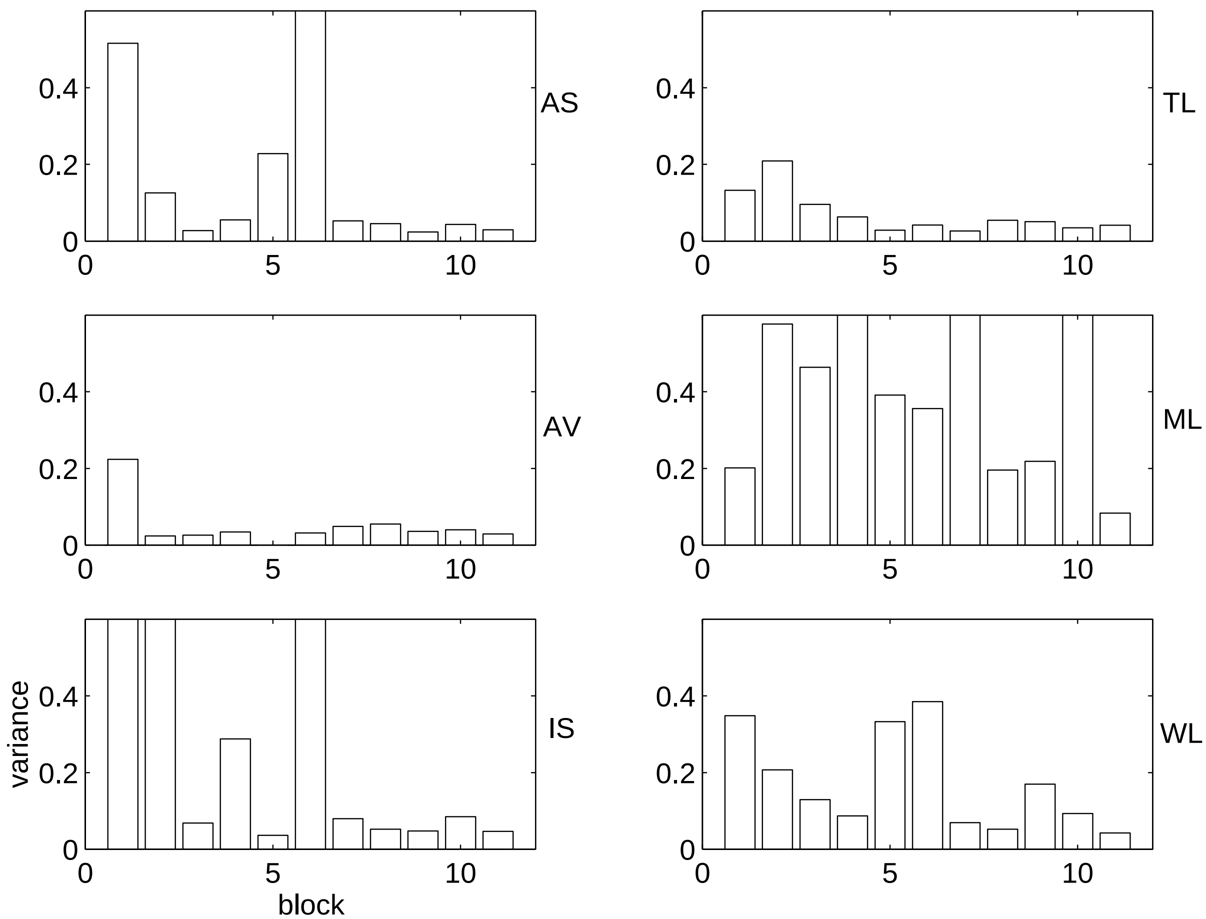

Figure 10: Variance of outputs (from Yu et al., 1998).

Figure 11 shows the results of the variability analysis at the action space level. Note that, at this level, the trajectories were not normalized with respect to the goal values for each trial. Such a normalization was not possible because there is no direct relationship between goal values and component settings. Thus, the variability analysis at this level was based on absolute setting values (with a compensation for the fact that different components have different scale values; see Yu et al., 1997).

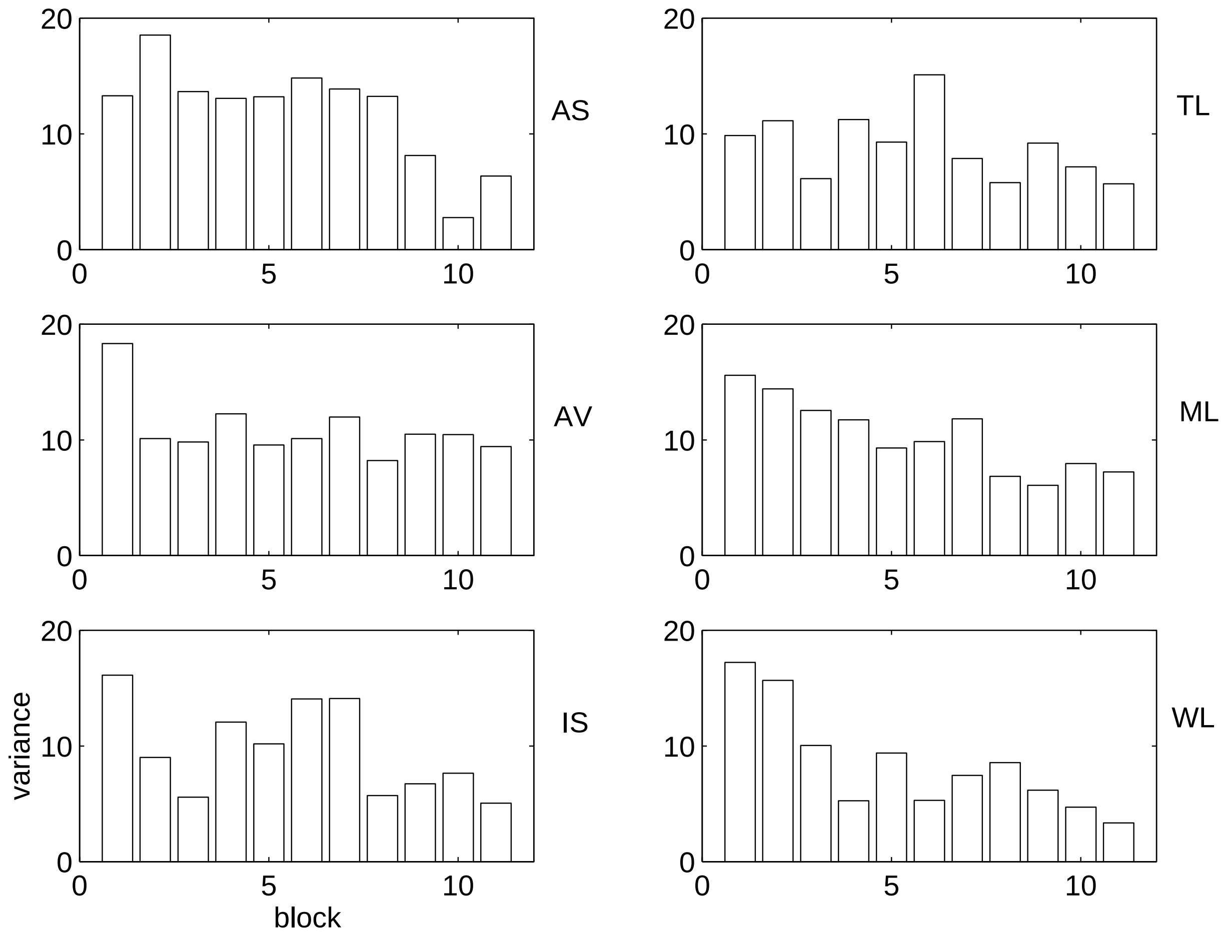

Figure 11: Variances in component settings (from Yu et al., 1998).

The results indicate that the least proficient participants in each group, AS and ML, had the highest action variability, presumably because they engaged in iterative, trial and error behavior. The primary contrast of interest, between AV and TL, is difficult to discern in Figure 11 because of the relatively small differences compared to the scale size. Figure 12 shows a direct comparison of the action variability for AV and TL during the last few blocks of the experiment. These data clearly show that TL's behavior was consistently less variable than AV's. Although the difference was not a large one, it is consistent with the observation that TL's behavior is driven more by a fixed set of specific actions than AV's (Yu et al., 1998).

Figure 13 shows the results of the variability analysis at the Generalized Function/ Component level. The trajectories in this frame of reference were also not normalized with respect to the goal values for each trial for the same reasons stated above. The results from this frame of reference are actually very similar to those in Figure 12, and thus add little to our understanding. In retrospect, the reason for this similarity is straightforward. These two levels, component settings and components flows, are actually strongly correlated. By inspecting the equations describing the system dynamics (see Yu et al., 1997), we see that there is a direct correspondence between these two sets of variables after the transient produced by a change in component setting. In other words, if the component settings were given we can usually uniquely derive the liquid flow rates and heat transfer rates (for normal trials). The only times during which this relationship was weakened was during the short period after an operator action. For example, if VB is closed and a subject sets it to 10, it will take approximately 15 seconds for the valve to actually reach that state because of the first order lag dynamics. After that lag, there will be a fixed relationship between valve settings and flow rates. However, during the preceding transient period while the valve is opening from its closed state to its final state, this relationship was not fixed. It is for these reasons that the results in Figure 13 are so similar to those in Figure 11.

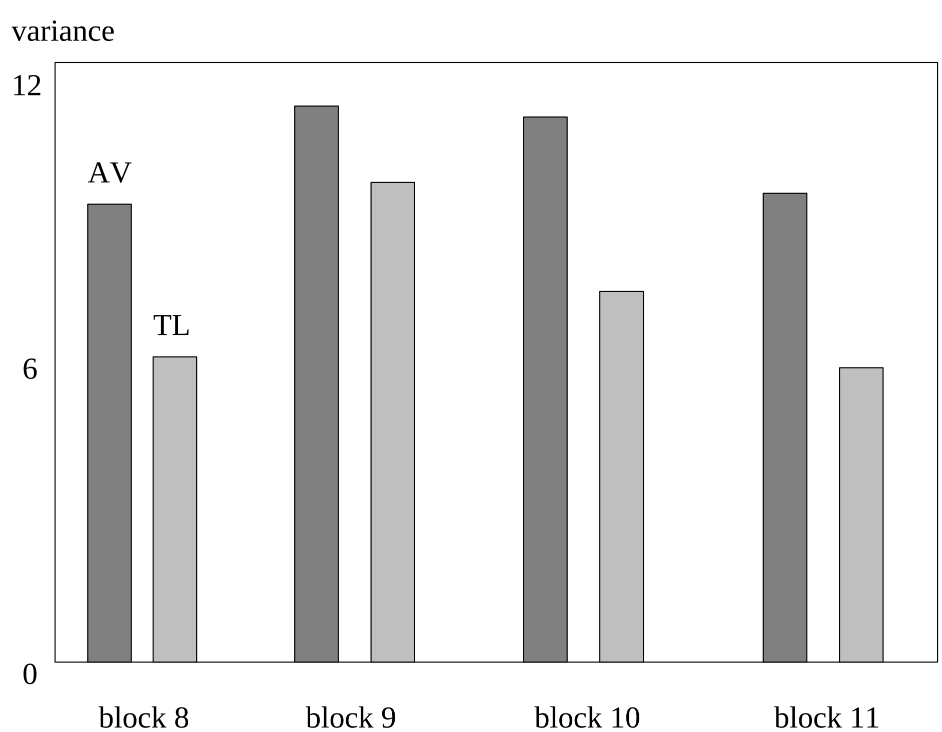

Figure 12: Comparison of variances of component settings for the last four blocks (from Yu et al., 1998).

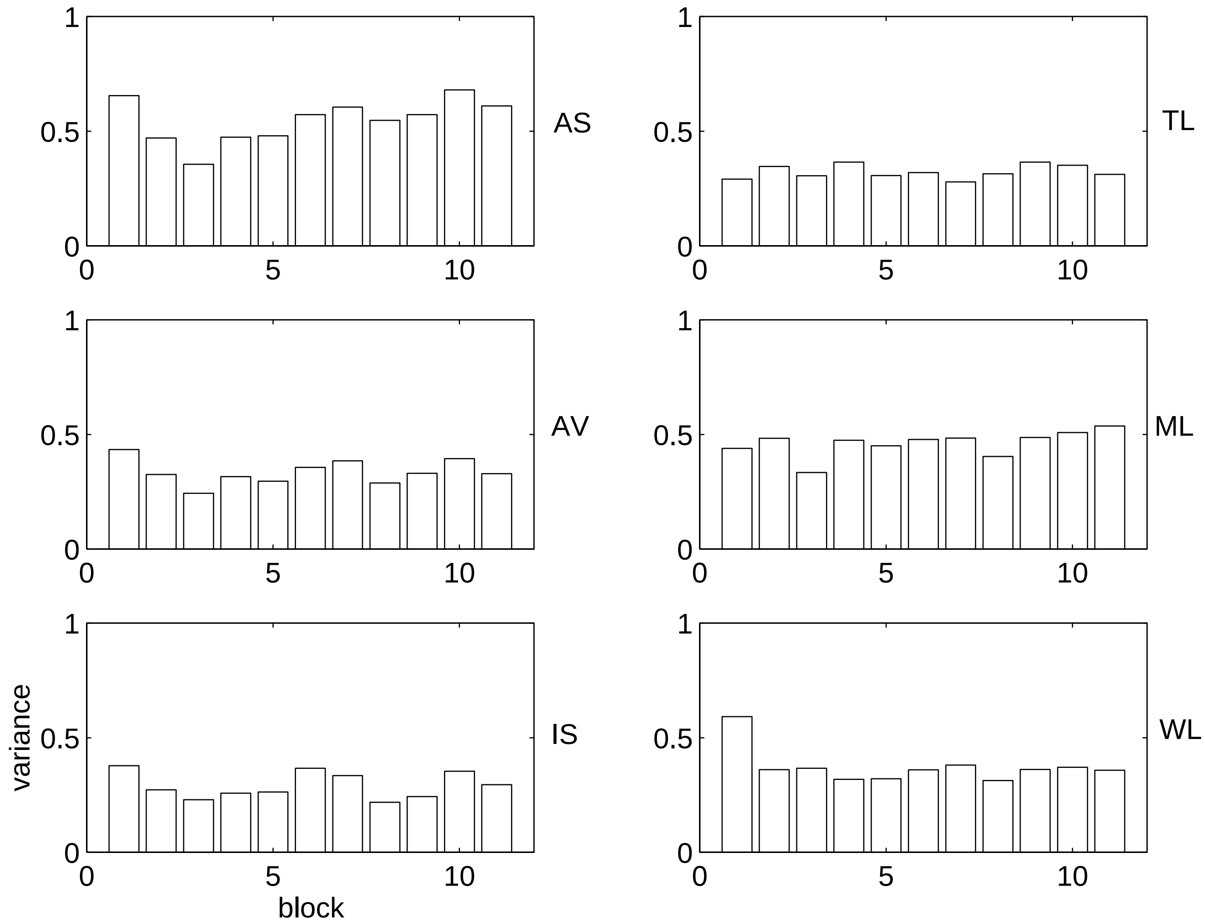

Figure 13: Variance of flows and heat transfer (from Yu et al., 1998).

The final set of AH variance analyses were conducted at the Abstract Function/Subsystem level. There were two interesting differences between this frame of reference and the last two we have just described. First, the measurement was taking place at an aggregate level. Variables were examined at the subsystem level, which were aggregates of the variables that were examined at the part-whole level of components. Second, measurement at this level was in terms of variables that describe the process in terms of first principles (i.e., mass and energy conservation laws). Because each trial had a different set of setpoints for the four goal variables, variance in the trajectories was expected for this reason alone. Although the trajectory for each trial began at the origin (because the system is shutdown), the end point for each trajectory (and presumably the trajectory itself) would be different for different trials as a function of the setpoints for that trials. If it was assumed that participants tried to stabilize both volume and temperature for each reservoir, then it was possible to correct the trajectories for differences in setpoint values across trials. This is accomplished by dividing the mass input and output flowrates (i.e., MI1, MO1, MI2, MO2) by the demand setpoints (D1, D2), and dividing the energy input and output flowrates (i.e., EI1, EO1, EI2, EO2) by the product of the demand setpoints (D1, D2) and the temperature setpoints (T1, T2). The details of these calculations are provided in Yu et al. (1997). Normalizing the trajectories in this fashion eliminated any variability caused solely by differences in setpoints across trials.

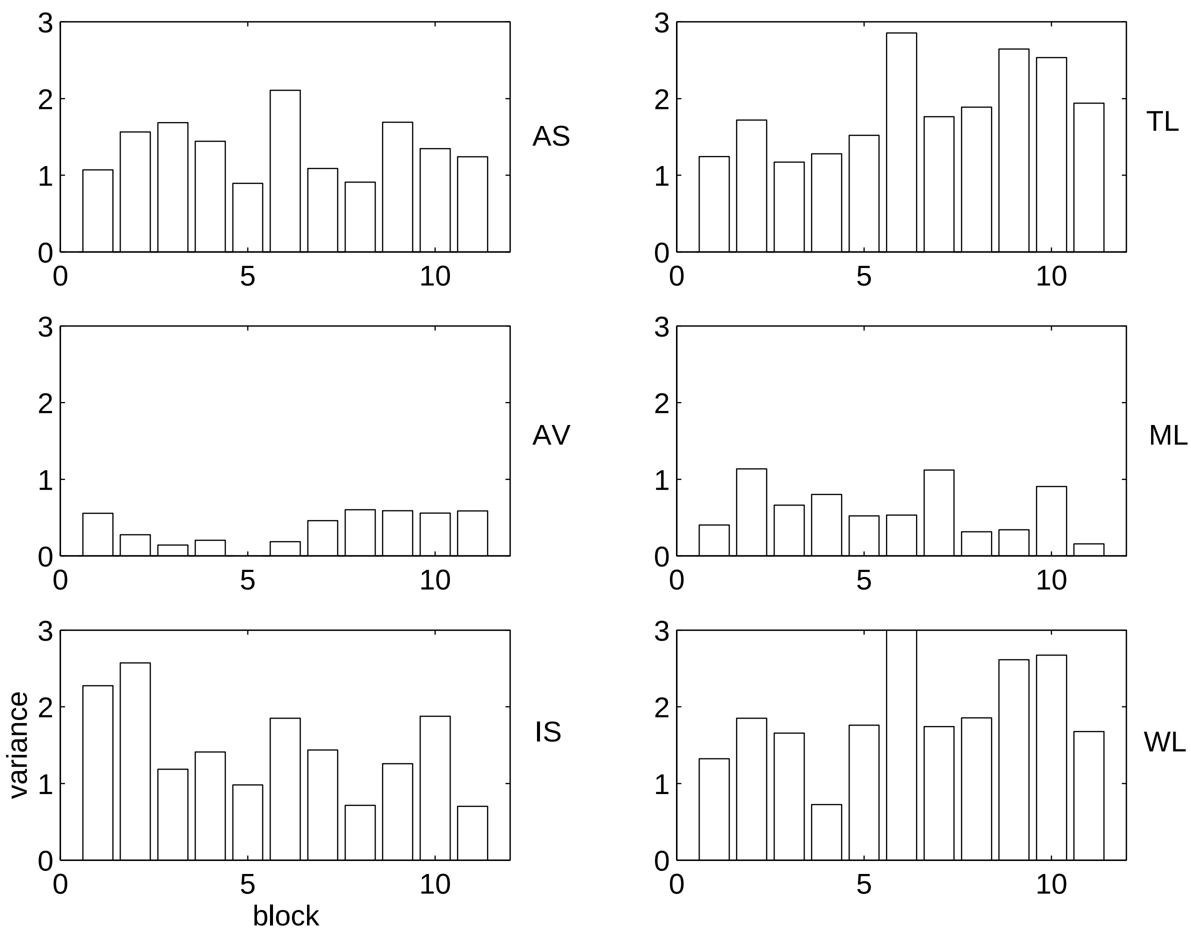

Figure 14: Variance of mass and energy normalized by both scale and goals (from Yu et al., 1998).

There are several interesting findings emerging from this way of looking at the data. The most important of all is the large difference between the variances for TL and AV, the most proficient participants in the P and P+F groups, respectively. From the beginning of the experiment, but especially in the second half, the trajectory variance for AV was much smaller than that for TL. AV was thinking about, and controlling, the work domain at a high level of abstraction, focusing on the mass and energy level. Moreover, he contextualized his control at this level based on the setpoint values for each trial. It is when the differences in setpoint values are compensated for that we see that, at high level of abstraction, AV was acting in a consistent fashion across trials. In contrast, the regularities in TL's behavior were more at the action level (Figure 11) where he exhibited a lower variance than AV. Because TL's actions were relatively similar for trials with different setpoints, his behavior was not as contextualized (or situated) as AV's. Thus, when TL's data was examined at a contextualized, functional level of abstraction, he exhibited less structure than did AV.

Another important result that can be seen in Figure 14 is that ML, perhaps the worst participant in the P group, also exhibited a very low variance at this level of abstraction. In fact, towards the end of the experiment, his variance was sometimes even lower than that of AV. This suggested that he too was controlling the system at a high level of abstraction.

4.2.2 EIP Adaptation

Another measure developed by Yu et al. (1997) to assess adaptive performance was Elementary Information Process (EIP) adaptation. This measure incorporates variables from many layers of the CWA framework (i.e., nested measure) and analyses the extent to which subjects adapt their behavior to their own information processing limitations. The approach was influenced by the work of Payne, Bettman, & Johnson (1993) on adaptive decision making. The fundamental insight behind this work was that people select their strategies based on a demand-resource tradeoff (see also Rasmussen, 1986). In many cases, this choice is adaptive because it allows people to deal with complex tasks despite the fact that they have a limited capacity for information processing. Payne et al.'s work was extended by showing how their ideas could be applied to study expertise a dynamic rather than a static work domain, i.e., DURESS II.

A specific decision task was chosen as the focus of analysis, feedwater stream (FWS) configuration, based on a cognitive work analysis of DURESS II (Vicente & Pawlak, 1994). To successfully complete a startup task, subjects were required to configure the FWS so that it supplied enough water to satisfy the output demand goals for both reservoirs (D1, D2).

The subjects were able to perform this subtask by changing the settings of the two pumps (PA, PB) and the six input valves (VA, VA1, VA2, VB, VB1, VB2). For the most part, the pumps were always turned on, creating flow through the process. Therefore, the analysis was focused on the use of the FWS valves. In particular, the goal of the analysis was to predict the final valve settings that subjects settled on by the end of a trial, and to explain the criteria behind their choice. Yu et al. were not concerned with predicting the intermediate values that the valve settings took on during the trial.

The problem was formulated as follows:

D1, D2 --------> f (VA, VA1, VA2, VB, VB1, VB2)

That is, given the output demand values for a particular trial, what function described how subjects chose to configure the FWS components? This problem was underspecified, with six unknowns and only two givens. There were not enough work domain constraints to identify a unique solution, resulting in an infinite number of solutions paths to the problem. Yu et al. analyzed this problem from a normative and descriptive perspective. The normative perspective outlined the criteria that operators should use to resolve the remaining degrees of freedom, based on the assumption of minimizing mental effort (i.e., number of EIP steps). The descriptive perspective outlined the criteria that operators actually used to resolve the remaining degrees of freedom.

The FWS configuration subtask was formulated more precisely by breaking it down into three, nested steps: a) the number of valves to use, b) which valves to use, and c) the quantitative value to set each valve by the end of the trial. Yu et al. postulated that the way in which subjects answered each of these steps was constrained by two factors: the characteristics of the work domain of DURESS II and psychological criteria.

Normative Analysis

Based on Payne et al.'s (1993) research, a normative analysis of this problem was conducted that was sensitive to psychological constraints. The guiding assumption was that adaptive operators would try to perform the FWS configuration task effectively while simultaneously trying to minimize the information processing demands they experienced. This would require sensitivity to both work domain constraints and psychological constraints. Since subjects could not change the way in which DURESS II was built, they had to obey the work domain constraints if they were to be able to satisfy the task goals. Thus, these work domain constraints are described first.

Work domain constraints. The constraints that the work domain itself imposed on the FWS configuration subtask could be understood by considering the three nested steps previously identified. For the first step, the number of valves to use was constrained by the capacity of the two FWSs. Each pump could produce a maximum of 10 units/s of flow. As a result, each FWS could only supply 10 units/s of flow to the two reservoirs. Because of this structural work domain constraint, the minimum number of input valves (out of 6) that could be used to perform the task could be determined, as a function of the demand values (Vicente & Pawlak, 1994). Three, mutually exclusive and exhaustive cases were distinguished:

Mode 1 - *If (D1 + D2) < 10, then min # of valves = 3*

Mode 2 - *If (D1 + D2) > 10 AND (D1 <= 10) AND (D2 <= 10), then min # of valves = 4*

Mode 3 - *If (D1 > 10) OR (D2 > 10), then min # of valves = 5.*

Of course, subjects were free to use more than the minimum number of valves in any mode. Thus, there was no unique solution to this problem.

For the second step, which valves to use, the work domain imposed an additional set of structural constraints because of the topological connections between valves and reservoirs. More specifically:

If D1 > 0, then (VA AND VA1) OR (VB AND VB1) must be used

If D2 > 0, then (VA AND VA2) OR (VB AND VB2) must be used.

Once again, the work domain constraints limit, but do not uniquely specify, what the operator should do.

For the third and final step, the value that each valve should be set at, there were several quantitative work domain constraints that must be taken into account:

MI1 = FA1 + FB1 – conservation of mass into reservoir 1

MI2 = FA2 + FB2 – conservation of mass into reservoir 2

FA = FA1 + FA2 – conservation of mass in FWS A

FB = FB1 + FB2 – conservation of mass in FWS B

If (VA1 + VA2) > VA then FA = VA – flow is constrained upstream in FWS A

else FA = VA1 + VA2 – flow is constrained downstream FWS A

$FA1 = \dfrac{FA * VA1}{VA1 + VA2}$ *– flow split relation in FWS B*

If (VB1 + VB2) > VB then FB = VB – flow is constrained upstream in FWS B

else FB = VB1 + VB2 – flow is constrained downstream FWS B

$FB1 = \dfrac{FB * VB1}{VB1 + VB2}$ *– flow split relation in FWS B*

The first four are additive constraints that represent the fact that water is neither created nor destroyed in the system. The next four constraints describe the relationships between the valve settings and the flows through the valves. There are two sets of such relationships, one for the case where flow is being constrained downstream and another for when the flow is being constrained upstream. Each can be illustrated by example. In the downstream case, if the initial valve, VA, is set to 10 and VA1 and VA2 are set to 1 and 4, respectively, then the flows will be as follows: FA = 5, FA1 = 1, and FA2 = 4. In the upstream case, if the initial valve, VB, is set to 9 and VB1 and VB2 are set to 4 and 8, respectively, then the flows will be as follows: FB = 9, FB1 = 3, and FB2 = 6. Thus, in the downstream case, the relationship between the valve settings (VA1, VA2) and the flows (FA1, FA2) is a simple one, namely 1:1. In the upstream case, on the other hand, this relationship is more complex, being determined by the ratio of the relative settings of the valves. This difference has important psychological implications (see below).

Psychological constraints. The work domain constraints were not sufficient to provide a unique solution to the FWS configuration subtask. As a result, additional criteria needed to be introduced to deal with the remaining degrees of freedom. It was at this point that psychological constraints were introduced. For each of the three steps identified earlier, normative psychological criteria were proposed that could be used by operators to resolve the remaining degrees of freedom in a way that minimizes their information processing demands.

The first step was how many valves to use. Although there was no unique solution to the problem based on work domain constraints alone, psychological constraints could be used to narrow down the possibilities in a meaningful fashion. Each valve that was used added information processing demands. For example, operators would have to decide what value to set the valve at, and remember to monitor its state. Thus, it seems that an adaptive operator would minimize the number of valves for each of the demand pair modes identified earlier. However, there was a potential disadvantage associated with using the minimum 3 valves in the first mode. The flows to the two reservoirs become coupled because a single FWS was being used to feed both of them. This makes it more likely that the upstream valve mode would be experienced. As a result, a change in valve setting (e.g., VA1 or VA) to affect the flow going to one reservoir (e.g., reservoir 1) could unintentionally affect the other reservoir (e.g., reservoir 2). Using 4 valves obviates this problem because each FWS would be driving a single reservoir. There would be no interactions to worry about. Although it is difficult to be sure, it seems that the disadvantages of using only 3 valves outweigh the advantages. Thus, for Mode 1, it was predicted that adaptive operators would probably use 4 valves. For Mode 2, they should use the minimum 4 valves, and for Mode 3, they should use the minimum 5 valves.

The next step was to determine which valves should be used, given that the number of valves had been established. The work domain analysis (see above) showed that there were basically two options, so once again there was no unique solution. However, a cursory examination of the degrees of freedom allowed narrowing down the solution based on psychological criteria. In this case, stimulus-response compatibility was relied on. For one of the options, the upper (or lower) stream would feed the upper (or lower) reservoir. For example, VA & VA1 would be used to feed reservoir 1. For the other option, the upper (or lower) valve would feed the lower (or upper) reservoir. For example, VA & VA1 would be used to feed reservoir 2 instead. Clearly, the first option is far better since it makes control of the work domain much more natural because of the directness of the mapping (i.e., to control the top reservoir, use the top stream). In contrast, the second option – while equally viable from a work domain point of view – is problematic from a psychological viewpoint. It would result in an inverse mapping (e.g., to control the bottom reservoir, use the bottom stream). Thus, the choice for the adaptive operator seems clear.

The final step was to determine the quantitative values that the valves should be set at. In this case, the work domain analysis showed that there were a number of algebraic constraints that needed to be taken into account in making this decision. Nevertheless, there was still a bounded but infinite number of possibilities. To narrow down the possibilities, Yu et al. introduced additional psychological criteria to deal with the remaining degrees of freedom. The relevant criterion in this case was to minimize the computational effort involved. Recall that the FWS could be operated in two qualitatively different ways, one in which flow is constrained upstream and the other in which flow is constrained downstream. The upstream case was more complex, involving flow split ratio calculations, whereas the downstream case was considerably simpler because of the identity relationship between valve setting and flow rate. Note that these two cases were equivalent in terms of the range of demand pairs they could satisfy so, from a work domain perspective, there is no difference between them. But since the downstream configuration imposed considerably fewer computational demands (see below), it should be always be adopted by the adaptive operator.

In summary, this normative analysis of the FWS configuration task spanned both work domain and psychological constraints. Yu et al. began by identifying the constraints that the work domain imposed on successful performance, since these constraints needed to be respected if they were to perform the task successfully – the laws of physics could be ignored but not escaped. A number of work domain constraints that needed to be taken into account were described, but there were still an infinite number of ways of accomplishing the task. From a work domain perspective, there was no basis for choosing between these different solutions – they were all equivalent. But since DURESS II was being controlled by human operators, psychological criteria were introduced to close the remaining degrees of freedom. The assumption was that an adaptive operator would select solutions that, not only got the job done, but got it done as economically and easily as possible. Table 2 summarizes the criteria that were identified for each of the steps associated with the FWS configuration subtask. These criteria identify, out of all of the feasible ways of performing the task, the way(s) that an adaptive operator was most likely to adopt.

Degrees of Freedom	Psychological Criteria
How many valves?	minimize monitoring demands & interactions
Which valves?	stimulus-response compatibility
What settings?	minimize computations

Table 2: A summary of the psychological criteria used for the normative analysis of the FWS configuration subtask.

If all of the considerations summarized in Table 2 were integrated, along with the fact that D1 never exceeded 10 in our experiment, all of the degrees of freedom in performing the task were used up. In other words, a normative algorithm could be identified that an adaptive operator might use to accomplish the FWS configuration subtask. To show that this was the case, Yu et al. borrowed from the work of Payne et al. (1993) by defining a set of elementary information processes (EIPs) that could be used as a basis for the specification of such an algorithm. Table 3 shows the EIPs that were developed for this particular task. Table 4 shows how these EIPs were used to compose a normative algorithm for the FWS configuration subtask, based on the criteria discussed above. This algorithm provided one possible answer to the question posed at the outset – what function does (or should) an operator use to go from a set of output demand pairs to a set of valve settings?

EIP	Description
READ	Read the demand goals D1 and D2 into short-term memory.
ADD	Add two values.
DIVIDE	Divide value by 2.
SUBTRACT	Subtract one value from another.
COMPARE	Compare two values (greater than or less than).
ASSIGN	Finding appropriate component and assigning a setting value to it.

Table 3: EIPs for performing the FWS configuration subtask in DURESS II (from Yu et al., 1997).

ASSIGN	*VA = 10*	
ASSIGN	*VB = 10*	
READ	*D1*	
READ	*D2*	
ASSIGN	*VA1 = D1*	
COMPARE	*D2 to 10*	
if D2 ≤ 10		
	ASSIGN	*VB2 = D2*
else (D2 >10)		
	ASSIGN	*VB2 = 10*
	SUBTRACT	*D2 – 10*
	ASSIGN	*VA2 = (D2 - 10)*

Table 4: Normative algorithm for the FWS configuration subtask, based on the adaptive criteria discussed above (from Yu et al., 1997).

Results

Yu et al. compared the subjects' actions with the predictions made by the normative analysis. They calculated an error score to measure the difference between the valve settings used by subjects and the valve settings predicted by our analysis. An error score of zero indicated perfect agreement with the normative analysis. They then plotted this error score vs. trial number for each subject. The results are shown in Figure 15. The normative analysis predicted the behavior of TL quite well. Although he started off with a different strategy, the behavior of IS was also well accounted for during the latter third of the experiment. As for AV, there was good agreement during the first third of the experiment and moderate agreement during the remainder. The data for the remaining subjects (AS, ML, and WL) clearly showed that they were not following the normative strategy.

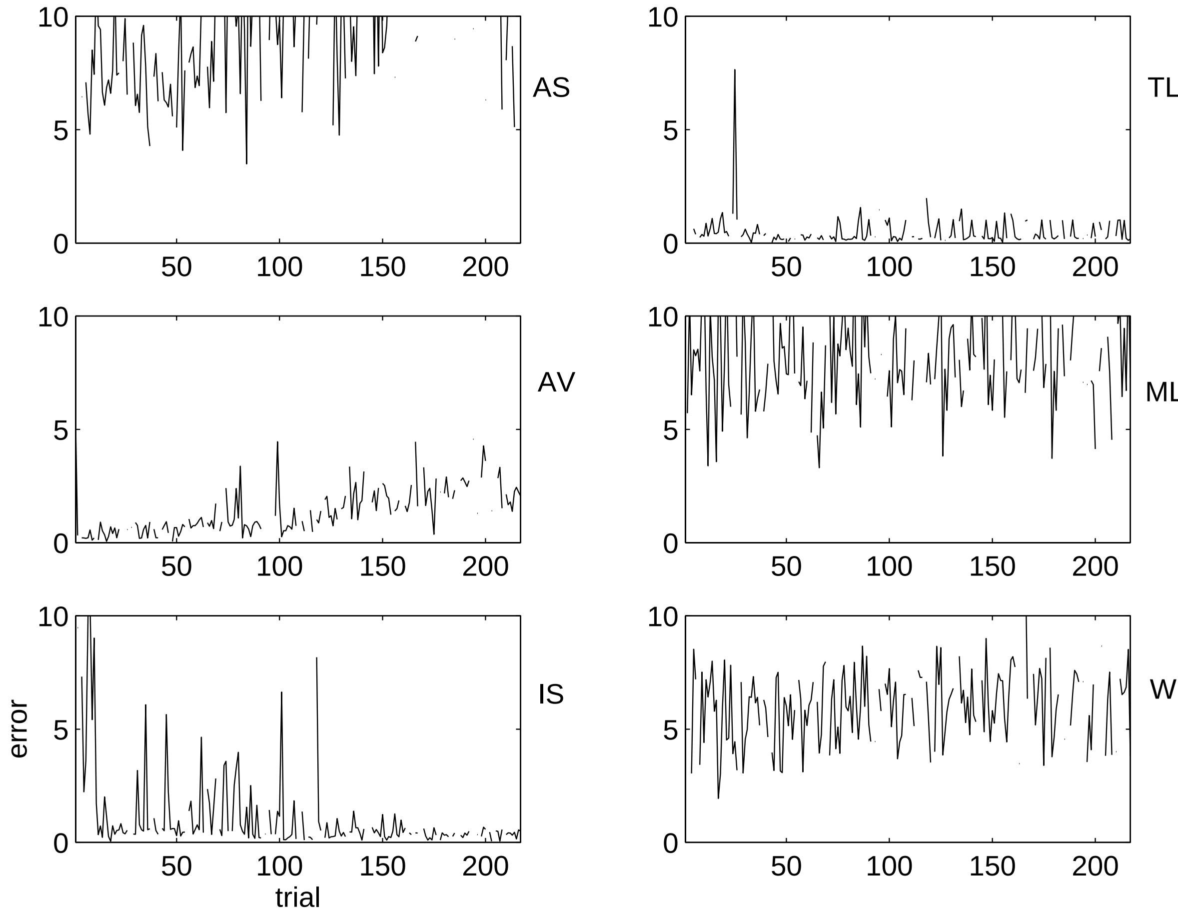

Figure 15: Error between actual valve settings and those predicted by Table 4 (from Yu et al., 1997).

Interestingly, the degree of fit to the adaptive norm seemed to be best for the most proficient subjects. TL was clearly the best in the P group, AV usually the best in the P+F group, and IS a close second (see the performance data in Yu et al., 1997). This observation suggests that there may be a connection between subjects' performance and the extent to which their behavior was adapted to the work domain constraints <u>and</u> their own information processing limitations. This conjecture is described in more detail in the descriptive analysis.

Descriptive Analysis

Using the three steps identified in Table 2, a more detailed descriptive analyses was conducted to inductively determine to what extent subjects were following the adaptation criteria identified in the normative analysis. This analysis resulted in the identification of the strategies subjects were following. The first step analyzed how many valves were used. Figures 16, 17, and 18 plot the number of valves subjects used per trial for trials with demand pairs of Mode 1, Mode 2, and Mode 3 (see above), respectively. From the normative analysis, the minimum number of valves for Mode 1 was 3. Figure 16 illustrates that no subject ever used this strategy. Nevertheless, the data in Figure 16 reveal several interesting facts. First, subjects were remarkably consistent in terms of the number of valves that they used. There was no learning whatsoever. The number of valves that subjects started off using was the number that they kept on using for the 6 month duration of the experiment. Second, three subjects (AV, IS, and TL) consistently used only 4 valves. Given the normative analysis, it seems that these subjects adopted the criteria of minimizing monitoring demands and interactions in determining how many valves to use. The other three subjects (AS, ML, and WL) almost always used all 6 valves to configure the FWS. They did not seem to adapt their strategy to the unique possibilities associated with Mode 1 demands.

Figure 17 shows the data for the trials with Mode 2 demand pairs. The minimum number of valves for this Mode was 4. As with the previous graph, all subjects were very consistent in the number of valves that they decided to use. AV, IS, and TL generally used only 4 valves to perform the task, just as in Mode 1. In contrast, AS, ML, and WL used all 6 valves for Mode 2 trials, even though they could have accomplished the task with only 4. Thus, these subjects did not exploit the economies that were available in Mode 2 demands.

Figure 18 shows the results for the trials with Mode 3 demand pairs. The minimum number of valves for this Mode was 5. Once again, subjects were very consistent in their actions. AV, IS, and TL always used only 5 valves to perform the task. On the other hand, AS, ML, and WL almost always used 6 valves, one more than was necessary to complete the task. Collectively, these results show that AV, IS, and TL adapted their actions according to the demands imposed by the demand pairs, thereby minimizing their information processing requirements whenever the trials allowed it. In contrast, AS, ML, and WL showed no such sensitivity. Their actions were consistent, independent of the demand pairs. As a result, they unintentionally experienced greater information processing demands than the other subjects.

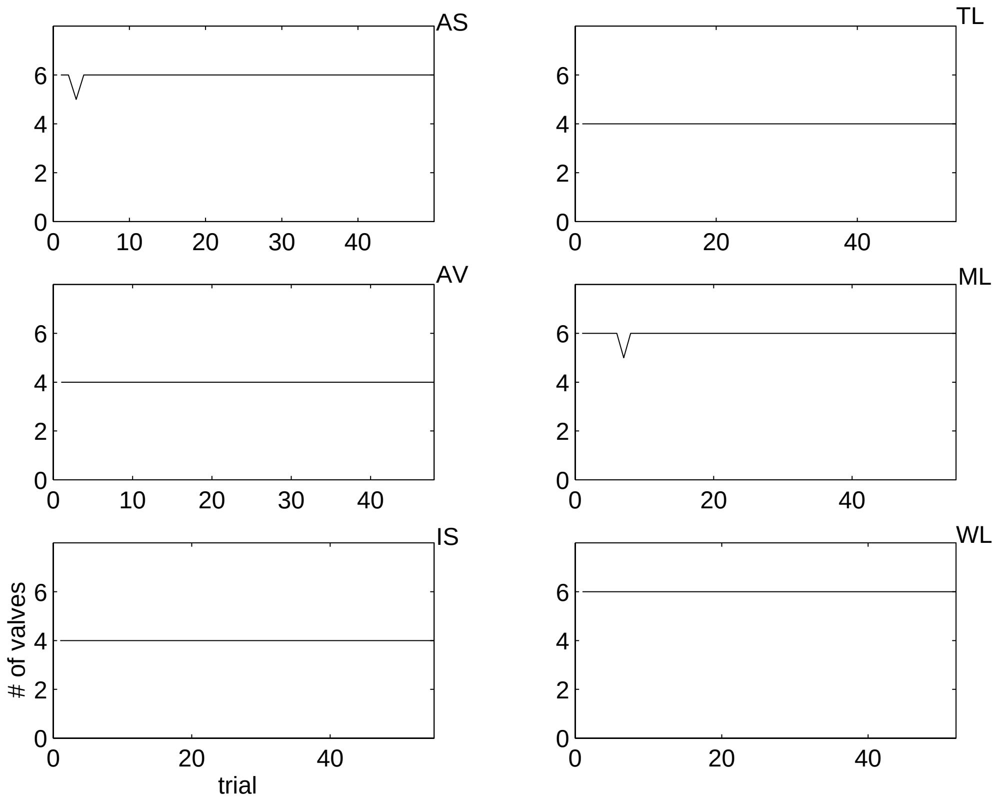

Figure 16: Number of valves used at trials with demand pairs of mode 1 (from Yu et al., 1997).

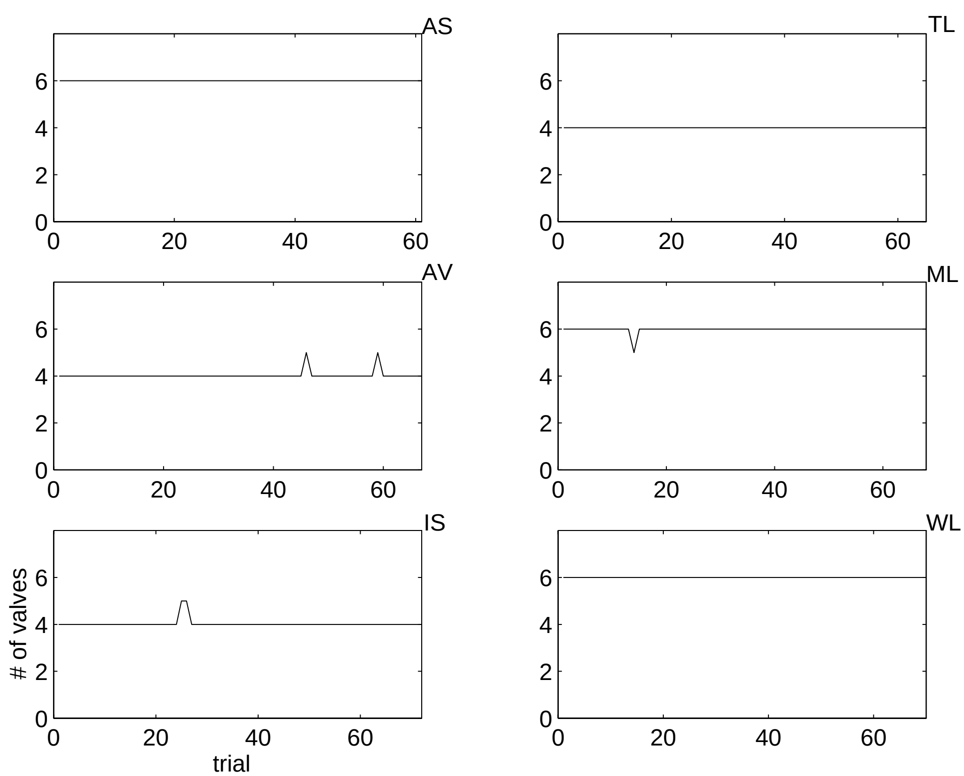

Figure 17: Number of valves used at trials with demand pairs of mode 2 (from Yu et al., 1997).

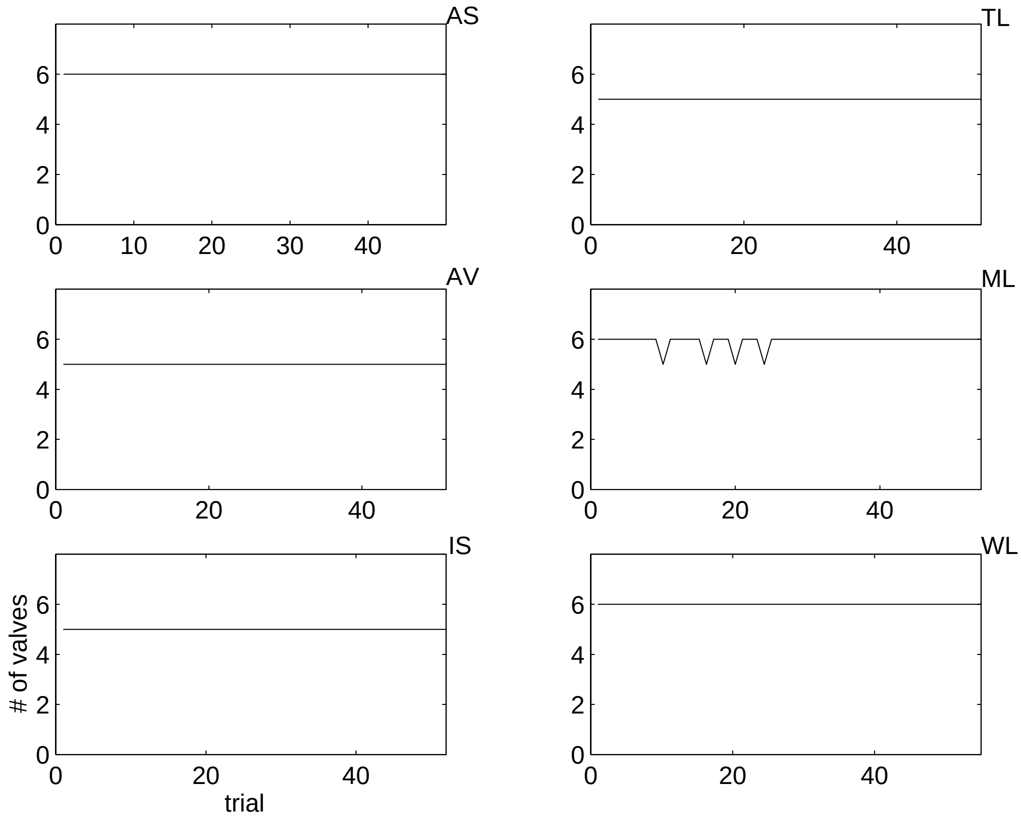

Figure 18: Number of valves used at trials with demand pairs of mode 3 (from Yu et al., 1997).

A second set of analyses was conducted to address the issue of which valves were used. This analysis was only pertinent to AV, IS, and TL because the other subjects always used all 6 valves. The results were very straightforward. When these subjects used only 4 valves, VA2 and VB1 were always the valves that were not used. When they used 5 valves, VB1 was always the one that was not used. These findings were in perfect agreement with the stimulus-response compatibility criterion identified earlier (see Table 2). Subjects always chose the direct mapping, thereby avoiding the non-intuitive reverse mapping.

A third set of analyses was conducted to tackle the question of what valve settings were used. As indicated by Figure 15, the data for AV, IS, and TL were well-captured by the normative strategy. Yu et al. tried to inductively identify algorithms that would capture the behavior of the other three subjects by replaying their trials and by relying on their verbal reports and control recipes (Christoffersen et al., 1994). The algorithm that was identified for ML is listed in Table 5, using the EIPs in Table 3.

READ	*D1*	
READ	*D2*	
ADD	*D1 + D2*	
DIVIDE	*(D1 + D2) / 2*	
ASSIGN	*VA = (D1 + D2) / 2*	
ASSIGN	*VB = (D1 + D2) / 2*	
COMPARE	*D2 to 10*	
if D2 $\leq$ *10*		
	ASSIGN	*VA1 = D1*
	ASSIGN	*VB1 = D1*
	ASSIGN	*VA2 = D2*
	ASSIGN	*VB2 = D2*
else		
	DIVIDE	*D1 / 2*
	ASSIGN	*VA1 = D1 / 2*
	ASSIGN	*VB1 = D1 / 2*
	DIVIDE	*D2 / 2*
	ASSIGN	*VA2 = D2 / 2*
	ASSIGN	*VB2 = D2 / 2*

Table 5:. The algorithm inductively developed to account for ML's strategy for performing the FWS configuration subtask (from Yu et al., 1997).

An error score was calculated and plotted against trial number to see how well this algorithm accounted for the subjects' behavior. The results for all 6 subjects are given in Figure 19.

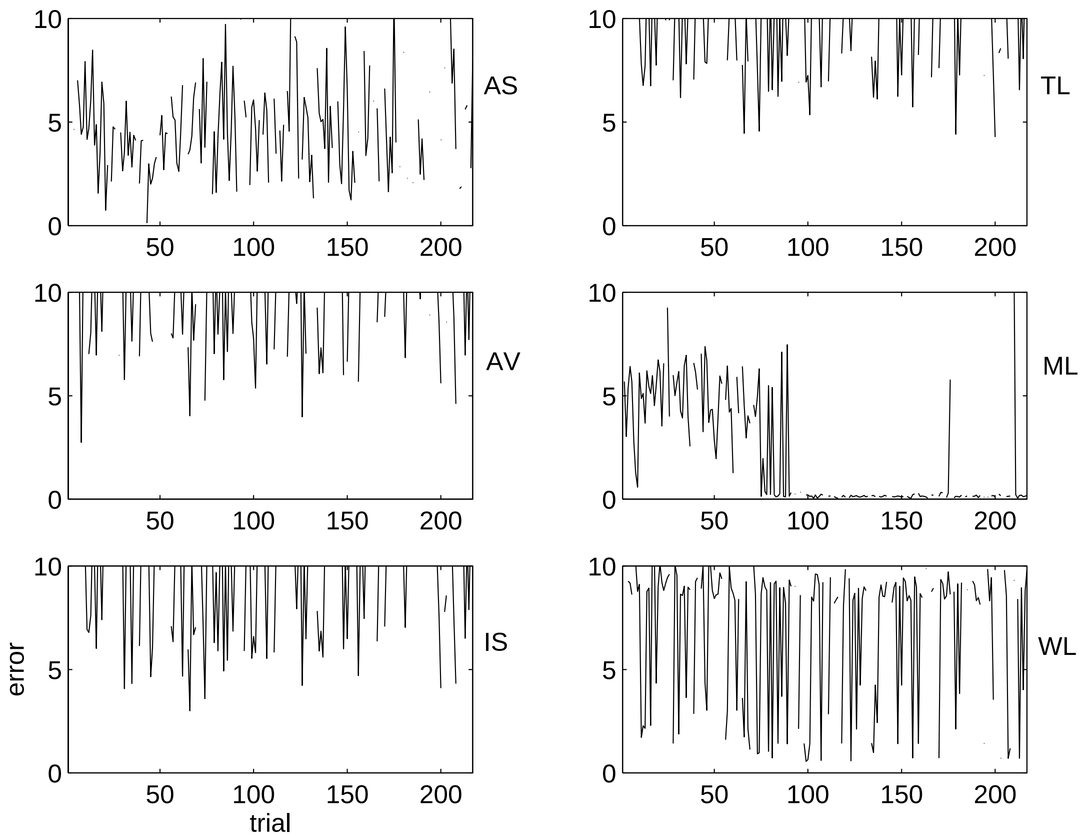

Figure 19: Error between actual valve settings and those predicted by Table 5 (from Yu et al., 1997).

As expected, the algorithm in Table 5 provided an excellent fit to the data of ML, beginning at around trial #100. The fit with the data for most of the other subjects was quite poor. The sole exception was WL, whose behavior on certain trials was moderately well captured by the algorithm in Table 5. This occasional fit occurs because WL's strategy is similar to ML's under certain conditions (see below).

The algorithm that was inductively identified for WL is listed in Table 6. Figure 20 shows how well this algorithm fit the subjects' data. The agreement with WL's behaviour is very good and quite consistent. ML's behaviour is well accounted for on some trials for the reason already mentioned. The fit to the behavior of all of the other subjects is clearly inadequate.

```
ASSIGN      VA = 10
ASSIGN      VB = 10
READ        D1
READ        D2
COMPARE     D1 to 1
COMPARE     D2 to 2
if D1 = 1
        ASSIGN      VA1 = 1
        DIVIDE      D2 / 2
        ASSIGN      VA2 = D2 / 2
        ASSIGN      VB2 = D2 / 2
if D2 = 1
        ASSIGN      VA2 = 1
        DIVIDE      D1 / 2
        ASSIGN      VA1 = D1 / 2
        ASSIGN      VB1 = D1 / 2
else (D1 & D2 > 1)
        DIVIDE      D1 / 2
        ASSIGN      VA1 = D1 / 2
        ASSIGN      VB1 = D1 / 2
        DIVIDE      D2 / 2
        ASSIGN      VA2 = D2 / 2
        ASSIGN      VB2 = D2 / 2
        ASSIGN      VA = 10
```

```
ASSIGN      VB = 10
READ        D1
READ        D2
COMPARE     D1 to 1
COMPARE     D2 to 2
if D1 = 1
        ASSIGN      VA1 = 1
        DIVIDE      D2 / 2
        ASSIGN      VA2 = D2 / 2
        ASSIGN      VB2 = D2 / 2
if D2 = 1
        ASSIGN      VA2 = 1
        DIVIDE      D1 / 2
        ASSIGN      VA1 = D1 / 2
        ASSIGN      VB1 = D1 / 2
else (D1 & D2 > 1)
        DIVIDE      D1 / 2
        ASSIGN      VA1 = D1 / 2
        ASSIGN      VB1 = D1 / 2
        DIVIDE      D2 / 2
        ASSIGN      VA2 = D2 / 2
        ASSIGN      VB2 = D2 / 2
```

Table 6: The algorithm inductively developed to account for WL's strategy for performing the FWS configuration subtask (from Yu et al., 1997).

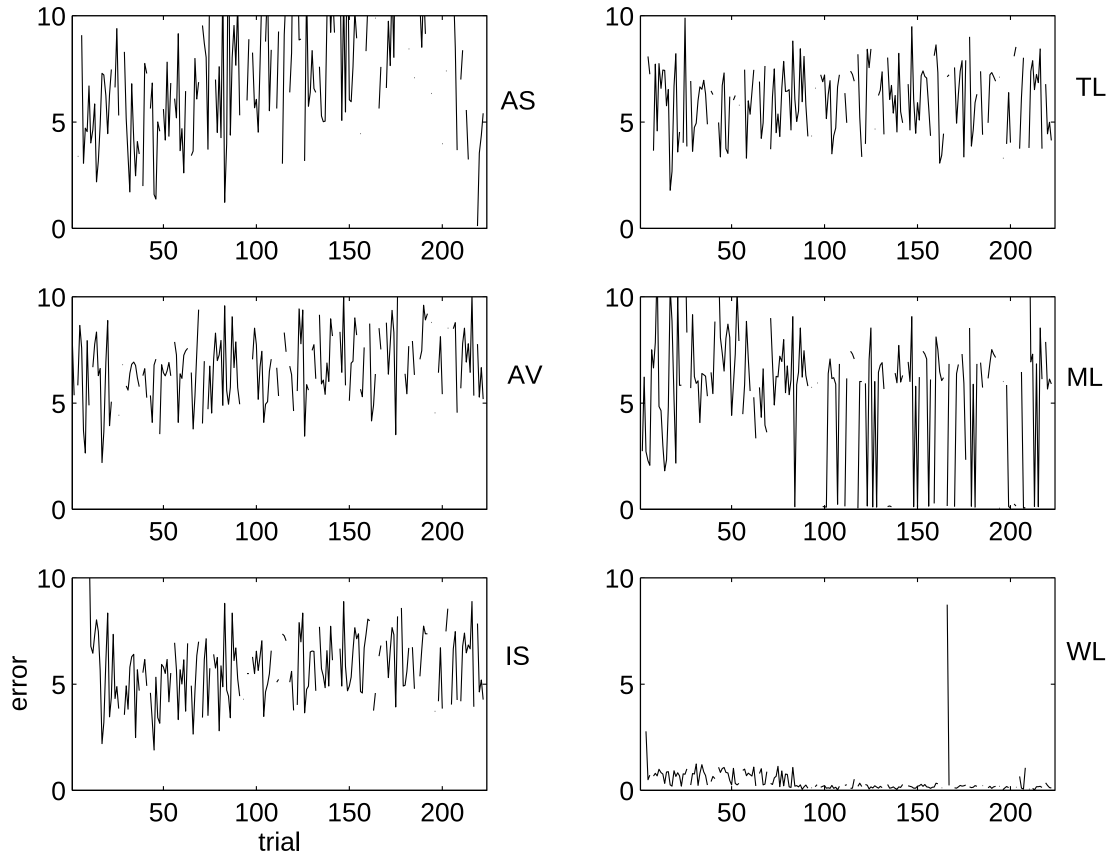

Figure 20: Error between actual valve settings and those predicted by Table 6 (from Yu et al., 1997).

Yu et al. also tried to inductively determine what strategy AS was using, but they were unable to do so. Figures 15, 19, and 20 clearly reveal that AS was not following any of the previously identified algorithms. Although it was possible that there was some unknown consistency in his behavior, as far as they were able to tell, his actions were driven by a trial and error strategy across trials.

The differences in the ways in which the subjects approached the issue of valve settings could be summarized by tabulating the number of EIPs associated with the algorithms listed in Tables 4 to 6. The results are listed in Table 7. Because the issue of settings is nested under the issues of how many valves to use and which valves to use, this summary actually represents the overall differences in strategies across subjects. The strategy used by AV, IS, and TL was more adaptive in the sense that it minimizes the number of EIPs required to do the task. The other subjects were doing more work than was required to perform the task.

EIP	AV, IS, & TL	ML	WL
ADD	0	1	0
ASSIGN	4 or 5	6	5 or 6
COMPARE	1	1	2
DIVIDE	0	1 or 3	1 or 2
READ	2	2	2
SUBTRACT	0 or 1	0	0
TOTAL	7 or 9	11 or 13	10 or 12

Table 7: A comparison of the EIP count for the FWS configuration subtask strategies for all subjects, except for AS (from Yu et al., 1997).

Although there were probably differences in the amount of effort associated with different EIPs (e.g., READ is probably easier than DIVIDE), Yu et al. did not try to weight the different types of EIPs to account for these differences (cf. Payne et al., 1993). If weights were added, it may very well be that the differences in information processing demands between the strategies in Table 7 are even greater than the totals would indicate.

From the analyses conducted using the EIP adaptation measure, it was determined that almost all of the subjects exhibited very consistent strategies for dealing with the FWS configuration subtask. A normative analysis revealed that it was possible to introduce psychological criteria to remove essentially all of the available degrees of freedom, thereby leading to the identification of a minimal effort solution to the task. By comparing this normative algorithm to the data exhibited by subjects throughout the experiment, Yu et al. found that three of the subjects showed signs of following this algorithm. AV, IS, and TL all adapted their behavior to get the task done in a way that was consistent with the psychological criteria that were identified. In short, these subjects were adapted, not only to the work domain constraints, but to their own information processing limitations as well. They tried to make the task as simple as possible.

Yu et al. also conducted a descriptive analysis to see if they could identify the strategies being used by the remaining three subjects. In one case, AS, they were not successful. There did not seem to be any consistent pattern in the way in which AS dealt with the FWS configuration subtask across trials. In the other two case, ML and WL, they had more success. Algorithms were developed that provided very good fits to the behavioral data. An examination of the information processing demands imposed by these algorithms shows, however, a poor degree of adaptation to minimal effort criteria. These subjects were performing the task in a way that was more difficult than it needed to be. They were not sensitive to the differences between trial Modes, and they used strategies that required a larger number of EIPs than was necessary. Thus, while they were able to perform the task successfully, they did not do so in a cognitively economic fashion.

If these results are compared to the findings from the performance analyses presented by Yu et al. (1997), an obvious pattern comes to the fore. The three subjects who exhibited signs of adaptation to information processing limitations were generally the most proficient performers. TL was always the best in the P group, AV was usually the best in the P+F group, and IS was often a close second. The three subjects who made the FWS configuration subtask more difficult than it needed to be, on the other hand, were the least proficient performers. WL and ML alternated being the worst in the P group, and AS was clearly the worst in the P+F group. Thus, there seems to be a correlation between performance and adaptation to information processing limitations.

Although there was no direct evidence to causally link these two variables, it seemed very likely that such a link exists. All subjects have information processing limitations. Furthermore, all subjects were given the same task to perform on the same system. The constraints imposed by these factors cannot be escaped. The only place where there was discretion was in how to deal with the task demands. Some subjects dealt with those demands in a way that showed that they were sensitive to making the task as easy as possible. This involved systematically adapting their behavior as a function of the context. Other subjects dealt with the task demands in a way that made the task more difficult than it needed to be. Their behavior was not as systematically adapted as a function of the context. Thus, it is plausible that the subjects who performed the task in a more economic fashion would outperform those who performed the task in a more challenging fashion.

***

The next part of this section discusses some proposed measures that may be useful for assessing adaptive performance to complement the measures that have already been developed from previous research.

4.3 Proposed Measures and Directions for Future Measures

In this subsection, we outline a number of proposed measures that may be useful for capturing adaptive performance across a number of the behavior-shaping constraints identified in the CWA framework. These measures are divided into three categories: dynamical systems and stability concepts, and distributed moments.

4.3.1 Dynamical Systems and Stability Measures

A set of measures based on the concepts of dynamical systems and stability concepts can be used to study adaptive performance in complex work situations. These types of measures have been used extensively in physics and biology to study the behavior of nonlinear systems (e.g., Abraham & Shaw, 1982, 1983, 1984, 1988; Ashby 1952, 1956; Haken, 1988). They are now becoming incorporated in psychology (e.g., Thelen & Smith, 1994; Kelso, 1995; Port & van Gelder, 1995) and sociotechnical systems analysis (e.g., Vicente, in press; Yu et al., 1998). Before discussing this class of measures in detail, the motivation for using a dynamical systems approach for assessing adaptive performance is presented by discussing the ideas of strategy basins and catastrophe landscapes.

Strategy basins. From the viewpoint of the CWA framework, strategies can be considered as processes that identify how human operators can perform particular tasks (Vicente, in press). Figure 21 illustrates the concept of strategies from a constraint-based approach. Strategies may be considered as categories or bins of cognitive procedures, rather than specific behavior instances (i.e., trajectories). Action sequences can be considered as switching between instantiated strategies. From a dynamical systems perspective, strategies may be viewed as basins of attraction within a dynamically constrained work environment. Next, we discuss two examples to describe this concept of strategy basins: air traffic control and horse gait analysis.

The first example stems from a study of strategies for air traffic controllers (ATCs). In this work environment, ATCs adopt different strategies to perform any one task, and these strategies change based on the current situation. The work of Sperandio (1978) provides evidence supporting this claim from 10 years of studying ATCs in the field. He found that under a given work situation, certain operating procedures are more economical in terms of cognitive demands than others. Also, as task demands increased (e.g., number of aircraft to be controlled and monitored increased), ATCs would loosen the performance criteria they used to perform the tasks. In doing so, they adopted different strategies that accomplished the same task goals in a less effortful manner. Figure 22 shows these qualitative findings graphically. As task demands increased, there were abrupt transition points where ATCs switched to qualitatively different procedures or methods to manage the changing demands.

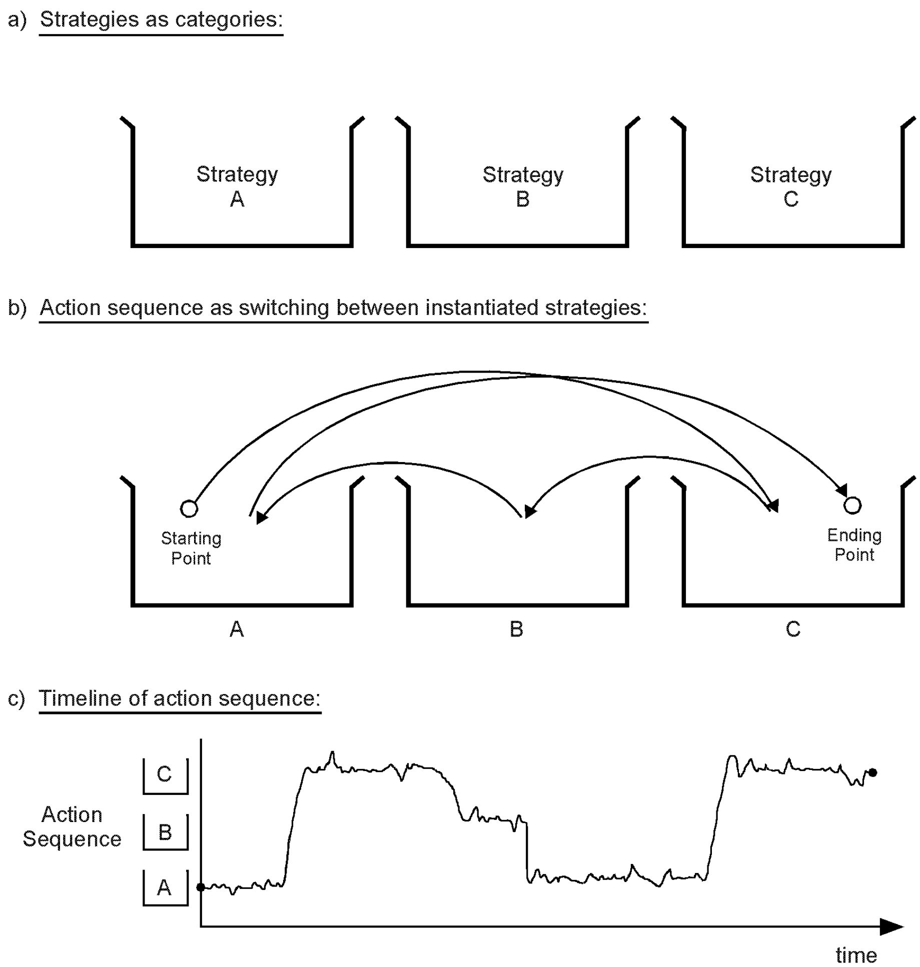

Figure 21: Strategies in complex work environments (from Vicente, in press).

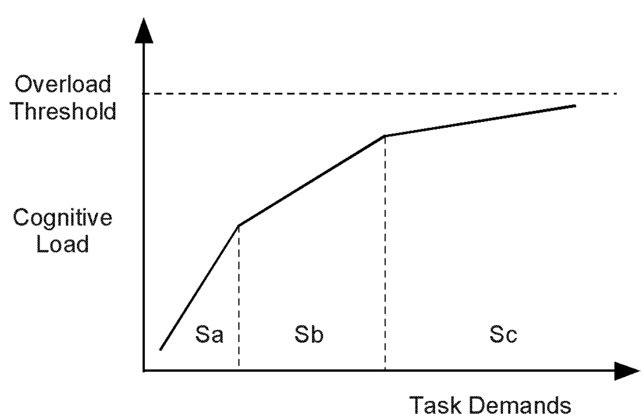

Figure 22: Adaptive regulation of cognitive load via the use of successively less demanding strategies (Sa, Sb, Sc) by air traffic controllers (from Sperandio, 1978).

In the first stage (Sa), where there are about 1 to 3 planes to control, the task demands are relatively low and the ATC has time to calculate the optimal flight path for each individual plane based on a number of factors (e.g., speed, course, altitude). When the task demands have increased to about 4 to 6 airplanes, there is an abrupt transition to the second stage (Sb). The ATC adopts a cognitively more economical strategy, where uniform speeds and stereotypical paths are adopted instead of optimal paths for each airplane. As the task demands further increase to more than 6 planes, a transition to another strategy (Sc) occurs. Under these situations, the ATC creates buffers that consist of streams of airplanes. When the ATC is ready, they would bring an airplane out off the buffer and towards the runway at a generally uniform speed and descent path. The primary concern for this strategy is safety, not efficiency.

As shown in Figure 23, another example of qualitative transitions in behavior comes from horse gait analysis by Hoyt & Taylor (1981). They measured the amount of oxygen consumed for the horse to move 1 meter as a function of running speed.

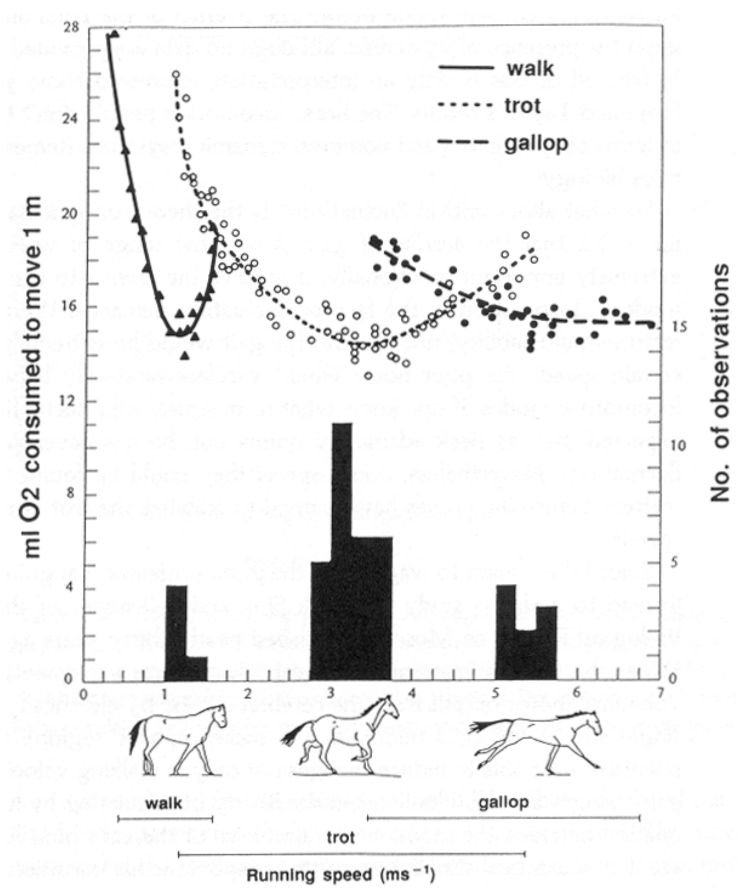

Figure 23: Oxygen consumption and preferred speed of walk, trot, and gallop of ponies (from Hoyt & Taylor, 1981).

The results showed that, at certain critical values of speed, the horse made a transition in the type of (or strategies for) locomotion from walking to trotting to galloping. These transitions corresponded to critical points of high oxygen consumption roughly defined by the intersection of the parabolic relations in Figure 23, where a transition to a different type of locomotion resulted in improved efficiency in oxygen consumption. Also, at the bottom of the graph are histogram relations indicating the frequencies with which the horse was observed to move at particular speeds. The three distinct peaks correspond to the lowest levels of oxygen consumption for the different types of locomotion, indicating an adaptation to local minima of work (i.e., oxygen consumption).

What are the similarities between these two examples from seemingly diverse application domains? How do they relate to dynamical systems and stability concepts? These examples provide evidence that when particular situations occur in the work environment (e.g., changing task demands), different strategies are adopted to manage the complexity of the environment in an economic manner (e.g., in terms of the mental or physical energy expended). Transitions to more efficient strategies occur when certain constraint boundaries are reached (e.g., human information processing limitations). In this sense, strategies may be viewed as basins of attraction, as previously discussed. The interaction between organism and environment constraints induce movement to and from different attractor basins as a constraint boundary is approached. In the case of ATCs, as the task demands increase (i.e., increased number of airplanes to monitor and control), transitions occur resulting in more efficient strategies of operation, in terms of cognitive load. In the case of gait analysis of horses, as the demands on the horse increases (i.e., increased speed of horse), transitions in the sequencing of actions occur resulting in the walk, trot, and eventually, gallop strategies of locomotion. These transitions also result in more efficient strategies of operation, in this case in terms of physical load.

Catastrophe landscapes. One aspect of dynamical systems and stability theory is the concept of catastrophe landscapes that is based on a stationary state model called catastrophe theory (Woodcock, 1988). Catastrophe theory deals with the sudden changes in behavior. The theory is useful in cases in which smoothly changing forces can lead to both gradual and abrupt changes in the same system under different circumstances. For example, this approach has been used in biology to study evolution as a process of adaptation to fitness landscapes (e.g., Schluter & Grant, 1984).

The landscapes may be considered as the dynamic set of constraint boundaries that shape behavior (e.g., from the constraint layers of CWA). Depending on the initial conditions and the direction of movement, resulting behavior can be qualitatively different with the same or changing constraint landscape. For example, a particular catastrophe landscape (i.e., cusp catastrophe) may be conceptually visualized in Figure 24 for changes in two work domain variables, X and Y. The third axis may be indicative of a higher order work domain variable Z that results from the interaction of the two work domain variables.

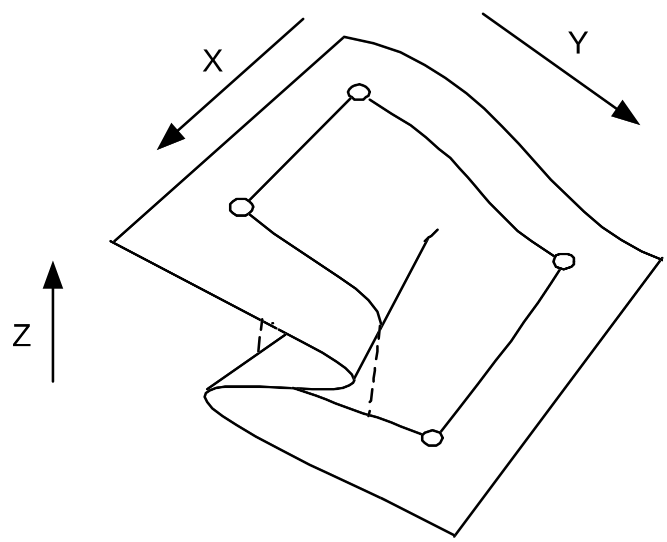

Figure 24: An example of a catastrophe landscape.

A number of situations may arise depending on the initial conditions and direction of movement of the work domain variables (Table 8). Situations 1 and 2 result in gradual changes in Z, while situations 3 and 4 result in abrupt (or 'catastrophic') changes in Z. The shape of the landscape greatly affects which situation will dominate. This is an important issue when the landscape becomes dynamic due to external factors (e.g., changing resource constraints that affect the shape of the landscape). Thus, it is important to be sensitive to these landscapes, representing a set of behavior-shaping constraints, to be able to adapt to changing circumstances. Given the need for adaptation to change and novelty described in the Introduction, the relevance of such landscapes to military operations should be clear.

Situation	Initial Conditions	Direction of Movement	Effect on Z
1	Any X, Low Y	Along X-axis	Gradual change throughout
2	Any Y, Low X	Along Y-axis	Gradual change throughout
3	Any X, High Y	Along X-axis	Abrupt change at transition points
4	Any Y, High X	Along Y-axis	Abrupt change at transition points

Table 8: The effect of situations (i.e., characteristics of X and Y) on Z from Figure 24.

As shown in Figure 25, a military example to illustrate how catastrophe landscapes may be conceptually applied stems from the work of Woodcock (1988).

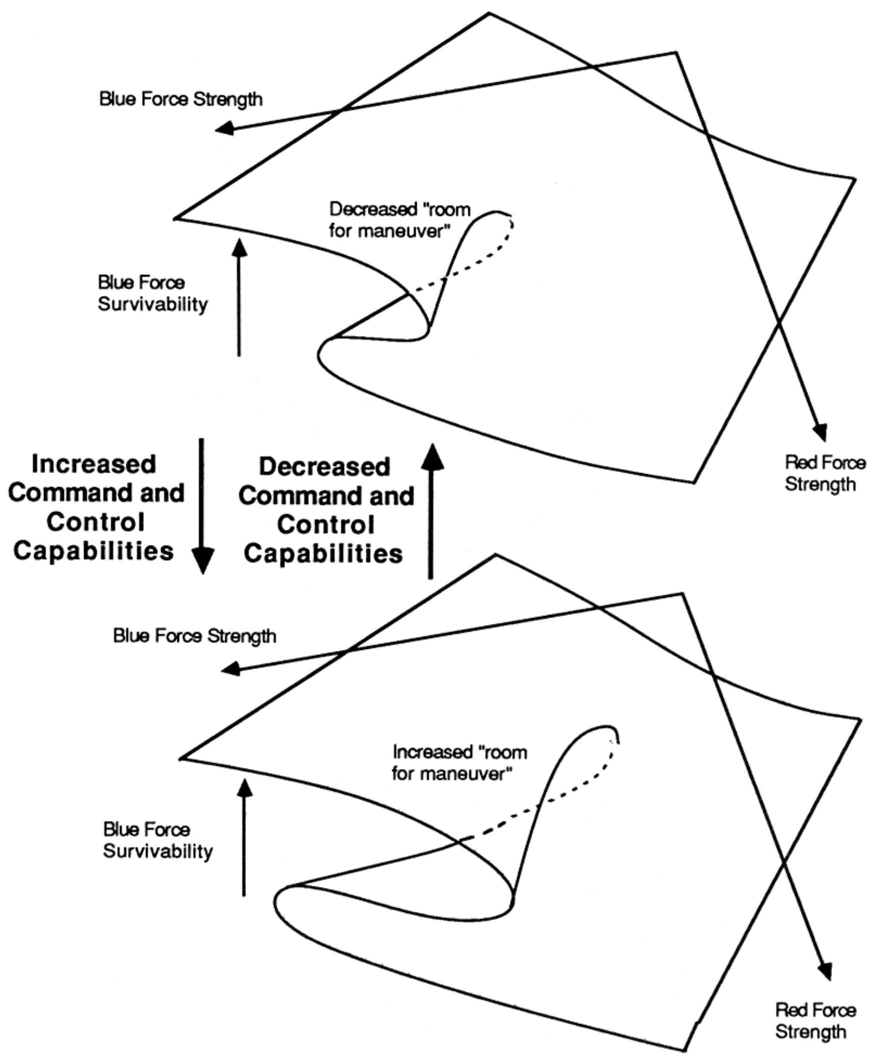

Figure 25: The impact of command and control on military force survivability (Woodcock, 1988).

In this hypothetical example, the inherent strength of opposing military forces (i.e., red forces) is considered to impose constraints modeled by these landscapes as a part of the work domain. The third dimension represents another work domain variable, survivability of the blue force, given particular levels of strength of the blue and red forces. The shape of landscape is affected by the command and control capabilities of the blue force. The combination of changing force strengths of blue and red forces, and command and control capabilities impact the survivability of the forces. By modeling these landscapes in terms of changing command and control capabilities, an understanding of the critical points may be developed to see how much room there is for maneuverability in the distribution of force strengths.

It would be useful to have a number of measures that would show when these transitions occur (i.e., between basins of attraction and catastrophes) and at what circumstances. The dynamical systems and stability concepts provide some potential measures that can assist the analyst in identifying the basins of attraction, catastrophe landscapes, and transition points to changing patterns of behavior as a result of constraints of the work environment (e.g., transition between strategy boundaries). The proposed approaches and measures outlined in this section include: hysteresis, bifurcation, relaxation time, and fluctuation. By identifying these transitions over time, one may be able to see how human operators acquire skill in the work environment in their ability to adapt to changing work situations.

Hysteresis is a phenonenon in nonlinear systems that can result in abrupt changes in behavior patterns when a control parameter is gradually modified. An example of this phenomenon is shown in Figure 26. As the control parameter is incrementally modified from lower values to higher values, at a particular point or transition, there is an abrupt change in behavior. A qualitatively similar phenomenon occurs when the control parameter is modified from higher values to lower values. The transition points (Ta, Tb) for the two situations may be different, and the loop that is formed is called the hysteresis loop.

This phenomenon can occur in situations where a catastrophe or enfolded landscape has formed as a result of the behavior-shaping constraints in the work environment (see Figure 24 and Figure 25). The hysteresis loop can be seen at a particular cross-section of this landscape. For example in Figure 24, if X is constant at a high value and Y is incrementally modified between lower and higher values, a hysteresis loop is generated. The characteristics of the hysteresis loop (i.e., location of transition points, distances between transition points and behavior patterns, and area of loop) depend on the shape of the landscape, which in turn depends on the constraints of the work environment.

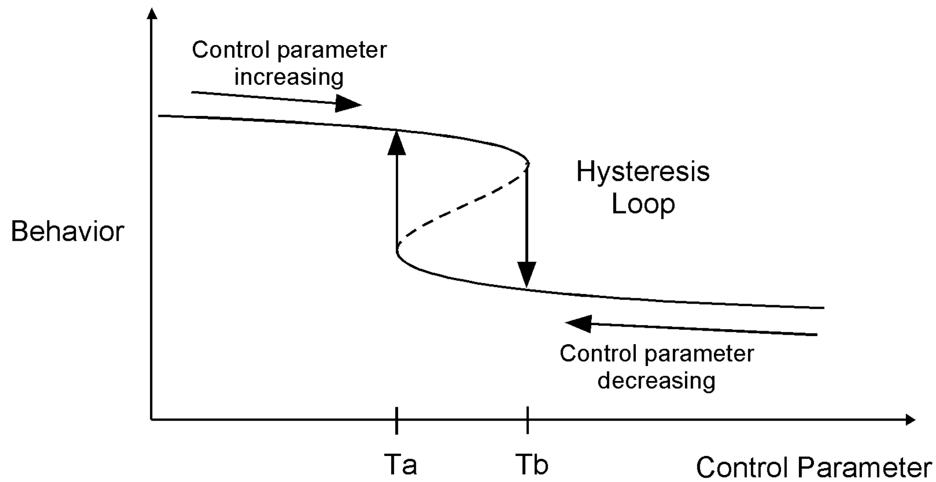

Figure 26: Example of the phenomenon of hysteresis.

From the perspective of the CWA framework, the control parameter may be considered as a particular property of the work environment that affects behavior. By systematically changing a particular constraint, we may be able to investigate the nature of behavioral changes due to this result. Characteristics of the hysteresis phenomenon may be useful for assessing particular aspects of adaptive performance in complex work environments. For example, the location and distances to transition points from the current state can be analyzed (e.g., in terms of actions and communications across actors) to determine to what extent individuals or teams adapt to the enfolding landscape. These measures may also be analyzed as the shape of the landscape changes (e.g., changes in the location, area, height, and width of the hysteresis loop), resulting from dynamic behavior-shaping constraints in the work environment.

Another phenomenon of a sudden change in the behavior of a system as a parameter or work domain variable is varied is called a bifurcation. Bifurcation diagrams are used for examining these changes in a dynamical system under these variations. As shown in Figure 27, some measure of system behavior is plotted as a function of a control parameter. As the control parameter increases, there is some critical point where behavior can change to a number of possible paths (both stable and unstable).

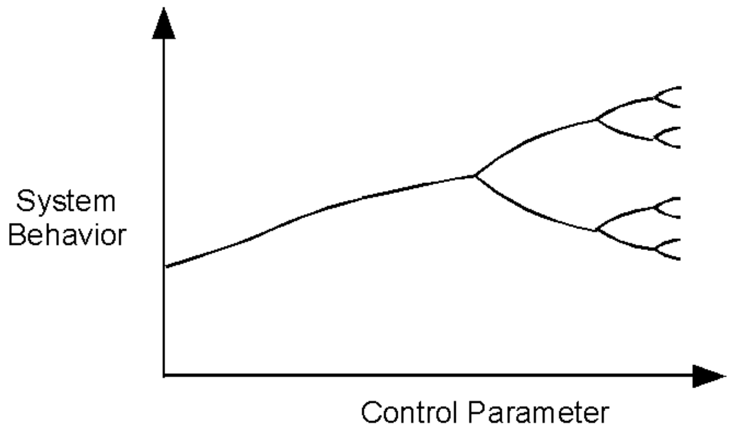

Figure 27: Example of a bifuraction diagram.

An example from physics to illustrate this phenomenon is the buckling of a metal beam (Figure 28). The independent variable or control parameter is the force exerted on the top of the beam and the dependent variable or behavior is the horizontal position of the point P on the beam. Initially, the beam is perfectly vertical. As force is exerted on the beam, it stays vertical. At a critical point C, any perturbation will cause the beam to bend in one of two directions (left or right). These correspond to the different paths on the bifurcation diagram.

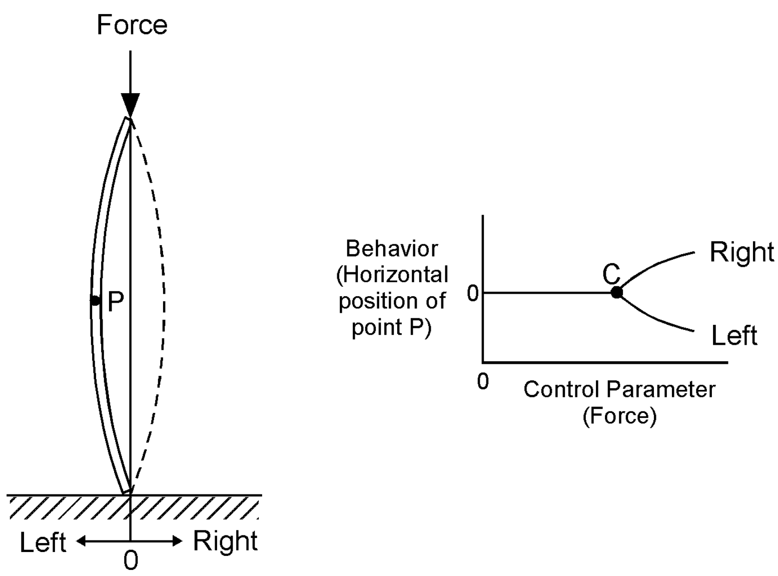

Figure 28: Example of bifurcation phenomenon of beam buckling.

The behavior of the beam is constrained by its properties and the forces exerted on it. In a similar way, behavior in complex work environments may be constrained by properties of particular work situations including factors that constrain behavior at each layer of the CWA framework. Adaptation in complex work environments may be seen as converging to one or more of these paths. This can assist in determining what the critical transition points are, as well as the bifurcation paths that may eventually lead to stable or unstable behavior.

The hysteresis and bifurcation measures can be potentially incorporated to all layers of constraint in the CWA framework. Control parameters can be constrained within the work domain, tasks, strategies, organizational structure, or competencies. Depending on the shape of the initial conditions and shape of the constraint surface, the transitions due to bifurcation or hysteresis may be investigated.

A couple of measures have been proposed by Kelso (1995) to investigate the onset of a transition by looking at the characteristics of instabilities: <u>relaxation time</u> and <u>fluctuation</u>. <u>Relaxation time</u> is the time required to recover back to the preset equilibrium position once a perturbation or disturbance has affected the system. For conditions closer to overall system instability or transition state, relaxation time increases as it becomes more difficult to return to a preset condition (see Figure 29). The phenomenon is also known as critical slowing down. <u>Fluctuation</u> refers to the variability in collective variables. As a transition point or instabilities are approached, large fluctuations in these variables may occur, anticipating a pattern change in behavior (see Figure 30). This phenomenon is also known as critical fluctuation. These measures may be useful for identifying and anticipating transitions in the work environment at various layers of the CWA framework. These transitions can then be assessed in terms of adaptation to the constraints of the work environment.

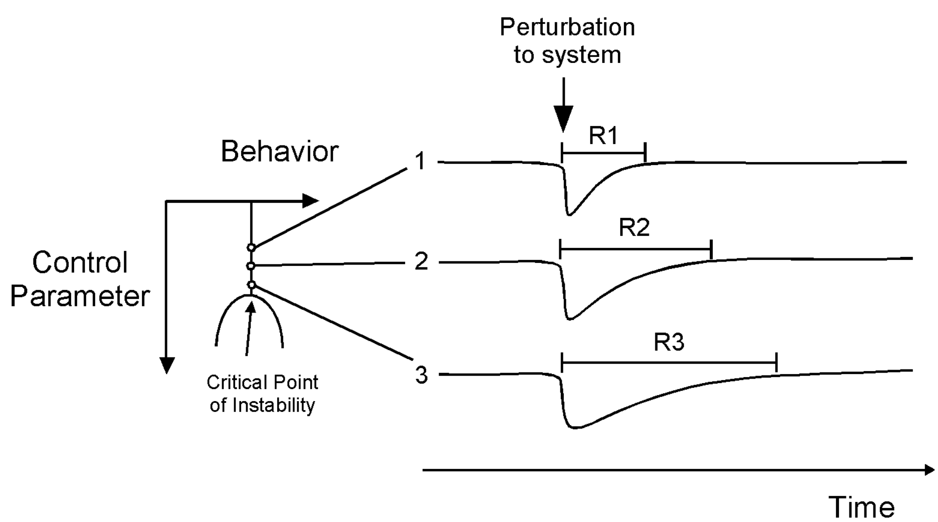

Figure 29: Changes in relaxation time (R1, R2, R3) as the system approaches a critical point of instability.

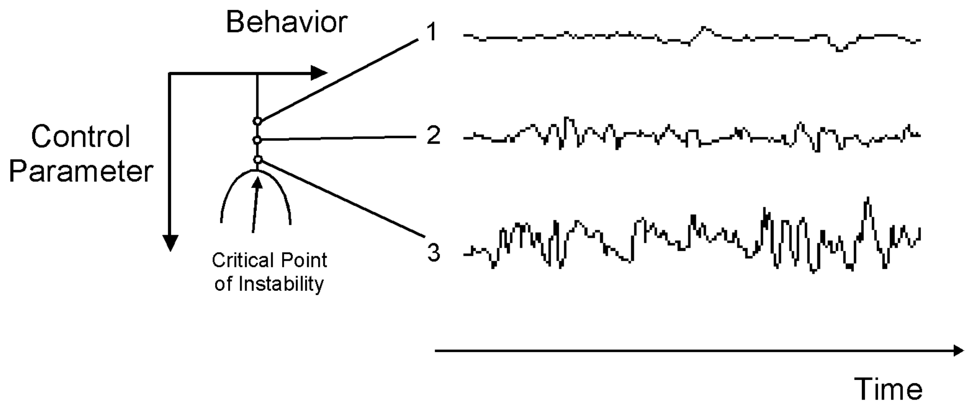

Figure 30: Changes in fluctuation of variables as the system approaches a critical point of instability.

4.3.2 Higher-Order Moments of Response Distributions

In Section 2, we used the simple example of automobile driving to illustrate the constraint-based approach. That example illustrates that response distributions (as opposed to single response trajectories) can provide insights into adaptive behavior by revealing sensitivity to constraint boundaries. Note that this insight is valid regardless of the particular type of constraint (i.e., work domain, control task, strategy, social coordination and cooperation, competencies) to which an operator may be sensitive. Accordingly, it would be useful to have a quantitative set of measures that would allow us to analyze rigorously the degree of adaptation in response distributions. Fortunately, Newell and Hancock (1984) provide a discussion of various summary statistics that can be used for this purpose. This subsection describes the quantitative measures identified in their article and shows the relevance of these measures to the assessment of adaptive behavior using a constraint-based approach.

Familiar Moments: Mean and Standard Deviation

A statistical distribution can be created by collecting a number of samples of some variable (in the case of cognitive engineering, usually a response variable). The familiar normal distribution is a prototypical example. There are various summary statistics that can be used to quantify the properties of a response distribution. The most familiar statistic is the first moment of the distribution, which is usually measured by calculating the mean of the sample. The mean provides a measure of central tendency, although other measures can serve the same purpose (e.g., mode, geometric mean, and harmonic mean). Another familiar statistic is based on the second moment of the distribution, which is usually measured by calculating the standard deviation of the sample. The standard deviation provides a measure of variability. Consequently, it provides insights that complement those that can be obtained by calculating the mean alone. After all, two distributions can have the same mean (i.e., the same central tendency) yet have different standard deviations (e.g., if one distribution is more variable, or "wider", than the other).

It is very rare for empirical studies in cognitive engineering to report information that goes beyond the mean or standard deviation of a sample. However, as Newell and Hancock (1984) pointed out, there are higher-order moments that provide additional insights into the characteristics of a response distribution. These "forgotten moments" turn out to be very valuable in measuring adaptation from a constraint-based perspective.

Forgotten Moments: Skewness and Kurtosis

One often-ignored descriptive statistic is based on the third moment of the distribution, which is usually measured by calculating the skewness of the sample. Skewness provides a measure of the degree of asymmetry about the mode of the distribution. If the distribution is symmetrical about the mean (as in the case of a normal distribution like that shown in Figure 32), then skewness takes on a value of zero. When the distribution "leans" towards the left, as in Figure 31, then it is said to be positively skewed. Conversely, when the distribution is "leans" in the opposite direction towards the left, then it is said to be negatively skewed. Skewness is distribution-independent in the sense that contrasts can be meaningfully made across different types of distributions (e.g., normal versus chi-squared). Thus, this measure is a robust descriptive statistic.

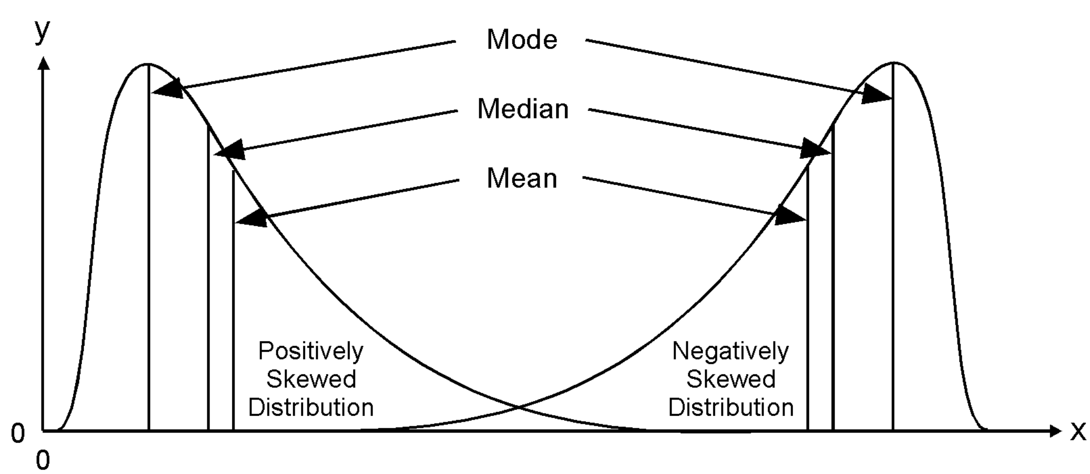

Figure 31: A sample of positively and negatively skewed distributions(adapted from Newell & Hancock 1984).

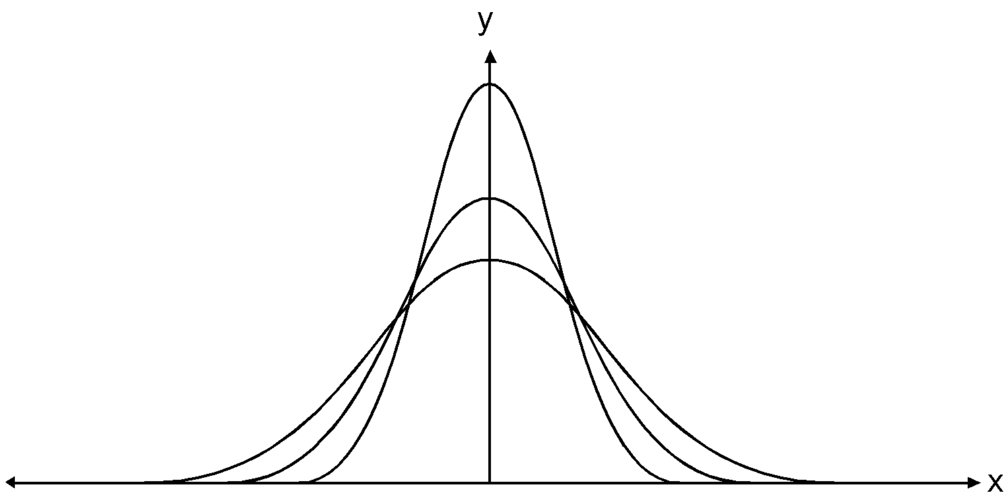

Figure 32: Response frequencies with systematic changes in kurtosis with skewness kept constant at zero (from Newell & Hancock 1984).

Another often-ignored summary statistic is based on the fourth moment of the distribution, which is usually measure by calculating the kurtosis of the sample. Kurtosis is a measure of the peakedness of the distribution. For purposes of comparison, a normal distribution has a kurtosis value of 3. High peakedness in a distribution leads to values greater than 3, whereas flatness in a distribution leads to values less than 3. Figure 32 shows changes in peakedness of a distribution with the central tendency held constant.

It should be obvious from these descriptions that skewness and kurtosis offer complementary information, the former measuring the extent to which a distribution "leans" in a positive or negative direction and the latter measuring the extent to which a distribution is "tall" or "flat". It should also be clear that skewness and kurtosis provide information that cannot be obtained from knowledge of the mean and standard deviation alone. As we will see next, this value added turns out to be particularly useful for measuring adaptation from a constraint-based perspective.

Relevance to Adaptation Measurement

Descriptive statistics based on higher-order moments of a distribution can provide quantitative measures of adaptation. The rationale behind this claim can be made clear by considering adaptation to two different types of constraints, a boundary constraint and a point constraint.

Adaptation to a boundary constraint. A boundary constraint represents a case where it is important for operators not to exceed a particular value. For example, when flying in international air space, it may be important for pilots to fly as close possible to the national air space of a particular country without actually transgressing into that air space (e.g., to gather intelligence information). In these cases, skewness provides a quantitative measure of adaptation to a boundary constraint. Figure 31 shows that positive skewness (i.e., leaning to the left towards smaller scale values) is indicative of greater sensitivity to the constraint on the low end of the dimension in question. The more positive the skewness, the closer the operators are to working at the very edge of that constraint (assuming that the boundary is not crossed). In contrast, smaller value of skewness would indicate a lower level of sensitivity to a boundary constraint on the low end of the dimension. A converse situation can be imagined where it is important to work closely to, but not exceed, a boundary constraint located at a high scale value. In this case, negative skewness (i.e., leaning to the right towards larger scale values) would be indicative of sensitivity to the boundary constraint. The lower the value, the closer the operators are to working at the very edge of the upper boundary on acceptable performance (assuming that the boundary is not crossed).

Adaptation to a point constraint. A point constraint represents a case where it is important for operators to achieve a particular point value. For example, when dropping off medical supplies from the air, it may be important for the supplies to hit (or be near to) a particular location on the ground. In these cases, kurtosis provides a quantitative measure of adaptation to a point constraint. Figure 32 shows that higher values of kurtosis (i.e., greater peakedness) show greater sensitivity to the constraint in question. The taller and narrower the distribution, the closer operators are to satisfying the constraint. In contrast, the shorter and wider the distribution, the farther operators are from the mark. In this case, we could say that they are not as sensitive (or less adapted) to the constraint in question.

Limitations

The primary limitation of these measures is that they require a rather large sample size to obtain reliable estimates. The higher the moment of the distribution, the greater the number of observations that is required to obtain a stable estimate. It will not always be possible to satisfy this requirement in the domain of military operations, although simulation data may provide a way around this limitation (Pew & Mavor, 1998).

5 Applying the Analysis Framework and Measurement Tools to Intentional Systems

In the previous sections of this report, we have outlined a theoretical approach and analysis framework that provide a general strategy for developing measures for assessing adaptive performance of individuals and teams in complex work situations. This approach and framework are based on the notion of behavior-shaping constraints associated with the work environment. It is important to note that an analysis of the different layers of constraint (i.e., CWA) needs to be performed for particular problems to select the appropriate dependent variables and measures for each case. The purpose of this section is to sketch some examples of how this might be done for military application domains.

To achieve this aim, we have first chosen the domain of ambulance dispatch management and then discuss the relevance of this example to military applications. There are several reasons for initially analyzing the ambulance dispatch domain. First, military application domains are very complex in nature and thus lead to unwieldy examples. The ambulance dispatch problem we have chosen is simpler, and thus, easier to explain as well as understand. Second, we do not have sufficient experience in terms of a cognitive work analysis for military work environments to use a detailed example based on our previous research. We have, therefore, borrowed from research conducted by other researchers for the domain of ambulance dispatch management (Wong, Sallis, & O'Hare, 1998). Third, military application domains and ambulance dispatch are both examples of intentional systems, work domains that are governed more by social constraints than by the physical constraints that govern causal systems (e.g., process control). As a result, the results from the analysis of ambulance dispatch should generalize to military work domains because the problems are of a similar qualitative nature. At the end of this section, we discuss a speculative example with a preliminary analysis of a particular military work domain for command and control to illustrate some of the similarities. Fourth, the applicability of CWA to intentional systems like military command and control was cast into doubt by Wong et al. (1998), who stated that "the approach appears not suited for use in human-activity-based intentional systems domains… these cause-and-effect relationships between the anomaly simply cannot be traced via the structural invariants identified through the work domain analysis". Thus, the ambulance dispatch management domain is a challenging test of our ideas. If we can show that CWA can be applied to the very same problem that Wong et al. were concerned with, then we will have demonstrated convincingly that their criticisms of CWA were not justified.

The remainder of this section is organized as follows. First, we discuss the general characteristics of intentional systems and how they can differ from causal systems. Second, we discuss the approach for structuring intentional systems, with a WDA for emergency ambulance dispatch management (EADM), an application domain with similar characteristics as some military work environments. Third, we present a hypothetical work situation for EADM to demonstrate how specific parts of the work domain structure may be used for problem solving in a particular scenario. Fourth, we discuss some measures of adaptive performance relevant to the EADM environment, using the results from the WDA. Finally, the relevance and transfer of the EADM example to military work environments is discussed in terms of a preliminary analysis of a focused military work domain for command and control.

5.1 Characterizing Environments for Work Domain Analysis

Before conducting a WDA for extracting dependent variables and measures of adaptive performance, it is useful to first determine the characteristics of the work environment. This will provide insight into how to represent the work domain by determining which constraints have a strong influence on behavior and how they can be modeled using the CWA framework.

Rasmussen et al. (1994) outline a taxonomy for characterizing work environments based on the relative degree of intentional and causal constraint. Figure 33 outlines a modified form of this taxonomy with examples of where particular work environments roughly fall on the continuum. There are two coupled axes in Figure 33. The first is the degree to which the environment is structured by actors' intentions or rules and practices (i.e., intentional constraints). The second axis is the degree to which the environment is structured by the physical laws of nature (i.e., causal constraints). Work environments with a higher degree of causal constraint generally have a lower degree of intentional constraint, and are referred to as causal systems (e.g., power generation, petrochemical, aviation, and flexible manufacturing systems). Conversely, work environments with a higher degree of intentional constraint generally have a lower degree of causal constraint, and are referred to as intentional systems (e.g., library information retrieval, finance, emergency dispatch management, and military command and control). It is important to note that both causal and intentional constraints exist in most sociotechnical systems, but to varying degrees.

It is important to consider the work environment in terms of these dimensions because they provide insight into how the WDA should proceed. For example, causal systems will predominantly be represented in terms of constraints associated with the laws of nature. Intentional systems will predominantly be represented in terms of constraints associated with actors' intentions, values, and rules and practices. This insight is useful to guide the process of WDA for intentional systems.

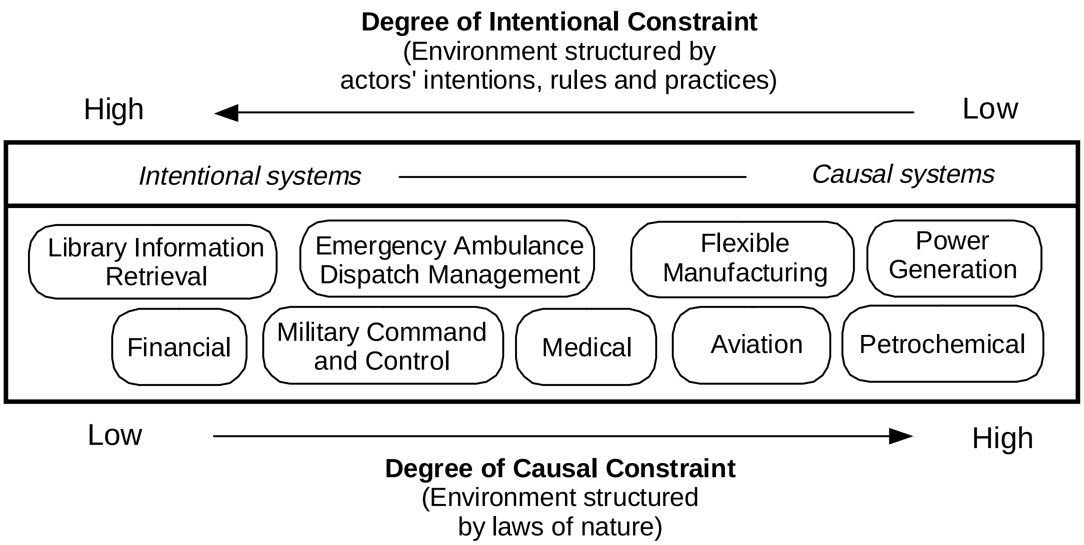

Figure 33: Characteristics of work environments based on the degree of intentional and causal constraint (adapted from Rasmussen et al., 1994).

5.2 Work Domain Analysis for Emergency Ambulance Dispatch Management

As mentioned earlier, we chose EADM as an application domain to illustrate how CWA can be applied to an intentional system. A description of this work environment is presented in Wong et al. (1998). The dispatch center has a team of operators that manage a network of distributed resources (e.g., ambulances) to handle emergencies within a particular geographic region. This system is intentional because the work environment is predominantly constrained by values, priorities, and rules and practices (e.g., management of emergencies according to an acceptable level of risk). The system boundaries subsume both the response resources and emergencies, embedded within the environment. There are also causal constraints that are important to model (e.g., balance and function of emergency response resources). However, for this application domain, the intentional constraints have a more significant role in shaping human behavior.

Figure 34 provides the general structure of our WDA for EADM using Rasmussen's (1985) abstraction hierarchy. The structure follows a similar one introduced by Rasmussen, Pedersen, & Grønberg (1987) for emergency management in non-nuclear industries, and by Moray, Sanderson, & Vicente (1992) for emergency management in the nuclear industry. The work environment was divided into two parts, the domain of potential risk associated with identifying and assessing emergencies and the domain of mitigating resources associated with response planning and execution. Dividing the work environment into these two "object worlds" (Bucciarelli, 1994) provides a way of distinguishing the intentional (i.e., priority functions) and causal constraints (e.g., resource balance functions). The work environment is also divided into various levels of abstraction, with the lower levels providing the structural means for achieving the higher level purposes (cf. Vicente, in press). The dispatcher will need information about both aspects of the work environment and the various levels of abstraction to dynamically make decisions on the allocation of resources and prioritization of emergencies.

Level of Abstraction	**Domain of Potential Risk**	**Domain of Mitigation Resources**
Functional Purpose	Goal: Maximize emergency health care response given the level of risk that is acceptable Constraints: The resources and abilities associated with the emergency response system	
Abstract Function	Priority criteria and given level of acceptable risk (e.g., survival, damages, public opinion, costs, rules and regulations)	Time, resource balance within region, probability of successful treatment
Generalized Function	Urgency of injury	Transportation, quality of care
Physical Function	Functional impairment, consequences of injury	Functional capability of resources (e.g., qualifications of medical personnel, speed of ambulance)
Physical Form	Location, # of people injured, type of injury	Locations, inventories, availability of various response mitigation resources (e.g., helicopters, ambulances, stations, hospitals, command posts, triage facilities, paramedic officers), roads, weather, terrain, traffic

Figure 34: Abstraction hierarchy for emergency ambulance dispatch management.

We will begin by describing the domain of mitigation resources, which identifies the capabilities and limitations of the emergency response object world. The physical form level at the bottom of the abstraction hierarchy identifies the appearance, condition, location, and spatial relationships of the response resources in the environment. The physical function level identifies the functional capabilities and limitations of these resources to handle emergencies. The generalized function level identifies the overall abilities for transportation and quality of care based on the functional capabilities and limitations of the resources. The abstract function level identifies the constraints associated with prioritizing resources including the probability of successful treatment, arriving at the emergency location within a particular time, and balancing resources across the geographic region. The functional purpose level identifies the purposes of the EADM work domain, to maximize emergency health care response given an acceptable level of risk and the constraints associated with the response resources.

An analogous representation may be formed for the domain of potential risk, which structures the risks and value judgements associated with emergencies in the region. The physical form level identifies the emergencies in terms of the location, appearance, number, and condition of the injured people in the environment. The physical function level identifies the functional impairment and consequences of the injuries. The generalized function level identifies the level of urgency associated with the functional impairment and consequences. The abstract function level identifies the criteria for prioritizing the emergencies (e.g., risks associated with survival and damage). The functional purpose level is the same as for the domain of mitigation resources.

5.4 Evaluating Situations with the Work Domain Representation

The value of the representation just described is perhaps best illustrated by an example showing how the work domain structure may be useful for evaluating and managing particular work situations. Using the results of the WDA from the previous subsection, we have instantiated the full work domain model for a specific albeit hypothetical work situation. The resulting situation model identifies the subset of the WDA that is relevant to, and instantiated by, the work situation in question. The hypothetical situation we chose has two available ambulances (one with a paramedic and another without) and two emergencies (one person having a heart attack and another person with a broken leg). How should the dispatcher allocate the ambulances to handle the emergencies? To assist in answering this question, the situation was mapped onto the work domain model, as shown in Figure 35.

At the physical form level, the characteristics of the situation are shown together for both object worlds on a common topographic map. The middle three levels are partitioned (i.e., by the dashed line) into the domains of potential risk and of mitigation resources. The top level is the same for the two object worlds, because there is a shared purpose of providing care for injured people based on acceptable risks and resource constraints.

In the domain of potential risk, the emergencies are identified at the physical form level (e.g., location and characteristics of the heart attack and broken leg victims). The physical function level maps the results of a functional assessment of the emergencies (e.g., chest pain and inability to walk). At the generalized function level, the urgency of these impairments are noted (e.g., high urgency for chest pain and low urgency for inability to walk). The abstract function level identifies the assessment of these emergencies based on the priority criteria of likelihood of survival and damage to the injured. The functional purpose level identifies the acceptable risks and value judgements associated with handling the emergencies, while taking into account the response resource constraints.

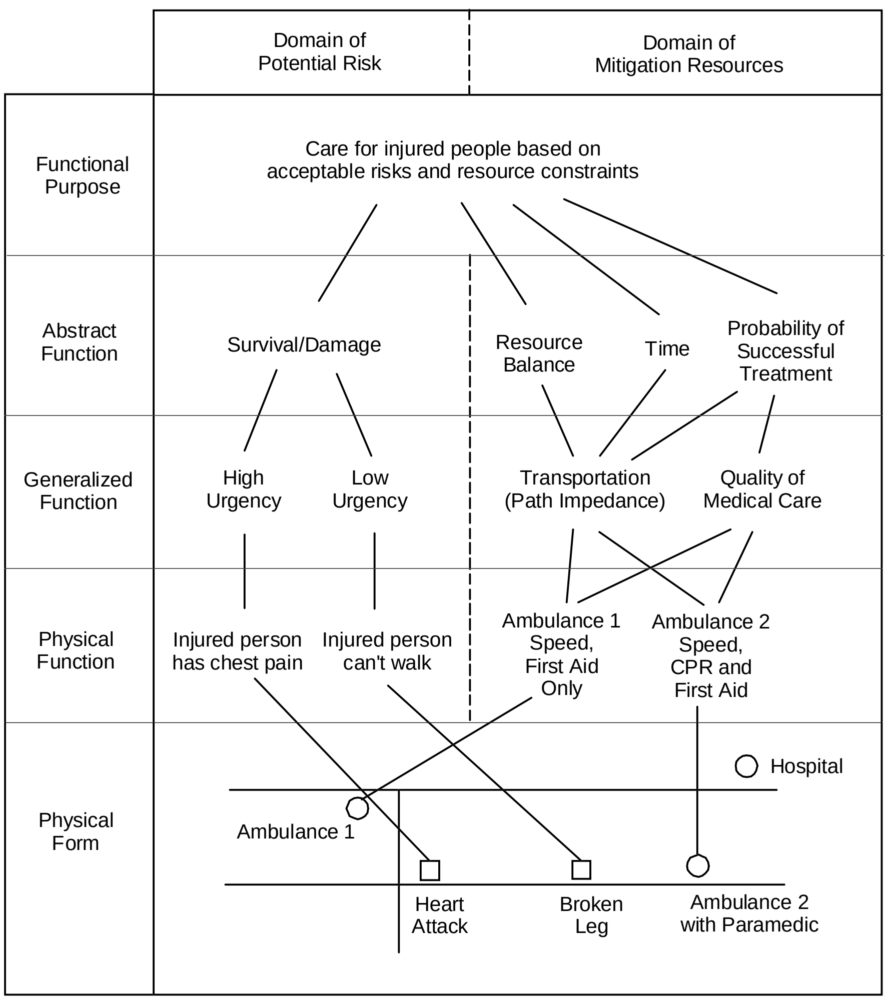

Figure 35: A hypothetical situation mapped onto the work domain model for emergency ambulance dispatch management.

The domain of mitigation resources illustrates the constraints associated with the control of risk within the environment. At the physical form level, the response resources are identified (i.e., ambulance teams at locations with particular characteristics). The physical function level maps the functional capabilities of the two ambulances and personnel (e.g., ability for personnel to conduct first aid and cardiopulmonary resuscitation (CPR), and capabilities of ambulances in terms of speed). The generalized function level maps the general transportation and quality of care capabilities of the two ambulance teams. For example, ambulance 2 can provide a better standard of care to the heart attack victim than ambulance 1, because the personnel on board can provide CPR. However, ambulance 1 is closer of the heart attack victim, and thus can arrive there more quickly. This situation sets up an interesting trade-off that must be resolved by referring to higher-level priority criteria. The abstract function level maps the time, resource balance, and probability of successful treatment in reference to the two ambulance teams. It is at this level that the aforementioned trade-off can be resolved. The probability of successful treatment will determine whether it is best to send the less qualified ambulance team to the heart attack victim (because the other team would arrive too late to care for the victim) or the more qualified team instead (if that team can get there in time to provide the required medical care). The functional purpose level is the same as is mentioned for the domain of potential risk.

An ambulance dispatcher may use the structure of the work domain model to assess emergency situations and allocate resources appropriately. In the situation described in this section, ambulance 1 is closer to the heart attack victim. However, the medical capabilities of ambulance 1 results in a lower quality of providing care for the heart attack victim, because they do not have the capabilities of providing CPR. The dispatcher is able to reason through these conflicting factors (i.e., time to reach the heart attack victim and quality of medical treatment) by integrating them into a higher-order functional concept (i.e., probability of successful treatment). This high-level, functional assessment allows them to make a judgement based on the priority criteria of survival and damage to the injured people.

5.5 Measures of Adaptive Performance

The work domain analysis for EADM can provide insight into developing measures of adaptive performance for both individuals and teams. This section outlines two measures, described in detail in the previous section, that may be useful in the context of EADM: AH variability and higher-order moments.

The AH variability measure discussed in Section 4.2.1 may be extended to EADM. Each level in the abstraction hierarchy defines a set of work domain variables that may be analyzed in terms of variability in state trajectories across various trials and scenarios. The types of variables are different for EADM compared with DURESS II because they have different work environment characteristics (i.e., EADM is primarily intentional and DURESS II is primarily causal). However, this measure may be applied to both work environments. The variables at different levels of abstraction discussed in Section 5.2 can be used to structure work domain state information similar to those shown in the scenario in Section 5.3. As before, each level of abstraction forms a different frame of reference that can be used for the measurement of adaptation. The states of each measured variable define trajectories in the work environment. As individuals or teams monitor and control the EADM work environment, adaptive performance may be assessed and compared in terms of the changes in variability at each level of abstraction.

Another set of measures that may be useful for EADM are related to the higher-order moments of response distributions discussed in Section 4.3.3. These measures can analyze response distributions (not trajectories) for adaptive behavior in sensitivity to the constraint boundaries defined from the work domain analysis. For example, the frequency of dispatcher communications and coordinated actions may change abruptly as a constraint boundary is approached as a result of an imbalance in the distribution of resources or a safety limit. The use of higher-order moments to quantify the sensitivity to constraint boundaries provides a measure to compare adaptive behavior across trials and scenarios.

In order to develop a comprehensive set of measures of adaptive performance for EADM, the other layers of behavior-shaping constraints need to be analyzed (i.e., other layers of CWA). The two measures outlined here are speculative and require evaluation with a detailed experimental investigation.

5.6 Military Applications

The command and control system used to manage military actions within a theater of operation represents another example of an intentional system. Because our analysis and measurement framework is quite general, it can be applied equally as well as to this and other domains of work. To illustrate the point, we provide a preliminary analysis for military command and control.

In general the command and control system incorporates the management tasks of planning, coordinating, decision making, and controlling both forces (assets) and operations (objectives and plans). For large-scale, complex operations, command and control as an intentional system can be describes at many different levels of granularity. For instance, a commander may be making management decisions about an entire air campaign that involves hundreds of airplanes and thousands of sorties spread over several days. At a different time the focus of a decision may be limited to the management of a single mission involving a hand-full of assets. Many decisions involve crisis management similar to the Emergency Ambulance Dispatch Management system.

We will briefly sketch the sources for intentional and causal constraint for command and control at a military air operations center. Then we will specialize the analysis for an event associated with an air interdiction mission by mapping the situation to the framework.

Air Interdiction is defined by Joint Pub 1-02: "Air operations conducted to destroy, neutralize, or delay the enemy's military potential before it can be brought to bear effectively against friendly forces…." Imagine that a strike package has been assembled to perform an air interdiction mission as part of the daily operations in a military campaign. Based on available information, it is believed that the enemy has active surface-to-air missile sites in the targeted area. Because these pose a high threat to the strike package, a support element (Suppression of Enemy Air Defenses) has also been included as part of the interdiction mission. SEAD support is provided by 4 F-16CJ, Block 50 airplanes. Twenty-five minutes before planned time-over target (TOT), however, the F-16 flight lead calls an air abort and returns to base. The Air Component Command in the Air Operations Center, who is managing all air campaign operations, is apprised of the situation. The commander must decide how to manage the new situation. One option may be to find out if other SEAD aircraft are available to handle this mission. Another option is to look for other types of assets. It may also be possible to adjust the TOT time. Other actions are also possible (e.g. request new intelligence to determine activity state of the SAMs). Further, mission abort is always an option. What action should the command take? What constraints shape the response?

As we have indicated earlier, our general analysis strategy is to partition the work domain along the intentional-causal constraint dimension and link the representation to the Abstraction Hierarchy. For military command and control example, the intentional constraint component may be called the Domain of Military Encounter. Engaging in military activity is an intentional act. Many factors shape the form of military action, including the even use military force.

Causal constraint aspects of the work environment might be called the Domain of Friendly Assets. There are limits on military resources that can be applied to the problem. And there are limits on how they can be used. These also factor into the equation of shaping an action plan and its execution.

Figure 36 summarizes an analysis of the military command problem in terms of the intentional domain and the causal domain as they are reflected at each level of the Abstraction Hierarchy. The command and control work domain is established and activated in order to achieve a military course of action that is deemed to be in the best interest of the country by national command authority. The goals, guidance, policies, and practices used to accomplish the objectives of a course of action are some of the intentional factors that compress the Domain of Military Encounter. In addition, priority measures may be imposed from such things as public opinion, congressional concerns, and the like. For example, injury to friendly forces and/or civilian casualties may serve as a cause to terminate military action before the objective is met. The command must balance these intentional issues. Similarly, time and resource issues over a campaign and within a mission must be balanced.

At the Generalized Function level, intentional constraints derive from concerns about assessing the military situation, acquiring and disseminating information, and coordinating diverse concerns among interrelated action organizations. The availability and relative location of the right people, equipment, and weapons over the duration of the campaign imposes causal constraints on action.

When considered at the Physical Function level, causal constraints are formed by the familiar functional characteristics of individual weapon systems and the staffs available to use and maintain them. Intentional constraints derive from attitudes toward the available information. Can the input be trusted? How good are the estimates? The projections? Is the enemy corrupting the information base?

Constraint shaping sources at the Physical Form level of description are obvious for causal constraints. The geographical distribution of assets relative to the military objective help bound the workspace. The medium of expression for information, as well as accessibility of it, serve as constraints on intentions as the medium influences communication, coordination, and collaboration.

This form of constraint-based analysis attempts to provide insight into the range over which an intelligent actor(s) can make adaptive adjustments to meet expected and unexpected situations on the way to achieving the objective. It directs the analyst's attention to the factors that shape the workspace and teases apart, to the best possible extend, the intentional and causal classes of constraints.

Level of Abstraction	Domain of Military Encounter	Domain of Friendly Resources
Functional Purpose	Goal: To accomplish a military course of action deemed to be in the best interest of the country by the national command authority	
Abstract Function	Priority criteria and given level of acceptable risk balance (e.g. public opinion; congressional support; international relations; rules;regulations)	Time, resource balance within engagement region
Generalized Function	Situation Assessment; information coordination; decision coordination and collaboration	Availability and locational distribution of people, equipment, and weapons
Physical Function	Trust/belief in information; risk/outcome estimation and projection; information assurance	Combatant skills; functional capabilities of equipment and weapons
Physical Form	Information location, method, and form of delivery; communication modes; collaboration modes	Location of people, equipment, weapons; geographic landscape; air space landscape

Figure 36: Constraint-based analysis framework applied to the military command and control work environment.

Some appreciation for how the framework guides effective analysis can be seen by inspecting the illustration of the air interdiction mission that included an air abort of the assets assigned to suppress enemy defenses in the vicinity of the target. Figure 37 shows a mapping of the mission (only the SEAD support is emphasized) onto the analysis framework.

When commander learns about the air mission, s/he must run through the complete Abstraction Hierarchy to reframe the situation in terms of intentional and causal constraints. It is possible that very little time will be needed at some levels (i.e. they are stable and still apply with little or no modification). There is a limited amount of time in which to make a decision. Constraints will shape how this time is spent. For example, one possibility is to execute the mission without SEAD suppression. If there is some question about the state of the SAM site (i.e. there is some doubt that it is active today), the Commander may wish to collect some new intelligence to reassess the site status. But if there are no Uninhabited Aerial vehicles in the area and time is very limited, or the information channel is slow and can only provide low resolution information, then this option may not be seriously considered. This might be the commander's first option but constraints remove it from consideration. As a result the workspace is such that only alternate support options are considered, and if none can be found, a mission abort order is issues.

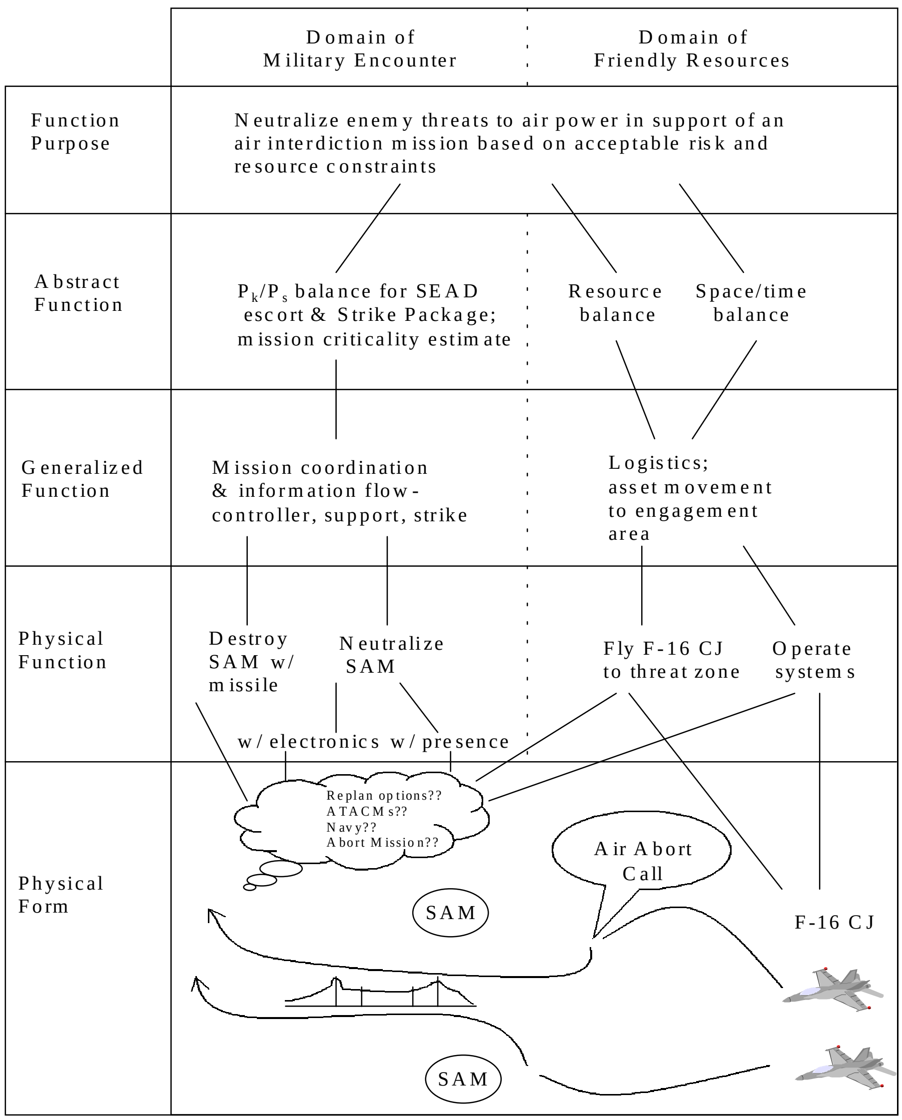

Figure 37: Mapping of threat suppression support (Air Interdiction Mission) to constraint framework.

6 A Proposal for Future Research

To evaluate the usefulness of the concepts presented in this report for military work domains, a research plan needs to be developed. This section outlines a direction suitable for evaluating the CWA framework and the associated measurement tools. It describes the chosen testbeds as well as the experimental and analytical methods that could be used.

There are four stages in the proposed plan: (1) choice of problem domain, (2) cognitive work analysis, (3) development of microworld simulation and interface design, and (4) experimental design and analysis. Each part of the proposed plan is discussed below.

6.1 Choice of problem domain

Before any measures are developed and tested, an appropriate choice of problem domain and scope must be selected. It is important that this domain be representative of (cf. Brunswik, 1956) some target situation or circumstance within the military work environment, since this is the application area for assessing the applicability of any measures of adaptive performance. In addition, it is also important to select the scope of the domain to be manageable with respect to complexity in conducting an experimental investigation.

In order to determine if a problem domain is appropriate for research, a review of current military operations (e.g., command and control) and field studies is recommended. Such a review would provide a basis for narrowing the scope of the research plan and experimental platform. System purposes, components, actors, and interconnections would all be defined during this scoping analysis.

6.2 Cognitive Work Analysis

After the work domain has been defined and narrowed in scope for the purposes of evaluating measures of adaptive performance, a CWA should be performed (Rasmussen et al., 1994; Vicente, in press) to identify all of the relevant behavior-shaping constraints for the selected work environment. This analysis is necessary to determine the dependent variables and appropriate measures for adaptive performance to be tested.

6.3 Microworld Simulation and Interface Design

After a CWA has been performed, it is proposed that a microworld simulation be developed as the testbed for assessing adaptive performance for a wide range of military scenarios for the chosen domain. The simulation would be designed to capture the relevant dependent variables of the problem domain, based on the CWA.

It is also proposed that at least two different interfaces be developed to assess and compare the relative adaptive performance of individuals and teams across many trials. These interfaces can be based on the typical configuration of current systems, the results of the CWA, or any other proposed displays/decision support tools. Because the purpose of this research is to evaluate the measures of adaptation, we would have to make sure that we chose two or more interfaces that are known to have different effects on worker adaptation. The experiment can then assess the sensitivity and diagnosticity of the measures with respect to these known differences. Also, it is essential that the work domain be the same for each interface developed. This will provide a control for experimental design and analysis.

6.4 Experimental Design and Analysis

A framework for developing a strategy for evaluating the proposed measures of adaptive performance is shown in Figure 38, based on the epistemological approach proposed by Xiao & Vicente (in press) for structuring empirical and field studies. There are five levels in this multi-staged framework that represent raw data at different levels of abstraction. The higher levels contain less reference to detail and allow for the integration and generalization of results from different studies. The lower levels contain more detailed information of the raw data and context of a particular study.

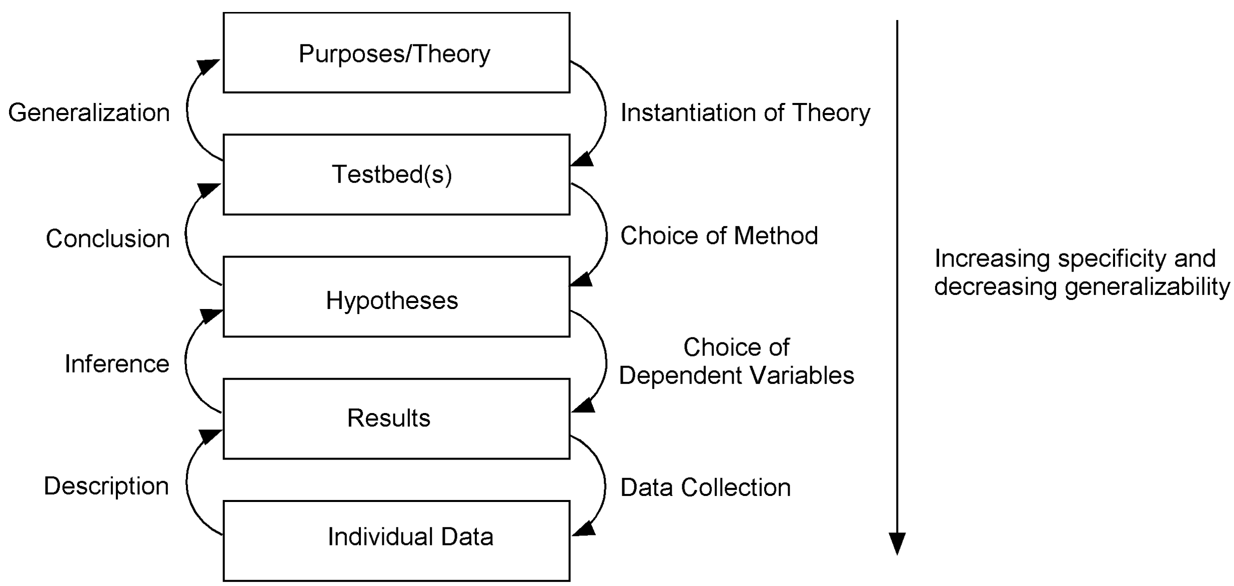

Figure 38: Epistemological approach for empirical studies proposed by Xiao and Vicente (in press).

Planning empirical studies (i.e., experiment design) to investigate a theory or principle is a top-down process with four stages requiring the selection of: (1) appropriate testbeds for instantiating the theory, (2) a method used for generating hypotheses, (3) dependent measures for generating results, and (4) data collection procedures for gathering raw data. Evaluating empirical studies (i.e., experiment analysis) is a bottom-up process with four stages in the reverse order: (1) description of the raw data to form results, (2) inference of the results to assess hypotheses, (3) determination of conclusions with respect to the testbeds, and (4) determination of any generalizations from the testbeds in terms of the theory or principles.

6.4.1 Experiment Design

In terms of future research, the purpose of the investigation is to evaluate proposed measures of adaptive performance of individuals and teams in military work domains. CWA is the guiding analysis framework in the development of these context specific measures of adaptation, taking into account the constraints of the work environment. The first step is to instantiate the theory with the development of testbeds, specific to military work domains. In Section 6.3 we proposed the development of a microworld simulation with alternative displays/decision support systems, designed to make individuals or teams adapt to the same work domain in different ways in order to eventually evaluate various proposed measures of adaptive performance. The second step is to choose a method for generating hypotheses regarding the sensitivity and diagnosticity of the measures of adaptive performance. We propose to use the CWA framework as the basis for predicting which measures will be sensitive and diagnostic, based on the degree and type of behavior-shaping constraint in the work environment for chosen work situations. The third step will require identification of dependent variables in the work environment that will be used for the various measures of adaptive performance. Finally, we propose that the last step, collection of data be automated by the microworld simulation platform.

In order to achieve objectives outlined in the experiment planning process, we need to first conceptualize the organization of the experiment in terms of the military work domain, interface, and individuals/teams, as shown in Figure 39. It is proposed that adaptive performance be assessed in a balanced experiment with different interfaces and different teams, adapting to similar situations in the same work domain. The same scenarios will be presented to each individual or team in order to be able to compare the results. It is important that these scenarios be representative of the target situation in order to be able to generalize any results regarding adaptive performance. Two classes of experiments are proposed: (1) an individual controls the work domain independent of other actors, and (2) a team coordinates actions and interacts to control the work domain.

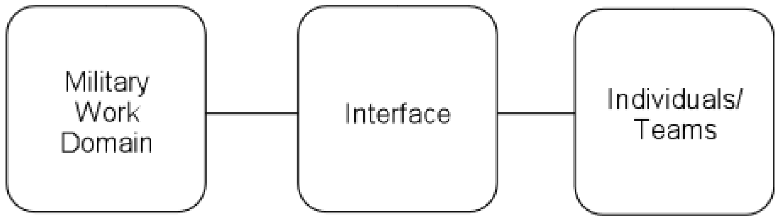

Figure 39: Conceptual structure of experiment.

Participants selected for the experiment should have some familiarity with military operations and the types of missions or scenarios presented. The scale of the experiment should include at least 3 people in each team environment coordinating tasks and controlling the work domain for each trial.

The CWA will be used to identify relevant dependent variables and measures of adaptive performance that are specific to the chosen testbed. It is proposed that AH variability, EIP adaptation, dynamical systems and stability, and higher-order moment measures be evaluated.

6.4.2 Experiment Analysis

The results of the experiment are proposed to be analyzed with the measures of adaptive performance. After the analysis, an assessment of the utility of these measures will be performed.

The analysis of the experiment data will proceed in a bottom-up process, defined by the framework in Figure 38. The first step will be to describe and incorporate the individual data into the proposed measures of adaptive performance. The second step is to compare the results and make inferences with the hypotheses. The primary question to be addressed is whether the measures based on CWA are sensitive and diagnostic with respect to the known impact(s) of the interfaces on worker adaptation. The third step is to generate some conclusions regarding these measures in terms of the testbeds investigated. The last step is to attempt generalizing which measures of adaptive performance will be sensitive and diagostic for military work domains, and under which conditions.

7 Conclusions

Military work domains of the future will put an increasing premium on flexible, adaptive behavior of individuals and teams. Accordingly, it is important that empirical measures of adaptation be developed to assess the effectiveness of different types of systems design interventions in these new environments. In this report, we have adopted Rasmussen's cognitive work analysis (CWA) framework as a basis for addressing this problem. CWA is a constraint-based approach that is tailored to the need for worker adaptation. We have shown how the layers in CWA can be used to develop and organize empirical measures of adaptation. Several example measures have been discussed, demonstrating the usefulness of the approach and directions for future consideration. In addition, case studies showing how our approach can be applied to command and control type problems have been presented for the domains of ambulance dispatch management and military command and control. Finally, we have briefly outlined an agenda for future research incorporating the concepts discussed in the report.

Acknowledgements

This research has been sponsored by a research contract from Logicon Technical Services at Wright-Patterson Air Force Base (Lawrence Wolpert, contract monitor). We would also like to thank Catherine M. Burns at the University of Waterloo for her contributions to the ambulance dispatch management example discussed in Section 5 of the report.

8 References

Abraham, R. H., & Shaw, C. D. (1982). Dynamics - The geometry of behavior, part 1: periodic behavior. Santa Cruz, CA: Aerial Press.

Abraham, R. H., & Shaw, C. D. (1983). Dynamics - The geometry of behavior, part 2: chaotic behavior. Santa Cruz, CA: Aerial Press.

Abraham, R. H., & Shaw, C. D. (1984). Dynamics - The geometry of behavior, part 3: global behavior. Santa Cruz, CA: Aerial Press.

Abraham, R. H., & Shaw, C. D. (1988). Dynamics - The geometry of behavior, part 4: bifurcation behavior. Santa Cruz, CA: Aerial Press.

Ashby, W. R. (1952). Design for a brain: The origin of adaptive behaviour. London: Chapman & Hall.

Ashby, W. R. (1956). An introduction to cybernetics. London: Chapman & Hall.

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

Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.). Berkeley, CA: University of California Press.

Bucciarelli, L. L. (1994). Designing engineers. Cambridge, MA: MIT Press.

Christoffersen, K., Hunter, C. N., & Vicente, K. J. (1994). Cognitive 'dipsticks': Knowledge elicitation techniques for cognitive engineering research (CEL 94-01). Toronto: University of Toronto, Cognitive Engineering Laboratory.

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

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

Flach, J., Hancock, P., Caird, J., & Vicente, K. J. (1995). Global perspectives on the ecology of human-machine systems. Hillsdale, NJ: Erlbaum.

Haken, H. (1988). Information and self-organization: A macroscopic approach to complex systems. Heidelberg: Springer-Verlag.

Hancock, P., Flach, J., Caird, J., & Vicente, K. J. (1995). Local applications of the ecological approach to human-machine systems. Hillsdale, NJ: Erlbaum.

Howie, D. E., & Vicente, K. J. (1998). Measures of operator performance in complex, dynamic microworlds: Advancing the state of the art. Ergonomics, 41, 385-400.

Hoyt, D. F., & Taylor, C. R. (1981). Gait and energetics of locomotion in horses. Nature, 292, 239-240.

Kelso, J. A. S. (1995). Dynamic patterns: The self-organization of brain and behavior. Cambridge, MA: MIT Press.

Kugler, P. N., Shaw, R. E., Vicente, K. J., & Kinsella-Shaw, J. (1990). Inquiry into intentional systems I: Issues in ecological physics. Psychological Research, 52, 98-121.

Morefield, C.L. (1995). Air Force information applications in the 21st Centrury. In G. H. McCall and J. A. Corder (Eds.) New world vistas: Air and space power for the 21st century – Information applications volume. USAF Scientific Advisory Board.

Moray, N., Lootsteen, P., & Pajak, J. (1986). Acquisition of process control skills. IEEE Transactions on Systems, Man, and Cybernetics, SMC-16, 497-504.

Moray, N., Sanderson, P. M., & Vicente, K. J. (1992). Cognitive task analysis of a complex work domain: A case study. Reliability Engineering and System Safety, 36, 207-216.

Newell, K. M., & Hancock, P. A. (1984). Forgotten moments: A note on skewness and kurtosis as influential factors in inferences extrapolated from response distributions. Journal of Motor Behavior, 16, 320-335.

O'Donnell, R. D. & Eggemeier, F.T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), Handbook of perception and human performance Vol. 2 (pp. 42-1 - 42-49). New York: Wiley.

Pawlak, W. S., & Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. International Journal of Human-Computer Studies, 44, 653-688.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. Cambridge: Cambridge University Press.

Pew, R. W., & Mavor, A. S. (1998). Modeling human and organizational behavior: Applications to military simulations. Washington, DC: National Academy Press.

Port, R. F., & van Gelder, T. (1995). Mind as motion: Explorations in the dynamics of cognition. Cambridge, MA: MIT Press.

Rasmussen, J. (1986). Information processing and human-machine interaction: An approach to cognitive engineering. Amsterdam: North-Holland.

Rasmussen, J., Pedersen, O. M., & Grønberg, C.D. (1987). Evaluation of the use of advanced information technology (expert systems) for data base system development and emergency management of non-nuclear industries (RISØ-M-2639). Roskilde, DK: Risø National Laboratory.

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

Sanderson, P. M., Verhage, A. G., & Fuld, R. B. (1989). State-space and verbal protocol methods for studying the human operator in process control. Ergonomics, 32, 1343-1372.

Schluter, D., & Grant, P. R. (1984). Determinants of morphological patterns in communities of Darwin's finches. The American Naturalist, 123, 175-196.

Sperandio, J.-C. (1978). The regulation of working methods as a function of work-load among air traffic controllers. Ergonomics, 21, 195-202.

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Vicente, K. J. (in press). Cognitive work analysis: Toward safe, productive, and healthy computer-based work. Mahwah: Erlbaum.

Vicente, K. J. (1996). Improving dynamic decision making in complex systems through ecological interface design: A research overview. System Dynamics Review, 12, 251-279.

Vicente, K. J. (1991). Supporting knowledge-based behaviour through ecological interface design. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Vicente, K. J., & Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating "direct perception" in complex work domains. Ecological Psychology, 2, 207-250.

Wickens, C. D. (1992). Engineering psychology and human performance (2nd ed.). New York: Harper-Collins.

Wong, W. B. L., Sallis, P. J., & O'Hare, D. (1998). The ecological approach to interface desogn: Applying the abstraction hierarchy to intentional domains. Proceedings of the Eighth Australian Conference on Computer-Human Interaction OzCHI'98. Australia: IEEE Computer Society Press.

Woodcock, A. E. R. (1988). Indications and warnings as an input to the C^3 Process. In S. E. Johnson and A. H. Levis (Eds.), Science of command and control: Coping with uncertainty (pp.32-47). Washington: AFCEA International Press.

Xiao & Vicente (in press). Epistemological analysis in empirical (laboratory and field) studies: Data and generalizations. Human Factors, Santa Monica, CA: HFES.

Yu, X., Chow, R., Jamieson, G. A., Khayat, R., Lau, E., Torenvliet, G. L., Vicente, K. J., & Carter, M. W. (1997). Research on the characteristics of long-term adaptation (JAERI Final Report, CEL 97-04). Toronto: University of Toronto, Cognitive Engineering Laboratory.

Yu, X., Lau, E., Vicente, K. J., & Carter, M. W. (1998). Advancing performance measurement in cognitive engineering: The abstraction hierarchy as a framework for dynamical systems analysis. In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting (pp. 359-363). Santa Monica, CA: HFES.

CEL TECHNICAL REPORT SERIES

CEL 96-02 "Skill, Participation, and Competence: Implications of Ecological Interface Design for Working Life"
• Peter Benda, Giuseppe Cioffi, & Kim J. Vicente

CEL 96-03 "Practical Problem Solving in a Design Microworld: An Exploratory Study'
• Klaus Christoffersen

CEL 96-04 "Review of Alarm Systems for Nuclear Power Plants"
• Kim J. Vicente

CEL 96-05 "DURESS II User's Manual: A Thermal-hydraulic Process Simulator for Research and Teaching"
• Lisa C. Orchanian, Thomas P. Smahel, Dianne E. Howie, & Kim J. Vicente

CEL 96-06 "Research on Factors Influencing Human Cognitive Behaviour (III)"
• Dianne E. Howie, Michael E. Janzen, & Kim J. Vicente

CEL 96-07 "Application of Ecological Interface Design to Aviation"
• Nick Dinadis & Kim J. Vicente

CEL 96-08 "Distributed Cognition Demands a Second Metaphor for Cognitive Science"
• Kim J. Vicente

CEL 96-09 "An Experimental Evaluation of Functional Displays in Process Supervision and Control"
• Catherine M. Burns and Kim J. Vicente

CEL 96-10 "The Design and Evaluation of Transparent User Interfaces: From Theory to Practice"
• Beverly L. Harrison

CEL 97-01 "Cognitive Functioning of Control Room Operators: Final Phase"
• Kim J. Vicente, Randall J. Mumaw, & Emilie M. Roth

CEL 97-02 "Applying Human Factors Engineering to Medical Device Design: An Empirical Evaluation of Two Patient-Controlled Analgesia Machine Interfaces"
• Laura Lin

CEL 97-03 "ADAPT User's Manual: A Data Analysis Tool for Human Performance Evaluation in Dynamic Systems"
• Xinyao Yu, Farzad S. Khan, Elfreda Lau, Kim J. Vicente, & Michael W. Carter

CEL 97-04 "Research on the Characteristics of Long-Term Adaptation"
• Xinyao Yu, Renée Chow, Greg A. Jamieson, Rasha Khayat, Elfreda Lau, Gerard Torenvliet, Kim J. Vicente, & Michael W. Carter

CEL 98-01 "Applying Human Factors Engineering to Medical Device Design: An Empirical Evaluation of Patient-Controlled Analgesia Machine Interfaces"
• Laura Lin

CEL 98-02 "Building an Ecological Foundation for Experimental Psychology: Beyond the Lens Model and Direct Perception"
• Kim J. Vicente

CEL 98-03 "Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-based Work"
• Kim J. Vicente

CEL 98-04 "Ecological Interface Design for Petrochemical Processing Applications"
• Greg. A. Jamieson & Kim J. Vicente

CEL 98-05 "The Effects of Spatial and Temporal Proximity of Means-end Information in Ecological Display Design for an Industrial Simulation"

CEL 98-06 "Research on Characteristics of Long-term Adaptation (II)"
• Xinyao Yu, Gerard L. Torenvliet, & Kim J. Vicente

CEL 98-07 "Integrated Abstraction Hierarchy and Plan-Goal Graph Model for the DURESS II System: A Test Case for Unified System- and Task-based Modeling and Interface Design"
• Christopher A. Miller & Kim J. Vicente

CEL 98-08 "Comparative Analysis of Display Requirements Generated via Task-Based and Work Domain-based Analyses: A Test Case Using DURESS II"
• Christopher A. Miller & Kim J. Vicente

CEL 98-09 "Abstraction Decomposition Space Analysis for NOVA's E1 Acetylene Hydrogenation Reactor"
• Christopher A. Miller & Kim J. Vicente

CEL 99-01 "Development of an analytical framework and measurement tools to assess adaptive performance of individual and teams in military work domains"
• John R. Hajdukiewicz, Kim J. Vicente, & Robert G. Eggleston