

HUMAN SUPERVISORY CONTROL ISSUES IN NETWORK CENTRIC WARFARE

**P.J. MITCHELL
M.L. CUMMINGS
T.B. SHERIDAN**

MASSACHUSETTS INSTITUTE OF TECHNOLOGY*

PREPARED FOR BOEING, PHANTOM WORKS

HAL2004-01

SEPTEMBER, 2004



<http://halab.mit.edu>

e-mail: halab@mit.edu

*Department of Aeronautics and Astronautics, Cambridge, MA 02139

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
INTRODUCTION	2
CENTRAL NCW HUMAN SUPERVISORY CONTROL ISSUES - OUTLINE.....	4
KEY DEFINITIONS	6
CENTRAL NCW HUMAN SUPERVISORY CONTROL ISSUES - DETAIL.....	7
<i>1. Appropriate Levels of Automation</i>	<i>7</i>
<i>2. Information Overload</i>	<i>10</i>
<i>3. Adaptive Automation</i>	<i>12</i>
<i>4. Distributed Decision-making and Team Coordination.....</i>	<i>15</i>
<i>5. Complexity Measures.....</i>	<i>16</i>
<i>6. Decision Biases</i>	<i>18</i>
<i>7. Attention Allocation</i>	<i>19</i>
<i>8. Supervisory Monitoring of Operators</i>	<i>21</i>
REFERENCES	24
APPENDICES	27
<i>Appendix A – MIT Candidate NCW Experimental Platforms</i>	<i>27</i>
<i>Appendix B – Command and Control Research Groups Related to NCW</i>	<i>29</i>

EXECUTIVE SUMMARY

Network centric warfare (NCW) is a concept of operations that seeks to increase combat power by linking battlespace entities to effectively leverage information superiority. A network centric force must be supported by sophisticated automated systems, so human-computer interactions are an important aspect of overall performance. These interactions are examples of human supervisory control (HSC), in which a human operator intermittently interacts with a computer, receiving feedback from and providing commands to a controlled process or task environment, which is connected to that computer. The Department of Defense (DoD) has recognized that a lack of understanding of HSC issues relevant to NCW is a significant barrier limiting NCW's potential benefits. This report identifies eight central HSC issues that could significantly impact operator performance in NCW: Appropriate levels of automation, information overload, adaptive automation, distributed decision-making through team coordination, complexity measures, decision biases, attention allocation, and supervisory monitoring of operators.

The adoption of NCW principles is often misunderstood as requiring increased *levels of automation*, which makes this a particularly acute problem as NCW is implemented. For the average operator, implementation of NCW will exponentially add to the number of available information sources as well as the volume of information flow. Without measures to mediate this volume, *information overload* will occur much more often than in the past, as it will be far easier for operators to obtain or be given more information than they can adequately handle. One way to alleviate this problem is through *adaptive automation*, which has been shown in certain cases to lower workload. There will also be a corresponding increase in information complexity, quantified by *complexity measures*, which can cause a loss of situation awareness or an unmanageable increase in mental workload. It is therefore essential that the interfaces with which NCW operators interact help to reduce and manage this increased level of data complexity.

A more fundamental issue associated with the increase in the number of available information sources, volume of information, and operational tempo under NCW are operator *attention allocation* strategies. NCW hinges on successful information sharing, so knowledge of the relationship between perceived and actual high priority tasks and associated time management strategies, as well as the impact of task disruptions is critical. As a result of NCW information sharing, command and control (C2) structures will change significantly. Traditional methods where commands are passed down from higher levels in a command hierarchy will, at least, be partially replaced by *distributed decision-making* and low-level *team coordination*. Therefore, understanding how to make effective, time-pressured decisions within these organizational structures takes on greater importance in NCW. These redefined C2 structures will drive an increase in information-sharing tempo and rapid decision-making. Under these time pressures, the use of heuristics and other naturalistic decision-making methods may be subject to undesirable *decision biases*, both for individuals and groups. Lastly, how automated technology can be leveraged in order to observe and diagnose HSC issues during *supervisory monitoring of operators* is another significant area of concern since NCW will contain embedded HSC systems.

INTRODUCTION

What is Network Centric Warfare?

As stated by the Department of Defense (DoD) in their 2001 Report to Congress, Network Centric Warfare (NCW) is ‘no less than the embodiment of an Information Age transformation of the DoD’ (DoD 2001). It is a concept of operations envisioned to increase combat power by effectively linking or networking knowledgeable entities in a battlespace. Greater combat power is generated through the creation of shared situational awareness, increased speed of command, self-synchronization*, and higher operational tempo, lethality and survivability (Alberts et al. 2000).

NCW’s basic tenets (Figure 1) are as follows (DoD 2001):

- (1) A robustly networked force improves information sharing.
- (2) Information sharing and collaboration enhance the quality of information and shared situational awareness.
- (3) Shared situational awareness enables self-synchronization.
- (4) These, in turn, dramatically increase mission effectiveness.

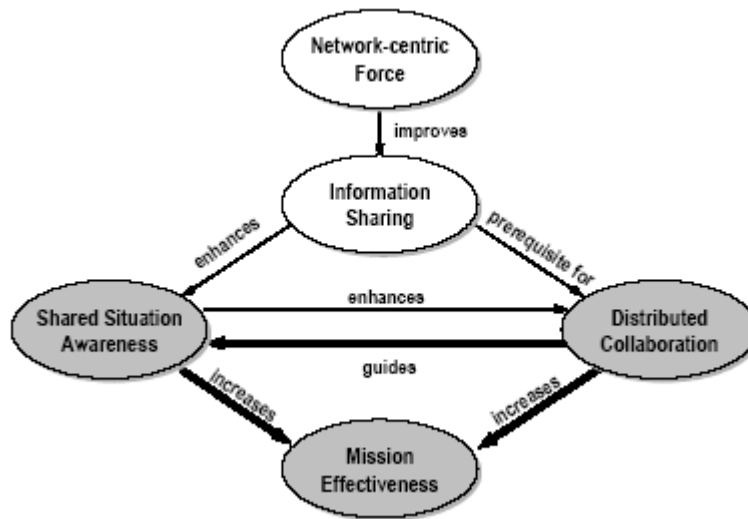


Figure 1. Tenets of NCW (Madni and Madni 2004)

A force with these capabilities is therefore able to increase combat power by leveraging information superiority, rather than through the traditional method of sheer numerical superiority.

Almost as important as what NCW is, it must be said what it is *not*, as there are many myths surrounding the concept. As Alberts (2000) outlined, the most common one is that NCW focuses primarily on the network. In actuality, NCW is about utilizing networking principles to increase combat power. Having a robust network in place to enable NCW is a prerequisite to achieving

* The ability of a force to organize and synchronize complex warfare activities from the bottom up.

NCW's goals, but it is only a part of the larger picture. Second, there is a widespread misconception that NCW represents an attempt to automate warfare. NCW is not about giving responsibility to the "network" or machines to run warfare. Rather, it is about trying to exploit information to more effectively bring all of a force's assets to bear on a particular situation. There will be logical places for automation to help these processes occur more smoothly, but these are a consequence of NCW, not a reason for it in the first place.

What is Human Supervisory Control?

Human supervisory control (HSC) is the process by which a human operator intermittently interacts with a computer, receiving feedback from and providing commands to a controlled process or task environment, which is connected to that computer. Figure 2, adapted from Sheridan (1992) illustrates this concept.

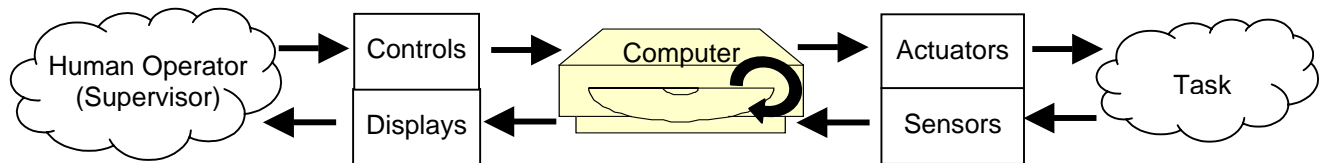


Figure 2. Human Supervisory Control (Sheridan 1992)

Human supervisory control is comprised of five generic functions, usually accomplished in the following cyclical order: planning a computer-based task, teaching the computer what was planned through various inputs, monitoring the computer's actions for errors and/or failures, intervening when the plan has been completed or the computer requires assistance, and then the human learns from the experience (Sheridan 1992).

The Role of Human Supervisory Control in Network Centric Warfare

NCW is rooted in information superiority, generated by a robustly networked force that is supported by automated technologies and systems. The resulting increase in use of automation is a fundamental component of NCW; thus in the context of human interaction, NCW is a high level human supervisory control problem. The number and types of human-machine supervisory interfaces will expand accordingly, as full-automation (no human involvement) of most tasks is difficult to achieve. Fully automated systems require highly reliable error handling capabilities and the ability to effectively deal with unforeseen circumstances and faults (Parasuraman et al. 2000). This level of reliability is difficult if not impossible to achieve in the NCW domain, and even if this level of automation were possible, it may not be implemented due to users' lack of trust.

Due to the increasing importance of HSC in NCW, the DoD has recognized that a lack of automation reliability and understanding of relevant HSC issues, as experienced both by individuals and teams, are among the primary barriers limiting exploitation of the full potential of NCW (DoD 2001). The following report seeks to identify, describe and prioritize HSC tasks within existing NCW systems that exhibit or are at most risk of exhibiting degraded performance due to the complexity of networked systems. In addition, MIT NCW research capabilities are discussed along with an outline of those current agencies that are conducting similar research in command and control arenas.

CENTRAL NCW HUMAN SUPERVISORY CONTROL ISSUES - OUTLINE

Eight major human supervisory control issues have been identified as having the most impact on operator performance during NCW:

- **Appropriate Levels of Automation**
 - What level of automation is most useful for different aspects of NCW tasks, particularly data fusion?
- **Information Overload**
 - At what point does information cognitively saturate operators even though it is perfectly reliable?
- **Adaptive Automation**
 - How should NCW operators deal with transient information overload? Moreover, should the computer determine this or should the human adjust workload levels?
- **Distributed Decision-Making and Team Coordination**
 - Does adding more decision makers help or hinder NCW/HSC processes, and how can team members effectively share the load and maintain coordination?
- **Complexity Measures**
 - How can complexity of NCW tasks be characterized and how does this impact human performance?
- **Decision Biases**
 - When there are uncertainties and multiple plausible hypotheses, how can technology mitigate human tendencies to use inappropriate heuristics and become biased in making decisions?
- **Attention Allocation**
 - When there is a mix of different preview times (time to consider the signals before something has to be done), namely some urgent, some not so urgent, how can a human best allocate attention, and how can the computer help?
- **Supervisory Monitoring of Operators**
 - How can a supervisor (either a computer or a human) effectively detect decreases in performance across multiple human operators, and how should the supervisor respond to them?

Table 1 demonstrates how the issues above relate to the primary HSC elements of planning, communicating the plan, monitoring the plan's execution, intervening when necessary, and learning. It is salient to note that planning has been recognized as a central part of nearly all issues, while learning plays a more minor role in our focus on supporting relevant technologies. This does not mean that the learning elements of HSC are not important, just that they are less relevant in the context of the technology focus of this investigation.

Table 1. Major HSC issues in NCW and their primary elements

Issue	Primary HSC Elements				
	Planning	Communicating	Monitoring	Intervention	Learning
Appropriate Levels of Automation	x		x	x	x
Information Overload	x				
Adaptive Automation	x	x		x	
Distributed Decision-Making and Team Coordination	x		x		x
Complexity Measures	x	x	x	x	
Attention Allocation			x		
Decision Biases and Heuristics				x	
Supervisory Monitoring of Operators	x	x	x	x	x

KEY DEFINITIONS

TASK: Specified variables must be brought to specified states under given time constraints.

AUTOMATION: When some variables are in specified states, IF...THEN mechanisms are applied and change certain (and possibly other) variables in specified ways.

ADAPTIVE AUTOMATION: Either the IF...THEN rules change by themselves based on the operator's state or performance, or the operator can change the IF...THEN rules at will.

NETWORK CENTRIC: Two or more human operators access information from each other and the system about remote activities. The communication can be subject to noise and/or constraints of delay, etc.

SUPERVISORY CONTROL: Operators activate and deactivate automation (and change the rules insofar as it is adaptive) in an effort to perform a given task.

MANUAL CONTROL: Operators use available controls to perform the task directly without the aid of automation.

HIGH WORKLOAD or **COMPLEXITY:** Many tasks are being posed and/or the variables within these tasks are changing simultaneously and rapidly so that it is difficult for the operator to maintain understanding of the unfolding situation.

DATA FUSION: Putting together information in a more comprehensible format, possibly with some advice on what to do (the latter being **DECISION AIDING**).

1. APPROPRIATE LEVELS OF AUTOMATION

The Patriot missile system has a history of friendly fire incidents that can at least be partially attributed to a lack of understanding of human limitations in supervisory control. In the recent war on Iraq, there were 12 Patriot engagements—but three of them with our own planes. On March 23rd, 2003, a RAF Tornado GR4 was shot down in friendly airspace by a Patriot. Two days later, a USAF F-16 fighter pilot received a warning that he was being targeted by hostile radar, and fired a missile in self-defense. It was a Patriot missile battery aiming at him. On April 2nd, 2003, just nine days later, another Patriot shot down a US Navy F/A-18 returning home from a mission. The Patriot missile has two modes: semi-automatic (management by consent) and automatic (management by exception). These friendly fire incidents are believed to be a result of automatic mode, as there are known “ghosting” problems with the radar and occasional troubles with friend or foe aircraft identification responses. Under management by exception, a short time period of approximately 15 seconds is allowed to reject a computer’s decision, which may often be insufficient to solve such difficulties adequately. Despite obvious flaws in the operation of automatic mode, it is still frequently used by Patriot crews. Because of problems inherent in the system, it may be that more human involvement in the firing process will reduce the probability of accidents, which would constitute a lowering of the level of automation.

Automation is not simply ‘on’ or ‘off’; there is a range of levels where allocation of function is shared between man and machine. Sheridan and Verplank (1978) outlined a scale from 1-10 where each level represented the machine performing progressively more tasks than the previous one, as shown in Table 2.

Table 2. Levels of Automation (Sheridan and Verplank 1978)

Automation Level	Automation Description
1	The computer offers no assistance: human must take all decision and actions.
2	The computer offers a complete set of decision/action alternatives, or
3	narrows the selection down to a few, or
4	suggests one alternative, and
5	executes that suggestion if the human approves, or
6	allows the human a restricted time to veto before automatic execution, or
7	executes automatically, then necessarily informs humans, and
8	informs the human only if asked, or
9	informs the human only if it, the computer, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

Human interaction with automation represents a range of intermediate levels from 2-6 on this scale. For routine operations, higher levels of automation (LOAs) in general result in lower mental workload, while the opposite is true for low levels of automation (Kaber et al. 2000). However, if automation fails, workload can grow rapidly. Optimal human performance in monitoring occurs at moderate levels of workload; the relationship between workload and performance is given by the Yerkes-Dodson Law (Yerkes and Dodson 1908) shown in Figure 3.

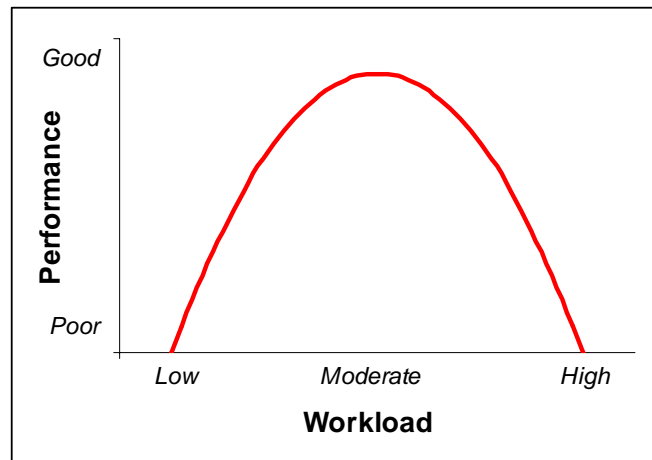


Figure 3. The Yerkes-Dodson Law

It is possible to have a LOA too high or too low, each with its own distinct set of problems (Billings 1997):

- HSC problems if LOA is too high
 - Manual or mental skill degradation
 - Loss of situational awareness due to lack of automation transparency, complexity, and inadequate feedback
 - More advanced automation issues such as brittleness & literalism; in other words, the automated system might not be able to handle novel or unexpected events, as well as operate effectively in conditions near or at the edge of the intended operating envelope
 - Time and difficulty to diagnose failures and manually take over
- HSC problems if LOA is too low
 - Cognitive and working memory overload in routine tasks under time pressure
 - Human decision biases and heuristics
 - Lack of repeatability and consistency
 - Complacency and boredom
 - Greater human interdependency and chaos when something fails, unless fail soft safeguards are in place

NCW is often misunderstood as focusing entirely on technology upgrades (Alberts et al. 2000) with the implied assumption that this will result in increased automation, and hence improved performance. There are several factors behind this misconception, including:

- The amount of information available under NCW is greater than ever before,
- The synthesis of this additional data to gain understanding and situational awareness is more abstract and therefore more difficult to interpret (e.g. interaction with geographically distant teams, etc),
- Expectations of increased combat tempo require faster decisions and actions that are perceived as needing additional automation to achieve, and
- The addition of automation on top of automation adds complexity, and thus confusion if the operator is called upon to diagnose and find problems.

These factors highlight both the intrinsic need for automation and the difficulties that will be encountered with complex automated systems. Care must be taken to consider each of the roles human and machine has in the collaborative interaction that is HSC, and automation only introduced when there is a specific need to do so (Billings 1997). This will ensure better overall performance across the dynamically changing set of responsibilities assigned to operators in NCW.

As automation has become more complex and capable of accomplishing a much wider range of tasks, a need for more flexible human-automation interaction has arisen. For this reason, Parasuraman et al. (2000) proposed that most tasks could be broken down into four separate information processing stages (information acquisition, information analysis, decision selection and action implementation), and that each could be assigned a level of automation separate and possibly different from the others. However, in the context of flexible human-automation interaction, subdividing a problem into these abstract stages may not go far enough. As proposed by Miller and Parasuraman (2003), each information processing task can be further divided into simple sub-tasks with differential levels of automation. For NCW, generalized cognitive tasks under these proposed information processing stages can be defined:

- *Information acquisition*
 - Monitoring resources (such as friendly forces)
 - Monitoring systems (such as surveillance networks)
 - Communications
- *Information analysis*
 - Data fusion & display techniques
- *Decision selection*
 - Planning
 - Re-planning
 - Rapid resource allocation
- *Action implementation*
 - Individual vs. team interaction

Human supervisory control interactions with automation primarily fall under the analysis and decision processes (Figure 4). Of these two stages, information analysis is the one most affected by conversion of the military to network-centric principles, as the number and variety of available information sources to a robustly networked force is expected to increase dramatically.

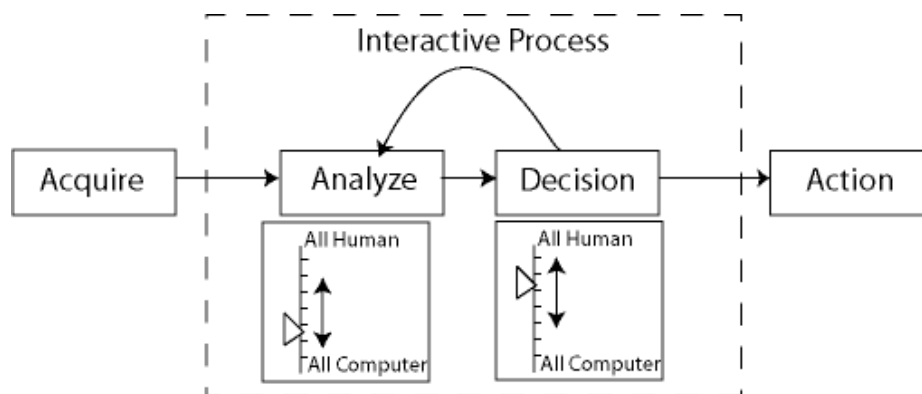


Figure 4. Proposed interactive process overlaying human information processing (Sheridan 2002)

The potential for cognitive overload in the information analysis phase for NCW is likely to be caused by problems with data fusion. Data fusion in this sense is defined as the process by which raw information from disparate sources is filtered and integrated into relevant groupings before being displayed to users. The LOA of data fusion can vary. Low levels of automation could include trend or predictive displays for single or multiple variables of interest, such as for tracking of enemy forces. Higher levels could include smart agents providing context dependent summaries of relevant information to users (Parasuraman et al. 2000).

Trust issues are a critical factor to consider when choosing an appropriate LOA for data fusion in NCW. The process of data fusion is not always perfect; the automated data fuser could, for example, omit important pieces of information in certain circumstances by misunderstanding the context, or combine data sources incorrectly. This has implications for performance, as to fully and appropriately use automation an operator's trust must be calibrated to a level corresponding to the machine's or function's trustworthiness (Muir 1987). In this case, if a NCW operator has too much trust in the automation driving data fusion, they may rely on it even when it has made obvious errors. The opposite may also be true. Distrusting the data fusion algorithm when it is perfectly capable may lead to disuse or deliberate misinterpretation of its results to fit the operator's mental models, even if this is very demanding and/or time consuming (Muir 1987). A possible solution to this is to introduce a human-computer collaborative element to data fusion, whereby the operator is allowed to determine what data elements are fused at what times, depending on their comfort level with the automation. Displaying the automation's confidence in various aspects of its display to users also would facilitate better calibration of trust and operator understanding of the data fusion process.

Research should also be undertaken to determine appropriate levels of automation for the aforementioned tasks, both individually and in combination with other tasks. A possibility would be to measure performance in a typical NCW situation with varying levels of automation. In addition to typical measures such as response time, accuracy of responses, etc, the ability to objectively gauge an operator's ability to understand what is happening in the network will be critical. Thus, an objective situation awareness (SA) metric will also need to be developed. Development of an increased level of shared situation awareness and knowledge (SSAK) is major high level tenet of NCW (DoD 2001), and is critical to realizing NCW's promise of substantial increases in combat power. Yet, a general objective measurement of situational awareness has been elusive. Direct system performance measures that use specific scenario manipulations to measure targeted aspects of SA are appropriate only in limited situations in which SA drives performance. However, the correlation between SA and performance in these cases is debated (Pew 2000). Direct experimental techniques, such as the SAGAT (Situation Awareness Global Assessment Technique) (Endsley 1988), have been used in command and control research, but are not without problems. Questions remain as to whether SAGAT measurement techniques disrupt task performance, and if expectation of probes alters peoples' natural behaviors (Pew 2000). Instead, subjective measures have traditionally been used in the human supervisory command and control domain, whereupon subjects self-rate their level of SA. It has been shown that these methods are of limited utility, providing more insight to judgment processes than situation awareness (Jones 2000).

2. INFORMATION OVERLOAD

On March 28th, 1979, the worst US commercial nuclear power plant accident in history occurred at Three Mile Island in Middletown, Pennsylvania. The problem began when a main feedwater pump failure caused the reactor to automatically shut-down. In response to this, a relief valve opened to reduce the pressure of the system, but stuck open. There was no indication to plant controllers that this had occurred. Due to the stuck valve, there was significant loss of reactor coolant water, subsequently causing the core of the reactor to overheat. There were no instruments that showed the level of coolant in the core, so it was thought to be acceptable based on the pressurizer coolant level. This led to a series of human actions that actually made

the problem worse, ending with a partial meltdown of the core. Operators were overwhelmed with alarms and warnings, numbering in the hundreds, over a very short period of time. They did not possess the cognitive capacity to adequately deal with the amount of information given to them during the unfolding events. Instead, they coped by focusing their efforts on several wrong hypotheses, ignoring some pieces of information that were inconsistent with their mental model.

According to the DoD, the Global Information Grid (GIG) will be the enabling building block behind NCW and its promised benefits of information and decision superiority over foes (Figure 5). The GIG is the end-to-end set of information capabilities, associated processes and personnel for collecting, processing, storing, and disseminating information to those who require it, on the battlefield or otherwise (DoD 2001).

Metcalf's Law states that the usefulness, or utility, of a network equals the square of the number of users (Shapiro and Varian 1999); what this means is that these promised networking capabilities will allow forces to have access to exponential amounts of information over today's forces. Thus, information intake for the average operator under NCW will be increased. Even if the information complexity does not increase (which is unlikely), mental workload will increase accordingly. As previously mentioned, the Yerkes-Dodson Law illustrates that beyond a task-dependent moderate level of arousal, an individual will become cognitively overloaded and their performance will drop. The problem is being able to predict when and how this overload will occur, so that the amount of information any single person or group is required to process is manageable.



Figure 5. The GIG as a building block of NCW (DoD 2001)

Shannon information theory (Shannon 1948) formally quantifies how much information is conveyed by an event. It is influenced by three variables: (1) The number of possible events, (2) The probabilities of those events, and (3) Sequential constraints, in other words, the context in which the events occur. In essence, this can be summarized by the following statement: The less the expectation of an event, the greater the information that is conveyed by it. For our purposes, we are interested in measuring the average information conveyed by a series of events with differing probabilities over time. The average bits of information (H_{ave}) is given by:

$$H_{ave} = \sum_{i=1}^n P_i \left(\log_2 \frac{1}{P_i} \right), \text{ where } P_i = \text{probability of event "i", given previous events}$$

Humans may be modeled as an information channel (Figure 6), where we want to know how much information is transmitted from stimulus to response, and the bandwidth. Important elements include stimulus inputs, such as display elements or voice transmissions (H_s), and response

information (H_R), such as decisions made due to this information. The response is composed of information faithfully transmitted (H_T) and irrelevant noise introduced into the signal.

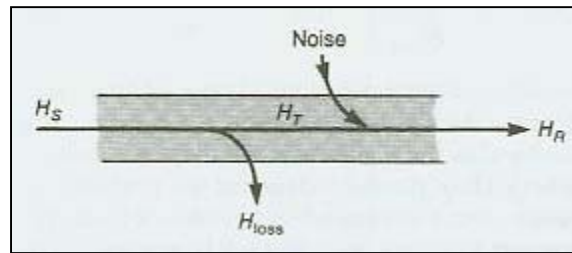


Figure 6. A model of information transmission (Wickens and Hollands 2000)

The stimulus information that is lost is H_{LOSS} ; this could represent information on a display presented to an operator, but not observed, or forgotten in later actions requiring this data. Essentially, variety in the stimuli does not show up in the response. This would be associated with cognitive overload, resulting from time pressure or a loss of SA. Each of these variables in the information processing model can be measured according to the formula above. Bandwidth is obtained by dividing H_T by response time.

Information theory could then be applied to studying the problem of information overload in NCW through several methods. Taking some generic NCW tasks from Section 1, experiments could be run in which only one of these tasks was performed by an operator, with the stimuli and possible range of responses carefully defined so as to provide meaningful measurements of the above variables. The information value of the stimuli could be increased during the experiment to discover a threshold above which H_R degrades and to observe *how* the operator's behavior deteriorates. This could be repeated with different generic tasks, and then gradually with multiple tasks. With this knowledge, a general guideline for maximum information intake could be generated for single and combinations of generalized NCW tasks, as well as providing better knowledge of information overload. Another approach would be to research information presentation that decreases information loss while still maintaining the same stimulus information content.

3. ADAPTIVE AUTOMATION

Instant messaging was a primary means of communication between Navy ships during Operation Iraqi Freedom in 2003. While instant messaging, otherwise known as chat, has many advantages for rapid response in critical time-pressure command and control situations, operational commanders have found it difficult to handle large amounts of information generated through chat, and then synthesize relevant knowledge from this information (Caterinicchia 2003). Chat enables faster responses in time critical command and control situations, but as described by the military, managing and synthesizing the large amount of information transmitted can sometimes be a "nightmare". The addition of instant messaging in the command and control loop requires a division of attention from the primary task, which may not always be appropriate. If the power of an intelligent automated and adaptive agent was harnessed so that the computer could determine more optimal scheduling patterns for the presentation of instant messages, it is possible that information would not be a detrimental interruption and overall human performance improved.

Warfare is often characterized by long periods of inactivity followed by intense periods of action during which time-critical decisions must be made. At these times performance is most critical, yet it will likely suffer due to the temporary information overload placed on the operator. With NCW and the emergence of a robustly networked force, the amount of information available to military personnel at all levels is exponentially greater. Therefore, the problem of information overload, particularly during brief bursts of actions, will become much more common.

One method to alleviate such problems is the use of adaptive automation (AA). AA has been shown to improve task performance (Hilburn et al. 1997; Prinzel et al. 2003), situational awareness (Kaber and Endsley 2004) and lower workload (Prinzel et al. 2003). Two important research questions to answer are (1) when to use adaptive automation to determine under what circumstances the LOA should change, and (2) whether the computer or the human decides to change the LOA. Changes in the level of automation can be driven by specific events in the task environment, models of operator performance and task load, physiological methods, or by changes in operator performance (Parasuraman et al. 1992).

Specific cues in the environment used to change the LOA may be either time or event-related. For example, it may be determined that in order to maintain vigilance levels in a task, automation will be turned off for a certain period of time once an hour, forcing the operator to execute manual control during that period. Alternatively, it may be known that particular events in the environment will cause higher levels of workload than desired for a short time, during which the automation can increase to compensate. An example of this would be an operator responsible for multiple UAVs performing a bombing mission. It is well known that the cruise and loiter phases of flight are low in workload, but that the approach and bombing phases require significant increase in human workload. As a preemptive measure, adaptive automation could increase the LOA during the periods of high workload automatically. This approach is problematic because it is scenario specific, will not handle unexpected situations well (desired changes in LOA must be pre-determined), and does not take into account operator variability. Some operators may have a much higher threshold for handling workload than others and thus may not require a change in LOA.

A related method of cueing AA is through models of operator performance, as they can be used to predict the effectiveness of humans during particular processes and behaviors. Therefore, the model's forecasted level of operator mental workload or performance on any number of specified tasks can be used to change the LOA. What is defined as an acceptable level of any measure predicted by the model must be carefully defined in advance. As defined by Laughery and Corker (1997), there are two main categories of human performance models: reductionist and first principles. Reductionist models break down expected behaviors into successively smaller series of tasks until a level of decomposition is reached that can provide reasonable estimates of human performance for these task elements. First principles models are based on structures that represent basic principles and processes of human performance. Performance models offer the advantage of flexibility in the sense that they can apply to a large range of situations, even unexpected ones, but often are costly and difficult to develop, especially if higher reliabilities are desired.

Psychophysiological measures such as the electroencephalogram (EEG), event-related brain potentials (ERPs), eye movements and electroculography (EOG), electrodermal activity (EDA), heart rate and heart rate variability (HRV), breathing rate and blood pressure have all been correlated with mental workload to varying degrees of success. Experimentally, these methods are advantageous because they are not task specific and they can continuously record data. The problem is that in the past, many devices used to take these measurements have been obtrusive and physically uncomfortable for subjects performing the experiments, creating a possible anxiety effect. While this is still a barrier to overcome, technologies such as the LifeShirt®, wireless EEG sensor headsets (Berka et al. 2004) and stereo cameras for tracking eye movements without the use of headgear can provide more ecological measurements of psychophysiological metrics. Other significant problems with psychophysiological measures are the large amount of noise present in readings, and extreme individual variability. A way to lessen these effects is to use combinations of measurements taken in concert, as done by Wilson and Russell (2003). They showed accuracies in excess of 85% classifying operator states in real-time using artificial neural networks trained on a battery of 43 physiological features. However, while psychophysiological measures have been used to adaptively allocate

automation functions in research environments, because of the previously discussed limitations, transferring experimental AA to operational AA has not yet been demonstrated.

Of all possible psychophysiological measures, analysis of brain waves has historically been the primary method of investigating neural indexes of cognition (Fabiani et al. 2000). Numerous studies have successfully used EEG measures of workload to discriminate between differences in operator attention and arousal (Berka et al. 2004; Kramer 1991; Pope et al. 1995). They all used engagement indexes based on the ratios of different EEG bands (alpha, beta, theta, etc), alpha suppression, or increased beta levels to detect changes in workload. However, as Prinzel et al. (2003) notes, EEG-based systems are only able to measure gross overall changes in arousal, not different types or finer levels of cognitive load. In contrast, the P300 component of ERPs has been associated with the availability of information processing resources, varying in amplitude as a function of primary task load, and its latency affected by stimulus evaluation time (Polich 1991; Rugg and Coles 1995). The P300 has been documented as an effective measure of mental workload (Donchin et al. 1986; Kramer 1991). Despite this, use of ERPs in non-laboratory settings to measure workload in real-time has proven difficult, as they are obtained from averaging of EEG signals over a number of trials. Methods for how to overcome this are under development, but Humphrey and Kramer (1994) were able to discriminate between different workload levels 90% of the time using only 1 to 11 seconds of ERP data (approximately 1-11 trials).

Eye movements have also been used extensively in a variety of studies. Their main value is in distraction studies to identify areas of attention, but measures of blink duration, frequency and pupil diameter have been correlated with visual and cognitive workloads. In general, eye blink duration and frequency decrease as both visual and/or cognitive workload increases (Hankins and Wilson 1998; Orden et al. 2001; Veltman and Gaillard 1998). Though pupil diameter is affected by ambient light levels, stimulus perception, and habituation, the degree of pupillary dilation has been shown to increase with higher cognitive loads (Andreassi 2000; Hess and Polt 1964; Orden et al. 2001). Specific characteristics of eye movements, such as fixation duration, dwell times and saccade durations are task-dependent and thus are extremely difficult to tie to changes in mental workload in a generalizable way. However, Simon et al. (1993) demonstrated that the more general visual behavior of scanning patterns becomes more organized when task difficulty increases.

Finally, AA may be based upon performance-based measures, whereupon some performance metric such as reaction time or task accuracy is used to determine mental workload. While generally easier to measure and quantify than physiological measures, performance measures are generally task-specific (not generalizable to other tasks) and often require the subjects to modify their natural task behavior to accommodate the experimental objectives. For instance, in monitoring tasks, there are few specific actions that may be linked to performance. Performance-based measures also may only give the experimenter discrete samplings of operator workload at specific intervals instead of a constant measurement. This could be inappropriate for some applications characterized by rapid changes in the environment, necessitating quick switches between automation modes.

It is clear that psychophysiological and performance-based measures complement each others' strengths and weaknesses. An experiment could be designed in which adaptive automation was driven by a combination of both, perhaps where the psychophysiological measure workload classification could be compared to one generated by the performance-based measure. This in turn would help provide a better estimate of operator workload so that the adaptive automation could more effectively support transient information overload in NCW.

4. DISTRIBUTED DECISION-MAKING AND TEAM COORDINATION

In 1994, Operation Provide Comfort was to provide humanitarian aid to over one million Kurdish refugees in northern Iraq in the wake of the first Gulf War. As part of this, the US sought to stop Iraqi attacks on the Kurds by enforcing a no-fly zone. The no-fly zone was patrolled by USAF fighters (F-15s), supported by airborne warning and control (AWAC) aircraft. On April 14, 1994, two US Army Black Hawk helicopters were transporting U.S., French, British, and Turkish commanders, as well as Kurdish para-military personnel across this zone when two US F-15 fighters shot them down, killing all 26 on board. The Black Hawks had previously contacted and received permission from the AWACs to enter the no-fly zone. Yet despite this, AWACs confirmed that there should be no flights in the area when the F-15s misidentified the US helicopters as Iraqi Hind helicopters. The teamwork displayed in this situation was a significant contributing factor to the friendly fire incident, as the F-15s never learned from AWACs that a friendly mission was supposed to be in the area. It was later determined that the F-15 wingman backed up the other F-15's decision that the targets were Iraqi forces despite being unsure, which was yet another breakdown in communication. Each team member did not share information effectively, resulting in the distributed decision making of the AWACs and F-15s pilots to come to incorrect and fatal conclusions.

Military forces in the 21st century are being thrust into a wider range of non-traditional roles, and so face more complex threats that often cannot be defeated by conventional tactics. Thus, it is critical that the military be able to leverage all of its available information, and to have sufficient agility to apply the relevant resources to bear on emerging situations. This is the driving force behind the US military's transformation into the Information Age and NCW. Unfortunately, the underlying traditional Industrial Age command and control processes of decomposition, specialization, hierarchy, optimization, deconfliction, centralized planning and decentralized execution (Alberts and Hayes 2003) are often at odds with this vision.

In many ways, platform-centric command and control (C2) in the past was about avoiding distributed decision-making and minimizing team coordination. Many decisions were made by a select few at the top level of command, and pains were taken to decompose various factions of the military into small, specialized niches that had little direct contact between one another (a hierarchical waterfall approach). This has begun to change in recent times, and a fully realized vision of NCW will require that teams of people, not a few people at the top of a hierarchy, make decisions under time-pressure. Therefore, understanding the issues unique to team-based coordination take on new importance in the context of NCW. The question is how to make effective decisions between and within distributed teams, particularly in the complex, data-rich, and time-compressed situations often seen in military C2/NCW scenarios.

Table 3. Team SA Requirements for Shared Information (Endsley and Jones 2001)

<p>Level 1 SA - Data</p> <ul style="list-style-type: none"> System Environment Other team members
<p>Level 2 SA - Comprehension</p> <ul style="list-style-type: none"> Status relevant to own goals/requirements Status relevant to other's goals/requirements Effect of own actions/changes on others Effect of other's actions on self and overall goal
<p>Level 3 SA - Projection</p> <ul style="list-style-type: none"> Actions of team members

A fundamental building block of good decision making is a high level of situation awareness (SA) (Endsley 1995), and in the case of distributed decision-making a high level of team SA or shared situation awareness (SSA). Three levels of individual SA have been defined: 1) perception of important environmental cues, 2) comprehension of the situation, and 3) projection of future events and dynamics (Endsley 1995). Team SA involves individual SA for each team member, plus the SA required for overlapping tasks and team interactions (Endsley 1995). Endsley and Jones (2001) expanded upon this definition by outlining a layered model of how teams achieve high levels of team SA. They detail what constitutes SA requirements in a team setting at each previously defined level (Table 3), the devices and mechanisms used to achieve shared SA, and SA processes that effective teams use. Team SA devices include spoken and non-verbal communications, visual and audio shared displays, and a shared environment.

Research could examine how individual and team SA can be enhanced through manipulation of these SA devices, mechanisms and processes in the context of information sharing in NCW tasks. Areas of particular relevance to NCW along these lines would be: data presentation, training, role allocation, team geographic distributions, communication, and team sizes.

Distributed decision-making and team coordination would also benefit from investigations into organizational design, as applied to team selection for NCW. Processing capabilities of humans are limited, so distribution of information, resources and tasks among decision-makers must be done in such a way that the cognitive load of each person is not exceeded. Research has shown that organizations operate most efficiently when their structures and processes match their mission environments (Levchuk et al. 2002). The traditional hierarchies omnipresent in the military today were designed for Industrial Age C2 processes and therefore will have reduced effectiveness when operating under NCW principles. Alberts and Hayes (2003) offer one solution with their proposed 'edge organizations', which are characterized by widespread information sharing and peer-to-peer relationships, but their design offers very few details on how to actually implement them. However, while team-based research is a high priority issue to examine in NCW, it is in general a lengthy proposition. Meaningful work is unlikely to be accomplished in the space of less than a year.

5. COMPLEXITY MEASURES

On July 3rd, 1998, 290 passengers and crew took off from Bandar Abbas airport in Iran on Iran Air Flight 655, bound for Dubai in the United Arab Emirates. Tragically, it never made it as the Aegis class cruiser USS Vincennes shot down the flight over the Strait of Hormuz, killing all on board. While many factors contributed to this accident, such as the tense atmosphere in the Gulf at that time due to the Iran-Iraq war, the root cause can be attributed to the complexity associated with USS Vincennes' advanced tracking radar. It was designed in the 1980s for open water battles with the Soviet Navy, and as such, was capable of tracking hundreds of missiles and airplanes simultaneously. Efforts that might have mitigated the high level of complexity of this system were not undertaken in favor of improving the detection capabilities of the system. Two pieces of wrongly interpreted data resulting from the complexity of the HMI interface caused the ship's commander to make the erroneous decision to fire on flight 655. First, flight 655 was reported as decreasing in altitude when it was in fact doing the opposite, representing that it was not in an attack profile. Second, the flight's Identification Friend or Foe (IFF) signal, designed to differentiate between civilian and military aircraft, was misidentified as being Mode II (military) instead of Mode III (civilian).

Information complexity is a growing problem in many domains, with particular applicability to NCW. Complexity will be impacted by both the amount and sources of information, and will be further exacerbated in the future as sensor technologies improve and the volume of available data continues to grow. NCW operators will be required to understand resultant critical relationships and behaviors of that data at the same or higher level than today. As put by Miller (2000), the impacts of increased complexity on the human will usually be increased workload and/or unpredictability of the

system, both of which have a negative effect on human and system performance. Greater system unpredictability is essentially equivalent to a reduced level of operator SA, and therefore to degraded ability to make informed, timely decisions. There are also well known limitations to human performance (see Section 1) that are likely to be encountered as workload increases. Therefore, it is important that the interfaces NCW operators interact with help to reduce and manage this increased level of data complexity.

How available information is presented to humans is a significant factor in mitigating information complexity. Properly designed human-machine interactive elements can ease workload by automating certain tasks such as data filtering, integration and decision support, but the opposite effect can occur in certain unexpected situations. One problem in mitigating information complexity is a lack of objective measures of complexity, for example, how a display supports knowledge-based reasoning. This work has been underway in air traffic control (ATC) for some time (see Majumdar and Ochieng (2002) or Mogford et al. (1995) for a review), but has yet to be explored in the NCW domain.

The first step to developing these objective measures of complexity is to define complexity as it relates to NCW processes. Complexity, as defined by Merriam-Webster, is “the quality or state of being hard to separate, analyze, or solve”. The use of the term ‘hard’ implies that complexity is relative, which was captured by Miller (2000) when he described the difference between actual and perceived complexity. Perceived complexity results from those elements of a task or situation that make it hard to deal with, or in other words, what makes a system seem difficult. Actual complexity implies the use of more objective criteria for complexity, and does not take into account humans’ perceptions of a task or situation. A lack of situation awareness can also unnecessarily complicate a person’s perception of a problem, increasing perceived complexity above actual complexity.

While perceived complexity and actual complexity are related, for human supervisory control interactions in NCW systems (and for most HCI applications) the designer should be concerned most about perceived complexity, and how to make it significantly lower than actual complexity without losing critical elements of the situation. For example, cognitive strategies such as grouping and rule formation can help to reduce perceived complexity. The primary characteristics underlying groupings in visual perception are proximity, similarity, common motion, symmetry, and good form. In addition, goal-oriented grouping that varies with the task at hand is common (Landry et al. 2001).

Perceived complexity can be divided into three general dimensions: 1) Component complexity: the number and diversity of components, 2) Relational complexity: the number and diversity of links between components, and 3) Behavioral complexity: the number and diversity of behaviors system components can exhibit (Miller 2000). Often, these dimensions are not independent and changes in complexity are driven by interactions between them. For example, a common task in the military is for an operator to track and interact with multiple entities. Examples of this type of action are in air defense warfare, where it could be an operator’s job to monitor and classify targets and tracks within a certain radius, or in land attack, where an operator could be in charge of targeting and re-tasking multiple missiles in flight. An increase in the number of targets would increase component complexity, but it is likely that the proximity of these new entities to the existing ones also would be important (an increase in relational complexity). It could also be possible that new tracks will exhibit some different behaviors than existing tracks (an increase in behavioral complexity), as would be the case with the entry of a new type of vehicle or missile in the previously mentioned scenarios.

One area of research could be evaluating how complexity varies in typical NCW tasks and how increasingly complex information affects human decision-making and performance. For example, control of four independent UAVs is likely to be less complex than control of four collaborative UAVs that have hidden states of knowledge and thus higher unpredictability. In this case, the

number of vehicles (components) is not driving complexity; instead the relations and behaviors between the UAVs are. Another significant area of exploration would be examining how receipt of different information types and combinations of types affects the perceived complexity of a situation. Information types may include probabilistic, discrete and continuous forms.

6. DECISION BIASES

On December 20th, 1995, American Airlines flight 965 was enroute to Cali, Colombia from Miami, Florida. Extremely behind schedule, as it approached Cali, the flight crew accepted a modified, unplanned and unfamiliar arrival route. Trusting the flight management system (FMS) to automatically come up with the next correct waypoint after having only entered the first letter of it, the captain mistakenly chose the computer's first choice, an incorrect waypoint 132 miles northeast of Cali. This resulted in a wide turn to the east, whereas the flight crew knew that they were supposed to be on a straight approach. Recognizing something was wrong, but not exactly what, the flight crew took manual control, turning right again towards Cali. Unfortunately, the approach into Cali was surrounded by mountains and flight 965 was now sufficiently off course that they flew into the side of a mountain, killing all but 4 aboard the Boeing 757. A clear contributor to this accident was the automation bias displayed by both the captain and first officer. They continued to rely on FMS-assisted navigation even when it became confusing and cognitively demanding during a critical segment of flight.

A defining characteristic of NCW is the expected increased information-sharing tempo over platform-centric forces of the past, which will require rapid decision-making. Humans in general, and especially under time pressure, do not make decisions according to rational theories. Rather they act in a naturalistic decision-making (NDM) setting in which experience, intuition, and heuristics play a dominant role. While NDM strategies are generally effective, the resulting decreased time of the OODA (observe-orient-decide-action) loop for decision-making under time pressure can cause these decisions to be subject to bias (reference the *USS Vincennes* incident.) It is well known that under uncertainty, humans generally employ heuristics as part of naturalistic decision-making in order to reduce cognitive load (Tversky and Kahneman 1974; Wickens and Hollands 2000). Heuristics can be useful, powerful tools, but also can introduce bias in decision making, especially when coupled with large amounts of information time pressures. There are three main classes of heuristics, all of which apply to NCW and have their potential drawbacks:

- *Representative heuristic* – probabilities are evaluated by the degree to which A resembles B
 - Potential problems
 - Insensitivity to prior probabilities of outcomes, sample size, and prediction
 - Misconceptions of chance
 - Can provide the illusion of validity
 - Tendencies of regression to the mean
- *Anchoring heuristic* – an initial guess is proposed and adjusted based on new information
 - Potential problems
 - Insufficient adjustment
 - Biases in the evaluation of conjunctive and disjunctive events
- *Availability heuristic* – probabilities of events are judged based on recency, simplicity
 - Potential problems
 - Biases due to retrievability of instances and/or effectiveness of a search set
 - Illusory correlation

Human are also prone to other decision biases to include:

- *Confirmation bias* – the decision-maker seeks information about a previously formed hypothesis, ignoring or discounting cues to the contrary
- *Automation bias* – an extension of conformation bias, the tendency of humans to trust computer recommendations without seeking contrary information.
- *Overconfidence bias* – confidence exceeds prediction accuracy, thus the decision-makers stop seeking confirmation of their hypotheses
- *Assimilation bias* - occurs when a person who is presented with new information that contradicts with a preexisting mental model assimilates the new information to fit into that mental model.

One important aspect of human supervisory control in NCW is the exploitation of the benefits of NDM without bias interference. To this end, research could explore what types of interface designs are best for this purpose, which types of biases are exhibited most strongly during each of the specific types of tasks detailed earlier, and how the introduction of automation that helps objectively assess probabilities changes operator decisions (one form of this is an optimal Bayesian agent).

7. ATTENTION ALLOCATION

A growing problem on roads today is driver distraction due to cognitive involvement in cell phone conversations. Although opinions vary on the severity of the problem, a recent study estimated a rate of 2,600 deaths a year in crashes on US roads due to cell phone use in 2002, as opposed to only 1,000 in 2000 (Coben and Graham 2003), identifying a disturbing trend. Coben and Graham also estimated that 570,000 injuries a year and 1.5 million crashes resulting in property damage can be blamed on wireless phone use. The cause of these accidents is inappropriate attention to the primary driving task, resulting in higher reaction times and lower situational awareness. In many cases, the cell phone conversation may become the driver's primary focus, instead of driving.

An important task in supervisory control is often one of how to allocate attention between a set of dynamic tasks. In deciding on an optimal allocation strategy, the operator acts to balance time constraints with relative importance of the required tasks. Due to the expected increases in the number of available information sources in NCW, volume of information and operational tempo (see sections 2, 3 and 6) will place greater attentional demands on operators. This is a fundamental and critical HSC problem in NCW. There are two general areas where attention allocation issues are likely to occur in NCW: Preview times/stopping rule generation and primary task disruption by secondary tasks.

Preview Times and Stopping Rules

A general situation illustrating attention allocation issues with preview times is one where an operator expects sensor information at established time intervals to accomplish some task, and must act on this information before a deadline. Entirely new tasks for the operator may also arrive at unexpected times. For example, an air defense warfare coordinator (AWC) on a Navy ship could be responsible for several tasks: Identifying unknown air tracks as friendly, enemy or commercial air, monitoring these identified tracks, providing warnings to enemy aircraft within a certain radius, and providing launch orders for additional defensive aircraft against encroaching enemies. Each of these tasks could involve numerous sub-tasks such as air traffic communications, visual confirmations, etc. Obviously some tasks are more important than others; shooting down threatening enemy aircraft is higher priority than tracking a commercial air flight. Enemy and commercial air flight launches or a sudden re-classification of a track could represent unpredictable increases to task load. Additionally, the AWC receives information updates only at discrete intervals as the radar sweeps by an area of

interest, or scheduled transmissions from equipment are received. Thus the AWC operator expects information to arrive in a certain time interval that could reduce uncertainty, but is sometimes faced with time-critical decisions that may or may not be able to wait for this information.

A central issue with the concept of preview times is how to maintain task priority when additional information is expected in the future, and how emergent situations influence an operator's ability to assimilate this preview information. Tulga and Sheridan (1980) investigated some aspects of this in a generic multi-task supervisory control paradigm. They found that at high workloads, the time subjects planned ahead was inversely proportional to the inter-arrival rate of new tasks. Using a similar paradigm, Moray et al. (1991) found that even if subjects were given an optimal scheduling rule, they were unable to implement it under enough time pressure, resorting instead to significantly non-optimal heuristic rules. In both experiments however, it was not possible to gain new information about specific tasks that would influence planning, nor were there unexpected events that significantly changed the nature of the task. It is clear from these initial efforts that more research is required to understand the effects of preview times, especially with information updates and unanticipated occurrences.

A related issue to preview times is that of stopping rule generation. Stopping rules are the criteria that individuals use to "satisfice" in uncertain situations, i.e. choosing the current best plan that is good enough. The general problem is as follows: Say that an operator has initial information, such as locations of friendly and enemy forces, and incoming information of various reliabilities and different times of arrivals, such as updates on enemy movements from sources such as voice communications and satellite images. The longer operators wait to make a decision on what to do with their forces, the more information they can gather (though not necessarily better due to its probabilistic nature), but they have a time limit in which to act. An individual's stopping rule would determine when the decision was made. Another interesting issue would be to observe how and if initial decisions were changed as more information was received and final time deadlines approached. A better understanding of the relationship between stopping rules and preview times in NCW is needed because by its very nature, NCW hinges upon successful information sharing. However, due to the stream of data from multiple sources and the need for rapid decisions, operators will have to weigh the benefits of gathering more information that will reduce uncertainty against the cost of a possibly delayed decision.

Primary Task Disruption by Secondary Tasks

Instead of a situation where an NCW operator has multiple dynamic tasks that vary in priority, consider a common scenario where the operator has a well-defined primary task along with secondary tasks. In time-pressure scenarios, interruptions of a primary task caused by interruption mechanisms such as secondary tasks can increase mental processing time and induce errors in the primary task (Cellier and Eyrolle 1992; Cummings 2004). In completing supervisory control tasks such as command and control or monitoring of displays, operators spend time monitoring unfolding events which may or may not be changing rapidly, and they also will periodically engage in interactive control tasks such as changing the course of UAVs or launching a missile. When task engagement occurs, operators must both concentrate attention on the primary task, but also be prepared for alerts for external events. This need to concentrate on a task, yet maintain a level of attention for alerts, causes operators to have a conflict in mental information processing. Concentration on a task requires "task-driven processing" which is likely to cause decreased sensitivity or attention to external events. Interrupt-driven processing, needed for monitoring alerts, occurs when people are sensitized and expecting distraction.

While interrupt and task driven processing can be present in a person, attention must be shared between the two and switching can incur cognitive costs that can potentially result in errors (Miyata

and Norman 1986). The conflict between focusing on tasks and switching attention to interruptions is a fundamental problem for operators attempting to supervise a complex system which requires dedicated attention but also requires operators to respond to secondary tasks, such as communications or alerts from non-critical sub-systems. In addition, Gopher et al. (1996) demonstrated that not only is there a measurable cost in response time and decision accuracy when switching attention between tasks, but costs are also incurred by the mere reconsideration of switching tasks. The potential cost of such induced errors in NCW supervisory control systems may be extremely high.

A more specific distraction problem likely to occur in NCW systems is that of operators becoming unintentionally fixated on secondary display information so that it becomes primary information. This means the operator misses critical data, leading to a degradation of SA and possibly lower overall performance. Many studies have investigated the effect of diverting attention to displays on the periphery (for examples see Maglio and Campbell (2000) or Somervell et al. (2001)), but very little work has determined how to design secondary displays so that they do not distract attention from the primary task. For the most part, secondary displays have been designed as standalone information visualizations. Somervell et al. (2002) begins to address this issue by identifying visual data density, presence time, and type of task the information is to be used for as influencing how secondary displays affect primary task performance. However, much work remains to be done in this area, particularly for NCW purposes.

8. SUPERVISORY MONITORING OF OPERATORS

October 29, 1998, two Boeing 737s were on standard air routes from Darwin to Adelaide and Ayers Rock to Sydney, respectively, at the same flight level of 37,000 feet. They were scheduled to conflict with each other, so protocol dictated that a 2,000 feet vertical separation standard be applied. This was noted by both the air traffic controller and a supervisor assigned to that particular sector 90 minutes before the conflict actually occurred, and marked for later action. In the next 90 minutes, traffic levels steadily increased and a third air traffic controller began to assist the other two already working in the conflict sector. The third controller assumed the coordinator position and attempted to deal with any items that seemingly required attention as he attempted to gain some idea of the traffic disposition. Despite the addition of a third coordinating ATC controller and the previous identification of the conflict, the pending conflict was subsequently overlooked. Instead, a violation of the minimum vertical separation distance occurred as one of the aircraft in question alerted ATC of the conflict at the last minute. This was an instance where the supervisor failed to detect a major failure of his supervisee, despite indications that one might occur.

A common operating structure in the military is one where a single supervisor oversees several human subordinates for the purpose of managing performance and relaying commands to the appropriate team members. Under information age C2 structures, the need for this second function will be reduced (even eliminated in some cases), but performance monitoring will still be required. Frequently, these operators will be engaged in HSC tasks, so it will be the job of a supervisor to observe and diagnose HSC issues in one or more teams.

HSC problems can sometimes be subtle in nature, and thus tend to be more difficult to detect than during many other types of operations. Most HSC tasks are primarily cognitive in nature, so the supervisor cannot easily infer accurate performance from physical actions of operators. Rather than being able to directly observe a task being completed by a human, the supervisor can only evaluate how an operator is interacting with automation that completes that same task, and once it is done, evaluate the results of that effort. Physical actions taken by operators are limited to activities like typing, button pushing, and body movements to position themselves for better screen viewing. Furthermore, the effects of operators' actions can occur in remote locations from both the supervisor and subordinates. This physical separation means that all people involved with the process

must form mental abstractions to envision a complete picture of the situation. Complicating this is that interaction is usually done through artifacts with inherent limitations, such as voice communication, data links, and 2-dimensional screens. While this is clearly a problem with individual operators (it is one of the primary considerations when designing automation of this type), it is an even larger one for supervisors, who must try to synthesize information from multiple operators at once. Furthermore, isolating a single cause for poor performance of an entire team can be difficult, especially in time-pressured environments characteristic of NCW environments. Lastly, decreases in performance may be the result of automation degradation and have nothing to do with the human. Supervisors may have difficulty separating the two.

The main problem is then how to support supervisors of HSC tasks so that they are better able to understand what their subordinates are doing. Many of the issues previously discussed in this report factor into this discussion. In order to quickly observe and diagnose HSC problems, supervisors must have a high level of SA, both for individuals and teams. Even more so than their subordinates, it is critical that HSC supervisors have a clear picture of the team's *overall* situation. The building block to achieving this superior level of SA is access to and absorption of all relevant data. Therefore, information overload will be a particularly acute problem, as a supervisor could be responsible for any or all of the information available to their numerous subordinates. Additionally, due to the greater range of information types received by HSC supervisors as compared to a single operator, the number of possible relationships and behaviors of this data is higher. This means that the complexity of the situation for the supervisor is raised along all three dimensions simultaneously.

Another issue with supervisory monitoring is how to rectify HSC problems once they are detected. There are several options, which may be applied singly or in concert to varying degrees:

1) *Provide a warning to the human whose performance is low at the time.*

It may be the case that operators do not realize that their performance has dropped below acceptable levels, and merely need to be reminded of it and/or motivated to get back on track. An operator's attention may be inappropriately allocated, so a warning provided by the supervisor (who is monitoring their performance) could cue them to correct it. Of course, if an operator is already cognitively overloaded then the warning could have no effect, or even a negative one due to the additional distraction it would provide.

2) *Redistribute the task load between existing team members.*

Workload could be unevenly distributed within a team, or various team members could be more skilled at certain tasks than others at certain times, so dynamically shifting tasks would allow other team members to pick up the slack. If all team members are equally busy or if others lack the expertise needed to perform tasks, this redistribution will not work. Additionally, the complexity of dynamically allocating tasks between team members puts a significant cognitive load on the supervisor.

3) *Change the number of team members or teams on the underperforming task.*

Potential and existing team members or teams could be humans, computers, or a combination of both. It is also important to remember that poor performance can result from underload as well as overload. In this case, the supervisor would observe an unacceptable level of performance from a subordinate and initiate a process to either relieve that operator of the most problematic aspects of their tasks, or add to their workload, as required. This change could be manifested in one of two ways: 1) A changing of the level of automation experienced by that operator (driven by a supervisor), or 2) The addition or

subtraction of a human or computer team member. Changing the number of team members requires planning, as the new team members must have access to the correct equipment, programming and training to be effective. As before, there must be an efficient way to reassign jobs between team members. In addition, while changing the number of team members may help in the longer term to reduce overall workload and/or improve performance, there will be a transition period with associated costs for team and individual SA as well as individual and team performance.

4) *Modify the mission objectives or timeline to accommodate lowered overall performance.*

This is a straightforward solution, but not one that will always be available in military situations. Many missions have time-sensitive deadlines that cannot be altered. Similarly, lower level mission objectives may be part of a larger operation, and therefore are not flexible.

Finally, the question of whether the supervisor should be a human or a computer should be discussed. Using a computer offers advantages and disadvantages in this regard. A computer would eliminate information capacity issues and the need for refined HCI designs. However, it would operate on defined rules and would be relatively inflexible in the face of unknown situations. The computer would also lack the capability of a human to predict future performance based on subjective judgments from visual observations and face-to-face interactions.

REFERENCES

- Alberts, D. S., Garstka, J. J., and Stein, F. P. (2000). *Network Centric Warfare: Developing and Leveraging Information Superiority*, Command and Control Research Program (CCRP), Washington, DC.
- Alberts, D. S., and Hayes, R. E. (2003). *Power to the Edge*, Command and Control Research Program (CCRP), Washington, DC.
- Andreassi, J. L. (2000). *Psychophysiology: Human Behavior & Physiological Response*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Berka, C., Levendowski, D. J., Cvetinovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., and Olmstead, R. (2004). "Real-Time Analysis of EEG Indexes of Alertness, Cognition, and Memory Acquired With a Wireless EEG Headset." *International Journal of Human-Computer Interaction*, 17(2), 151-170.
- Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centred Approach*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Caterinicchia, D. (2003). "DoD Chat Use Exploded in Iraq." Federal Computer Week.
- Cellier, J., and Eyrolle, H. (1992). "Interference Between Switched Tasks." *Ergonomics*, 35, 25-36.
- Cohen, J. T., and Graham, J. D. (2003). "A Revised Economic Analysis of Restrictions on the Use of Cell Phones While Driving." *Risk Analysis*, 23(1), 5-17.
- Cummings, M. L. "The Need for Command and Control Instant Message Adaptive Interfaces: Lessons Learned from Tactical Tomahawk Human-in-the-Loop Simulations." *Adaptive Displays Conference*, Los Angeles, CA.
- DoD, C. I. O. (2001). "Network Centric Warfare: Department of Defense Report to Congress." Office of the Secretary of Defense, Washington, D.C.
- Donchin, E., Kramer, A. F., and Wickens, C. D. (1986). "Applications of Brain Event-Related Potentials to Problems in Engineering Psychology." *Psychophysiology: Systems, Processes, and Applications*, S. Porges, ed., Till & Till, Middletown, NJ, 702-718.
- Endsley, M. R. "Design and Evaluation for Situation Awareness Enhancement." *Proceedings of the the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA, 97-101.
- Endsley, M. R. (1995). "Toward a Theory of Situation Awareness in Dynamic Systems." *Human Factors*, 37(1), 32-64.
- Endsley, M. R., and Jones, W. M. (2001). "A Model of Inter- and Intra-team Situation Awareness: Implications for Design, Training, and Measurement." *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*, M. R. Endsley, ed., Human Factors and Ergonomics Society (HFES), Santa Monica, CA.
- Fabiani, M., Gratton, G., and Coles, M. G. H. (2000). "Event-Related Brain Potentials." *Handbook of Psychophysiology*, G. G. Berntson, ed., Cambridge University Press, Cambridge, England.
- Gopher, D., Greenspan, Y., and Armony, L. "Switching Attention Between Tasks: Exploration of the Components of Executive Control and their Development with Training." *Human Factors and Ergonomics Society 40th Annual Meeting*, Philadelphia, PA.
- Hankins, T. C., and Wilson, G. F. (1998). "A Comparison of Heart Rate, Eye Activity, EEG and Subjective Measures of Pilot Mental Workload During Flight." *Aviation, Space and Environmental Medicine*, 69(4), 360-367.
- Hess, E. H., and Polt, J. M. (1964). "Pupil Size in Relation to Mental Activity During Simple Problem Solving." *Science*, 132, 349-350.
- Hilburn, B., Jorna, P. G., Byrne, E. A., and Parasuraman, R. (1997). "The Effect of Adaptive Air Traffic Control (ATC) Decision Aiding on Controller Mental Workload." *Human-automation Interaction: Research and Practice*, Lawrence Erlbaum, Mahwah, NJ, 84-91.
- Humphrey, D. G., and Kramer, A. F. (1994). "Toward a Psychophysiological Assessment of Dynamic Changes in Mental Workload." *Human Factors*, 36(1), 3-26.
- Jones, D. G. (2000). "Subjective Measures of Situation Awareness." *Situation Awareness Analysis and Measurement*, D. J. Garland, ed., Lawrence Erlbaum Associates, Mahwah, NJ, 113-128.
- Kaber, D. B., and Endsley, M. R. (2004). "The Effects of Level of Automation and Adaptive Automation on Human Performance, Situation Awareness and Workload in a Dynamic Control Task." *Theoretical Issues in Ergonomics Science*, 5(2), 113-153.
- Kaber, D. B., Endsley, M. R., and Onal, E. (2000). "Design of Automation for Telerobots and the Effect on Performance, Operator Situation Awareness and Subjective Workload." *Human Factors and Ergonomics in Manufacturing*, 10(4), 409-430.

- Kramer, A. F. (1991). "Physiological Metrics of Mental Workload: A Review of Recent Progress." Multiple Task Performance, D. L. Damos, ed., Taylor & Francis, Washington, DC, 279-328.
- Landry, S. J., Sheridan, T. B., and Yufik, Y. M. (2001). "A Methodology for Studying Cognitive Groupings in a Target-Tracking Task." *IEEE Transactions on Intelligent Transportation Systems*, 2(2), 92-100.
- Laughery, K. R., and Corker, K. (1997). "Computer Modeling and Simulation of Human/System Performance." Handbook of Human Factors and Ergonomics, G. Salvendy, ed., John Wiley & Sons, New York, NY, 1375-1408.
- Levchuk, G. M., Levchuk, Y. N., Luo, J., Pattipati, K. R., and Kleinman, D. L. (2002). "Normative Design of Organizations - Part I: Mission Planning." *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 32(3), 346-359.
- Madni, A. M. (2000). "Thriving on Change Through Process Support: The Evolution of the ProcessEdge™ Enterprise Suite and TeamEdge™." *International Journal of Information, Knowledge, Systems Management*, 2(1), 7-32.
- Madni, A. M., and Madni, C. C. "Context-Driven Collaboration During Mobile C2 Operations." *The Society for Modeling and Simulation International 2004 Western Simulation Multiconference*, San Diego, CA.
- Maglio, P. P., and Campbell, C. S. "Tradeoffs in Displaying Peripheral Information." *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2000)*.
- Majumdar, A., and Ochieng, W. Y. "The Factors Affecting Air Traffic Controller Workload: A Multivariate Analysis Based Upon Simulation Modeling of Controller Workload." *81st Annual Meeting of the Transportation Research Board*, Washington, DC.
- Miller, C. (2000). "The Human Factor in Complexity." Automation, Control and Complexity: New Developments and Directions, J. Weyrauch, ed., John Wiley, New York, NY.
- Miller, C. A., and Parasuraman, R. "Beyond Levels of Automation: An Architecture for More Flexible Human-Automation Collaboration." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Denver, CO.
- Miyata, Y., and Norman, D. A. (1986). "Psychological Issues in Support of Multiple Activities." User Centered System Design: New Perspectives on Human Computer Interaction, S. W. Draper, ed., Lawrence Erlbaum, Hillsdale, NJ, 265-284.
- Mogford, R., Guttman, J. A., Morrow, S. L., and Kopardekar, P. (1995). "The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature." *DOT/FAA/CT-95/22*, U.S. Department of Transportation, Atlantic City, NJ.
- Moray, N., Dessouky, M. I., and Kijowski, B. A. (1991). "Strategic Behavior, Workload, and Performance in Task Scheduling." *Human Factors*, 33(6), 607-629.
- Muir, B. M. (1987). "Trust Between Humans and Machines, and the Design of Decision Aids." *International Journal of Man-Machine Studies*, 27(5&6), 527-539.
- Orden, K. F., Limbert, W., Makeig, S., and Jung, T.-P. (2001). "Eye Activity Correlates of Workload during a Visuospatial Memory Task." *Human Factors*, 43(1), 111-121.
- Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., and Barnes, M. (1992). "Theory and Design of Adaptive Automation in Adaptive Systems." *Progress Report No. NAWCADWAR-92033-60*, Naval Air Warfare Center, Aircraft Division, Warminster, PA.
- Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3), 286-297.
- Pew, R. R. (2000). "The State of Situation Awareness Measurement: Heading Toward the Next Century." Situation Awareness Analysis and Measurement, D. J. Garland, ed., Lawrence Erlbaum Associates, Mahwah, NJ, 33-47.
- Polich, J. (1991). "P300 in Clinical Applications: Meaning, Method, and Measurement." *American Journal of EEG Technology*, 31, 201-231.
- Pope, A. T., Bogart, E. H., and Bartolome, D. (1995). "Biocybernetic System Evaluates Indices of Operator Engagement." *Biological Psychology*, 40, 187-196.
- Prinzel, L. J., III, Freeman, F. G., Scerbo, M. W., Mikulka, P. J., and Pope, A. T. (2003). "Effects of a Psychophysiological System for Adaptive Automation on Performance, Workload, and the Event-Related Potential P300 Component." *Human Factors*, 45(4), 601-613.
- Rugg, M. D., and Coles, M. G. H. (1995). "Electrophysiology of Mind: Event-Related Brain Potentials and Cognition." Oxford Psychology Series, Oxford University Press, New York, NY.
- Shannon, C. E. (1948). "A Mathematical Theory of Communication." *The Bell System Technical Journal*, 27, 379-423.
- Shapiro, C., and Varian, H. R. (1999). *Information Rules: A Strategic Guide to the Network Economy*, Harvard Business School Press, Boston, MA.
- Sheridan, T. B. (1992). *Telerobotics, Automation and Human Supervisory Control*, The MIT Press, Cambridge, MA.

- Sheridan, T. B. (2002). *Humans and Automation: System Design and Research Issues*, John Wiley & Sons, Inc., Santa Monica, CA.
- Sheridan, T. B., and Verplank, W. L. (1978). "Human and Computer Control of Undersea Teleoperators." MIT, Cambridge.
- Simon, P., Rousseau, F., and Angue, J.-C. "Quantitative Analysis of Mental Workload Influence on Eye Scanning Movements." *IEEE International Conference on Systems, Man and Cybernetics*, Le Touquet, France, 707-712.
- Somervell, J., McCrickard, D. S., North, C., and Shukla, M. "An Evaluation of Information Visualization in Attention-Limited Environments." *Joint Eurographics - IEEE TCVG Symposium on Visualization*, Barcelona, Spain, 211-216.
- Somervell, J., Srinivasan, R., Vasnaik, O., and Woods, K. "Measuring Distraction and Awareness Caused by Graphical and Textual Displays in the Periphery." *Proceedings of the 39th Annual ACM Southeast Conference*, Athens, GA.
- Tulga, M. K., and Sheridan, T. B. (1980). "Dynamic Decisions and Work Load in Multitask Supervisory Control." *IEEE Transactions on Systems, Man, and Cybernetics*, 10(5), 217-232.
- Tversky, A., and Kahneman, D. (1974). "Judgement Under Uncertainty: Heuristics and Biases." *Science*, 185, 1124-1131.
- Veltman, J. A., and Gaillard, A. W. K. (1998). "Physiological Workload Reactions to Increasing Levels of Task Difficulty." *Ergonomics*, 41(5), 656-669.
- Wickens, C. D., and Hollands, J. G. (2000). *Engineering Psychology and Human Performance*, Prentice Hall, Upper Saddle River, N.J.
- Wilson, G. F., and Russell, C. A. (2003). "Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks." *Human Factors*, 45(4), 635-643.
- Yerkes, R. M., and Dodson, J. D. (1908). "The Relation of Strength of Stimulus to Rapidity of Habit-Formation." *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Appendix A

MIT Candidate NCW Experimental Platforms

Multi-Modal Workstation (MMWS)

The Multi-Modal Workstation (Figure 7), developed by ONR/SPAWAR, is a platform available for research in the MIT Humans and Automation Lab, directed by Professor Cummings. The computer is a dual-processor 450MHz Pentium III machine that drives 4 Elo-Touch touch screen monitors, three 20" in size and one 15". Additional user interfaces include a wireless keyboard, trackball mouse and numeric keypad.



Figure 7. The Multi-Modal Workstation, as set-up at MIT

The MMWS is a versatile machine that could host a number of different experiments, each in some way exploring some aspects of the HSC issues described in the main body of the report. Two propriety applications, developed by SPAWAR for their experiments, are available. The first is called Air Defense Warfare (ADW), and involves managing all tasks associated with tracking, identifying and tasking all classes of aircraft within a particular radius of a US Navy fleet. The second application is called Land Attack Warfare (LAW). This program requires a user to coordinate all of the logistics of planning and re-planning, as well as monitor a battery of Tomahawk missiles launched from a US Navy vessel to numerous land attack targets. Either one or both applications may be used for NCW investigations.

Intelligent Systems Technology (ISTI) ProcessEdge™ and Knowport™ Software

The MIT Humans and Automation Lab has access to two ISTI applications that could be used in NCW experiments - ProcessEdge™ and Knowport™. ProcessEdge™ is an organizational design tool meant to facilitate process design, re-invention, analysis and visualization (Madni 2000). Through this software, research into optimal organizational structures to improve distributed decision making and team coordination could be performed. ProcessEdge™ is a standalone software package, making it relatively easy to set up and cheap to buy.

Knowport™ is a mobile platform application offering common knowledge access and decision support to distributed team members. It requires a server-client architecture be set up, in which the server contains a central data library accessed by smart software agents. The clients are mobile PDAs, and would be carried around in the field by soldiers and other military personnel. For experimental purposes, the MMWS could be used as the central depository of knowledge and team coordinator for the system. As should be clear from this brief description, Knowport™ will be relatively difficult to work with and expensive to buy, but the range of experimental possibilities for the purposes of NCW are exponentially greater than ProcessEdge™.