Reviews of Human Factors and Ergonomics, Volume 6

CHAPTER 2

Human Supervisory Control Challenges in Network Centric Operations

Mary L. (Missy) Cummings, Sylvain Bruni & Paul J. Mitchell

Running head: Human Supervisory Control

Key words: automation, unmanned vehicles, decision making, trust, attention, human performance, accountability, command and control

About the authors:

Mary L. (Missy) Cummings is an associate professor in the MIT Aeronautics and Astronautics Department, and is the Director of the Humans and Automation Laboratory. She holds joint appointments with the MIT Engineering Systems Division and the Computer Science and Artificial Intelligence Laboratory. Her research focuses on human supervisory control and human interaction with complex, autonomous systems, and she has published over 100 technical peer-reviewed papers and book chapters on these topics. Her current projects include human-unmanned vehicle interaction, bounded collaborative human-computer decision making, direct perception decision support, information complexity in displays, design of highly autonomous command centers, and the ethical and social impact of technology. She currently teaches an undergraduate course in Human Factors Engineering, and graduate classes in Human Supervisory Control and Design of Experiments. Professor Cummings received her B.S. in Mathematics from the United States Naval Academy in 1988, her M.S. in Space Systems Engineering from the Naval Postgraduate School in 1994, and her Ph.D. in Systems Engineering from the University of Virginia in 2003. A naval officer and military pilot from 1988-1999, she was one of the Navy's first female fighter pilots.

Sylvain Bruni is a Human Systems Engineer at Aptima, Inc., where he provides expertise in human-automation interaction, interface design, and the statistical design of experiments. His research focuses on the design and evaluation of multi-agent collaborative decision-support technologies, specifically in aeronautics, aerospace, and military environments. His work further includes the development of methodologies for the detection and classification of human operators' cognitive strategies as a means to assess system design and performance. Mr. Bruni's background also includes investigating the effects of information display and levels of automation on system performance, as well as defining evaluation tools for complex human-computer systems. Mr. Bruni holds a S.M. in Aeronautics and Astronautics from the Massachusetts Institute of Technology (MIT) and a Diplôme d'Ingénieur from the Ecole Supérieure d'Electricité (Supélec, France). He is currently a doctoral candidate in Aeronautics and Astronautics at MIT's Humans and Automation Laboratory. He is a member of the Human Factors and Ergonomics Society, the IEEE Systems, Man, and Cybernetics Society, the Association for Computing Machinery, the Usability Professionals Association, the American Society of Naval Engineers, the American Institute of Aeronautics and Astronautics, the Society for Automotive Engineers, and the Department of Defense Human Factors Engineering Technical Advisory Group.

Paul J. Mitchell received his B.S. degree in mathematics and engineering from Queen's University, Kingston, ON, Canada, in 2002 and his M.S. degree in Aeronautics and Astronautics from the Massachusetts Institute of Technology, Cambridge, in 2005. He is currently a Control Development Engineer with General Electric Energy in Power Generation Technology, Fairfield, CT. His research and professional interests include human supervisory control, human–computer interfaces and decision support, particularly for uninhabited vehicles, and the application of advanced control methods to aerospace and power generation systems.

CHAPTER 2

Human Supervisory Control Challenges in Network Centric Operations

Mary L. (Missy) Cummings, Sylvain Bruni & Paul J. Mitchell

ABSTRACT

Network-centric operations (NCO), envisioned for future command and control systems in military and civilian settings, must be supported by sophisticated automated systems, so human-computer interactions are an important aspect of overall system performance. This chapter identifies ten human supervisory control challenges that could significantly impact operator performance in NCO: information overload, attention allocation, decision biases, supervisory monitoring of operators, distributed decision making through team coordination, trust and reliability, the role of automation, adaptive automation, multimodal technologies, and accountability. Network-centric operations will bring increases in the number of information sources, volume of information, and operational tempo with significant uncertainty, all of which will place higher cognitive demands on operators. Thus, it is critical that NCO research focuses not only on technological innovations, but also on the strengths and limitations of human-automation interaction in a complex system.

In command and control settings, network-centric operations (NCO) is the concept of operations envisioned to increase combat power by effectively linking or networking knowledgeable entities in a battlespace. Mission success is achieved by leveraging information superiority through a network, rather than the traditional method of sheer numerical superiority through platforms and weapons. According to the United States Department of Defense (DoD), key components of NCO include information sharing and collaboration, which will promote shared situational awareness and overall mission success (DoD, 2001). To realize NCO, significant improvements must be made in areas of communications, sensor design, and intelligent automation. More importantly, whereas technological advances are important for realizing the concept of network centric operations, equally if not more critical is the need to understand how, when, where, and why the technology supports human decision makers and front line soldiers.

Command and control domains are complex socio-technical domains in that the technology is a means to an end (goal or mission), defined by human intentions. Thus, NCO systems are inherently socio-technical systems characterized by high complexity, high risk, time pressure, and dynamic goals. These same characteristics are hallmarks of naturalistic decision making, so any NCO technologies that are not designed with the express purpose of supporting military personnel decision making in these dynamic and uncertain situations with rapidly shifting goals are likely to fail. Even though this chapter will focus primarily on military NCO systems, since this is where the majority of research is primarily occurring, these results are extensible to all networked time-pressured systems such as first responder systems, real-time financial trading, and commercial air traffic control.

The move towards NCO represents a shift in the role of humans both in mission planning and actual operation. As has already been evidenced in the development of fly-by-wire controls, highly automated aircraft and missile systems, military operators are less in direct manual control of systems, but more involved in the higher levels of planning and decision making. One example of a military network application is the Tactical Tomahawk missile, which can be launched from a ship or submarine. Any time during its approximately two hour flight, control can be handed off to anyone in the distributed network (connected via satellite communications), and moreover, these missiles are typically fired in clusters so that one operator can theoretically control multiple missiles in flight (Cummings & Guerlain, 2007). An NCO example in the commercial realm includes the next-generation (Next-Gen) vision of the national aerospace system, which will include aircraft that can electronically deconflict with one another, while under positive air traffic control, as well as the presence of unmanned aerial vehicles (UAVs).

The shift in control from lower level skill-based behaviors to higher level knowledge-based behaviors, which is inherent to NCO, is known as human supervisory control (HSC). 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 2.1).

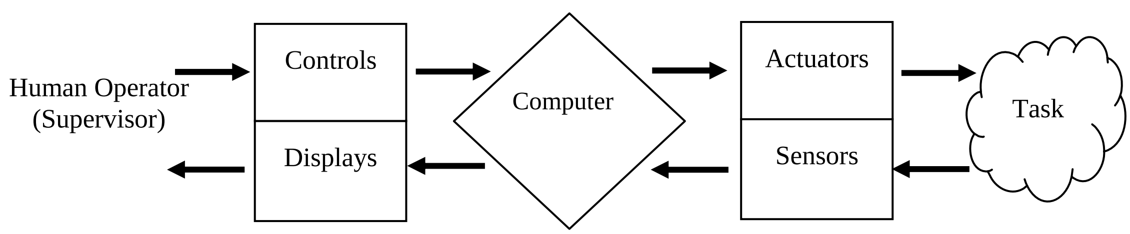

Figure 2.1. Human supervisory control (Sheridan, 1992).

HSC in military operations includes mission planning, passive and active intelligence operations, and payload delivery which could involve manned aircraft and ground vehicles, as well as unmanned air, ground, surface, and subsurface vehicles. The use of automated technologies and systems is a fundamental component of NCO; thus in the context of human interaction, NCO is a high-level human supervisory control problem. An example of a current NCO system is the use of multiple layers of unmanned and manned aircraft, as well as ground vehicles, to search for and destroy targets. It is not uncommon for the military to use unmanned aerial vehicles (UAVs) to search over long periods of time for hostile targets, and then once a target is found, coordinate through various agencies on the ground and in the air for threat neutralization. In some cases, these UAVs are operated from thousands of miles away (e.g., the Predator), so the DoD's reliance on networked agents who remotely coordinate is a reality.

These current NCO operations typically involve multiple personnel controlling a single UAV, who coordinate with various manned assets for mission success. Future NCO scenarios will invert this ratio, with small teams of people controlling an order of magnitude greater number of heterogeneous unmanned vehicles. For example, an operator aboard a Navy vessel may control a team of unmanned underwater vehicles (UUVs), as well as multiple UAVs engaged in coastal search and reconnaissance missions. This person will coordinate with Army personnel on land who supervise teams of unmanned ground vehicle (UGVs or robots) in minefield detection missions, all while coordinating with Air Force UAVs providing overhead security surveillance to ensure ground personnel safety.

As a consequence of this future vision of small teams of people controlling large teams of unmanned vehicles, there will be an exponential increase in incoming information. Given this influx of voluminous information from a variety of sensors (both human and automated), a particularly acute problem will be how to give operators enough information for well-informed decisions without reaching cognitive saturation. Moreover, there is little understanding of how functions should be allocated between humans and automation, what level of collaboration is needed, and how these functions can be supported with actual software and hardware (Sheridan & Parasuraman, 2006). These HSC NCO problems are further complicated by the dynamic, uncertain, and time-pressured elements typical of command and control environments. Due to the increasing importance of HSC in NCO, 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 the potential of NCO (DoD, 2001).

Using historical case studies, as well as previous and current research studies, ten major human supervisory control issues are identified in this chapter as those HSC issues that are likely to cause degraded performance for both the system and the operators/decision-makers in futuristic network centric operations. These issues, listed in Table 2.1, fall into two general categories, those that describe human performance and those that address HSC technologies critical to NCO operations. Within these two major categories, these issues are not mutually exclusive, are not rank ordered, and will often overlap theoretically as well as in actual design and testing. The tenth and final category, accountability, is set apart from the others since it represents the intersection of both humans and technology in a societal context, and thus is a meta-attribute. These ten challenges will be detailed in the following sections.

Table 2.1. Ten Supervisory Control Challenges in NCO

Human Performance
Information Overload
Attention Allocation
Decision Biases
Supervisory Monitoring of Operators
Distributed Decision Making through Team Coordination
Trust and Reliability
Technology
Role of Automation
Adaptive Automation
Multimodal Technologies
Accountability

THE HUMAN PERFORMANCE ASPECT OF SUPERVISORY CONTROL SYSTEMS

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 in the open position. 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 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 data and 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 incorrect mental model.

According to the DoD, the Global Information Grid (GIG), the actual information technology network that will link command and control agents, will be the enabling building block for NCO. 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 elsewhere (DoD, 2001). Metcalf's Law states that the usefulness, or utility, of a network equals the square of the number of users (Shapiro & Varian, 1999). For the purposes of this discussion, the term information also implicitly includes data which is the raw output of a human or machine sensor, whereas information is aggregated data that is combined in some way to relate to some goal-directed process (Pfautz, et al., 2006).

The GIG, and other such large networks, will give operators access to exponential amounts of information as compared to today's operations, and the information intake for the average NCO operator will be higher than ever before in the command and control environment. Even if the information complexity does not increase (which is unlikely), mental workload will increase accordingly. The problem is predicting when and how increased mental workload, as a result of information overload, will occur for a dynamic decision-making environment so that the amount of information any single person or group is required to process is manageable.

Predicting this point of saturation is difficult for operators in NCO systems because it is dependent on system-level attributes such as task, automation level, and operational tempo, as well as individual and team attributes such as training, experience, fatigue, etc. Modeling and simulation can aid in predicting where high mental workload points are likely occur, and there are a number of modeling techniques, including cognitive and psychophysiological (discussed in a subsequent section), that can be used to identify and predict high operator workload. However, we focus on the development of stochastic systems models that consider the performance of the human in conjunction with that of sensors, vehicles, and other sources of automation.

Identifying and predicting high workload as a result of information overload. Discussed in this section, modeling a network-centric environment with stochastic representations of both operators and autonomous processes can be used to predict operator mental workload or the impact of high workload on overall performance in supervisory control settings. Such models are *systems*-level models as opposed to the *information-processing* level models that are typically used in cognitive models. This distinction is important because although these two techniques can conceptually be used to predict the same outcome (occurrences of high mental workload that are associated with negative outcomes), they do so at two very different levels of granularity. Cognitive models account for low-level cognitive processes including working memory load, access to long-term memory, the use of contextual cues, etc. Stochastic operator-automation models assume an aggregation of these information processing functions with bands of variability and consider the performance of the human simultaneously with the performance of the relevant aspects of automated systems, which also has associated performance bounds. See Bryne and Pew (2009) for a more extensive discussion on human performance modeling.

One common task envisioned for an operator in a NCO environment is the control of multiple unmanned vehicles. A significant concern is that operators are limited by how much information they can effectively manage from multiple vehicles, ultimately limiting the number that they can control. For this single-operator multiple-vehicle problem, one stochastic supervisory control model that can be used to predict operator workload is a queuing-based model, as represented in Figure 2.2 (Cummings & Nehme, 2009). Such a model is a system-level model because it considers the attributes of both the human and the automated components, (i.e.,

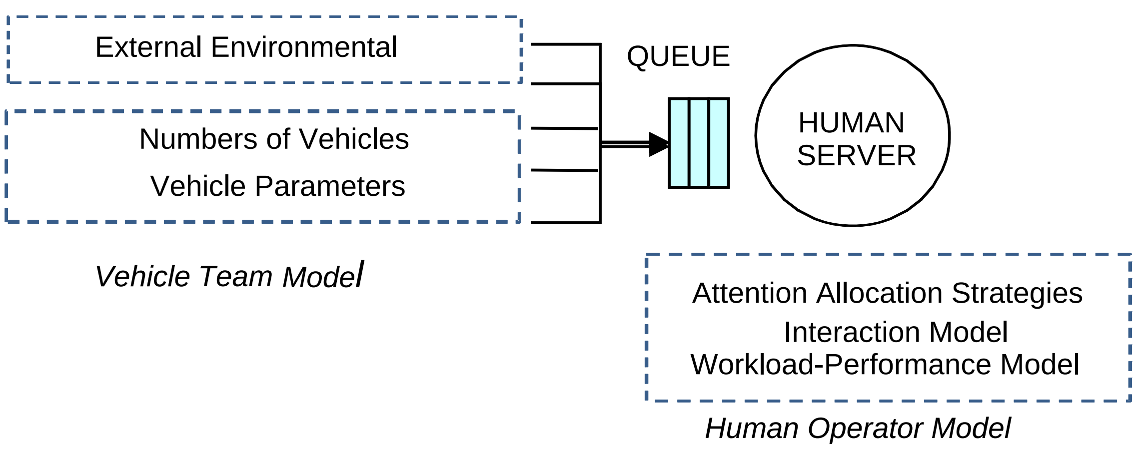

Figure 2.2. A queuing-based human-system performance model.

the vehicles), when attempting to make predictions about human or overall system performance.

The model in Figure 2.2 represents an event-based system, where events that are generated by vehicles and the human operator form a queue of operator tasks. For example, vehicles that require operators to input new waypoints, as well as more abstract tasks such as resource allocation of vehicles across a set of mission requirements, are considered events that require operator attention. The queuing model in Figure 2.2 assumes that the human "server" (the operator) is inherently limited in how fast tasks can be serviced, and that the operator is not always an efficient and correct server.

As seen in Figure 2.2, the model includes vehicle performance attributes (i.e., speed, range, etc.), as well as autonomy characterizations (how long and under what conditions vehicles can autonomously operate). Human models in these system account for operator strategies in managing multiple tasks (called attention allocation), how well they can do specific tasks (interaction models), and also, how performance can change as workload changes. This last workload relationship is typically represented as percent utilization (i.e., how busy an operator is over a given time period). In this modeling approach, the low-level cognitive interactions are subsumed by the interaction models and become just one of many variables in the overall system model.

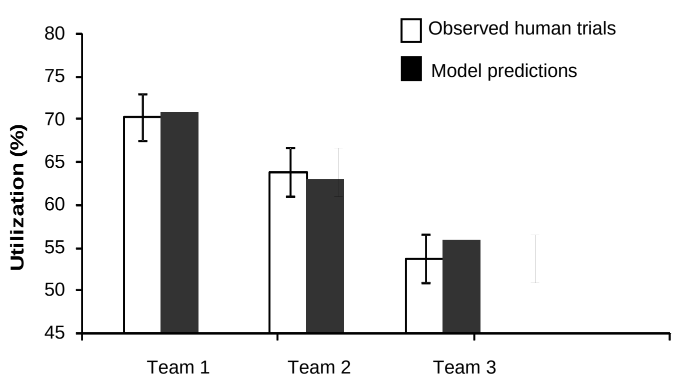

Figure 2.3. Stochastic model predictions of operator workload

Although queuing models have been used historically in supervisory control settings (Liu, 1997; Schmidt, 1978), the model in Figure 2.2 is unique in that it represents detailed models of both vehicles and human performance in a networked setting, so such a model can predict, for example, how many vehicles one person can control with an expected skill set and *a priori* vehicle parameters. It can also pinpoint at which utilization levels (i.e., mental workload), human *and* system performance begins to degrade, particularly as levels of incoming information increase. The model in Figure 2.2 is explicit for unmanned vehicle supervisory control, but it is generalizable to supervisory control settings with multiple complex tasks.

Figure 2.3 demonstrates this kind of modeling effectiveness, as it shows how well the model in Figure 2.2 can predict operator workload (as measured by percent utilization), given three different unmanned vehicle team architectures. The observed human performance measurements come from a multiple unmanned vehicle, single operator control simulation with 80 participants (Cummings & Nehme, 2009). Operators supervised multiple heterogeneous unmanned vehicles searching for targets, and then performed a visual identification task, similar to present-day intelligence, search, and reconnaissance (ISR) missions. All model predictions fell within the 95^{th} percent confidence intervals of the observed data, and it can be seen that team 1, the team that generated the most information for the operator to attend to, also required the most mental effort by operators. This approach to predicting operator workload as a function of unmanned vehicle team architectures is critical in allowing system designers to understand the implications on human performance for complex, networked systems.

There are other variations of stochastic supervisory control models that attempt to predict human and system performance through the incorporation of workload. Simulated annealing has been used with a similar embedded queuing model to predict performance of humans in concert with dynamic mission needs for networked unmanned aerial vehicles (Cummings, Nehme, & Crandall, 2007). Operator attention allocation strategies and interaction efficiencies have been modeled through stochastic representations in order to build models to predict overall human-unmanned vehicle team performance (Crandall & Cummings, 2007). In addition, other researchers have adapted the previously described operator utilization model in cooperative control models of multiple unmanned vehicles to improve overall model prediction for optimal vehicle coordination strategies (Savla, Nehme, Frazzoli, & Temple, 2008).

In summary, even though there is never one right modeling technique that is correct in all possible supervisory control settings, systems-theoretic models (i.e., those that consider all system entities in a holistic and integrated fashion), will be critical for building accurate models of network-centric environments. Moreover, because such environments contain significant uncertainty, both on the part of human operators as well as environmental events, stochastic modeling techniques that examine performance boundary limits in the presence of an optimization trade-space are likely to be more useful, especially given the large degrees of freedom inherent in NCO settings.

Mitigating information overload through design. The preceding discussion focused on the need for identification and prediction of high operator workload caused by information overload. Another equally relevant consideration is the need for better human-system interface designs that mitigate potential high workload situations. As will be discussed in a subsequent section, multimodal technologies are potential technological interventions to reduce workload and improve operator efficiency when attending to multiple sources of information. An additional design technique that can be used to mitigate the negative effects of information overload, particularly in the domain of visual interfaces which dominates NCO environments, is the application of direct-perception interaction displays.

Since perception is a direct, non-inferential process, given a well-designed interface that supports direct-perception interaction, operators can directly perceive the environment and what the various elements of the environment afford them, without needing to access costly cognitive resources (Gibson, 1979). Because direct perception does not require inference, it is fast, effortless, and proceeds in parallel unlike analytical cognition which is slow and error-prone (Vicente & Rasmussen, 1992). Moreover, it can be substituted for higher, more demanding cognitive tasks (Sanderson, Flach, Buttigieg, & Casey, 1989), and produces proficient performance from lower levels of cognitive control (Vicente & Rasmussen, 1992).

Numerous studies have shown that performance can improve when displays utilize direct perception-action visual representations that allow the user to employ the more efficient processes of perception rather than the cognitively demanding processes involved when relying on memory, integration, and inference (e.g., Bennett, 1992; Buttigieg & Sanderson, 1991). This kind of fast, low cognitive effort display aligns well with the dynamic, time-pressured NCO environments since the operator will need to take into account large amounts of information of varying degrees of uncertainty, but be expected to make high-consequence decisions accurately.

Direct-perception displays have been successfully used in NCO experimental settings such as providing decision support for multiple unmanned vehicle resource allocation (Cummings & Bruni, 2009), and tasking of multiple unmanned vehicles in real-time (Fisher, 2008). Although such tools are generally effective for improving operator performance both in terms of decision accuracy and time, previous research has demonstrated that the application of direct perception tools is not guaranteed to improve performance (Cummings, Brzezinski, & Lee, 2007). In the case of providing scheduling decision support for multiple UAVs, this research demonstrated that direct-perception visualization contributed to significantly degraded operator performance when applied using a globally optimal algorithm (across all vehicles), as compared to when the exact same visualization was powered by locally optimal algorithms (per vehicle). Thus, it is critical for designers to consider whether a direct-perception display (or any visualization designed to reduce workload) actually integrates information in a helpful manner, or adds to the workload by adding even more information for the operator to consider.

Attention Allocation

> *Instant messaging is rapidly becoming a primary means of communication between Navy ships and other operational units. Instant messaging, otherwise known as chat, has many advantages for rapid response in critical time-pressure command and control situations, but operational commanders have found it difficult to handle large amounts of information generated through chat, and then synthesize relevant knowledge from this information. 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" (Caterinicchia, 2003). 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. Laboratory research has shown that when given the opportunity to chat in NCO environments, operators will routinely neglect the primary task in favor of the secondary, lower priority task of chatting (Cummings & Guerlain, 2004). Designing systems to promote effective allocation of attention highlights a challenge for future network centric operations which inherently will require divided attention across often disparate tasks*

Due to the expected increases in the number of available information sources, volume of information, and operational tempo in NCO settings, greater attentional demands will be placed on operators. As a result, an important NCO issue is how to allocate attention across 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. There are three general attributes of attention allocation that need to be considered in NCO supervisory control systems: 1) task switching efficiency, i.e., selecting the right task at the right time to service, 2) minimizing switching costs, i.e., reducing the time it takes operators to reorient to the new situation, and 3) allocating just enough attention for the current task, i.e., once engaged in a task, how long should an operator spend gathering information on which to make a decision?

Task switching efficiency. The first attention allocation issue, task switching, is essentially a task prioritization problem with significant design implications. For example, an air traffic controller attends to several aircraft at once, particularly for enroute sectors. Controllers must be aware of which aircrafts require commands in both space and time. Procedures often aid operators in such supervisory control domains, such as “the aircraft that is closest to a sector edge should be serviced first.” However, there are also design strategies that can provide affordances (Norman, 1988), so that proper attention allocation is supported in a more intuitive manner.

For example, structure-based abstractions have been identified as key mental abstractions that help air traffic controllers effectively allocate their attention across multiple aircraft. These abstractions take on the form of standard flows (i.e., established traffic patterns), groupings (i.e., controlling groups of aircraft as one, effectively leveraging chunking strategies), and critical points (i.e., specific sector entry and exit points that all aircraft must pass through in three dimensions) (Histon, et al., 2002). To ensure that operators effectively attend to certain tasks at certain times, designing the airspace to leverage operators' use of these structure-based abstractions effectively promotes attention allocation.

Other design considerations for efficient and accurate task switching include designing effective automation cueing through warnings, alerts, critiques (Guerlain, 2000), and user-initiated notifications (Guerlain & Bullemer, 1996). Automation can be particularly useful for monitoring the large number of tasks that need attention; however, designers should also be mindful that alerts could represent interruptions to operators (Laughery & Wogalter, 2006; Trafton & Monk, 2007). Although such interruptions may be needed for high priority tasks that need immediate attention, overly salient alerts can be a nuisance and lead to inefficient attention allocation. Moreover, interruptions of a primary task by secondary task alerts can increase mental processing time and induce errors in the primary task (Cellier & Eyrolle, 1992).

For NCO supervisory control tasks that will likely include monitoring of displays, which may change rapidly, operators will periodically engage in interactive control tasks such as changing the course of UAVs or launching weapons. When task engagement occurs, operators must both concentrate attention on the primary task and be prepared for alerts to switch attention to 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 (Miyata & Norman, 1986).

While interrupt- and task-driven processing can occur simultaneously, attention must be shared between the two and switching can incur cognitive costs that can potentially result in errors (Miyata & 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. Moreover, often what appears to be an innocuous peripheral, secondary display feature such as scrolling of text in a chat window can have negative consequences because the distraction requires cognitive effort in considering whether or not it needs attention (Maglio & Campbell, 2000; Somervell, Srinivasan, Vasnaik, & Woods, 2001).

Switching times. The second area of concern in attention allocation for NCO time-pressured operators is minimizing the time needed to switch between tasks. Numerous studies have shown that there is a cognitive penalty when operators switch between tasks, i.e., there is a period of time where operators need to regain awareness of the nature of the task in order to develop a plan for resolution (Crandall, Goodrich, Olsen, & Nielsen, 2005). In addition, Gopher et al. (2000) 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 contemplation of switching tasks.

Interruptions inherently introduce additional switch time costs, and given the multi-tasking nature of NCO, this will likely be a significant future source of switching times. Altmann and Trafton (2002) have proposed that the interruption lag, the time between the interruption alert and the interruption event, is an intervention window that can be used to reduce task resumption time. However, the use of cues during the interruption lag has not been shown to decrease recovery time, and in one study, appeared to increase recovery time (Altmann & Trafton, 2004).

Increasing the role of automation, i.e., allowing a computer to determine appropriate task switch times is another possible intervention. Unfortunately, as will be described in detail in a subsequent section, increased automation often leads to degraded system performance as operators require more time to gain awareness of the need to switch tasks. This has been shown in a number of NCO laboratory experiments (Crandall, et al., 2005; Cummings & Mitchell, 2008; Squire, Trafton, & Parasuraman, 2006), so it is critical that designers understand the limits of automation and support mutual collaboration between human operators and automation. One example of such a collaborative tool is an intelligent aiding system that prioritizes incoming interruptions (McFarlane, 2002), allowing automation to assist operators in determining which task should receive attention, but ultimately giving operators final authority to determine which task should be serviced in the queue.

Allocating the right amount of attention. Attention allocation involves not only knowing when to switch attention to a relevant task, it also requires operators to be cognizant of how much attention they expect to allocate to a single task. In NCO, operators can expect sensor information at established time intervals to accomplish some task, and must act on this information, whether complete or not, before a deadline. 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. Some tasks are more important than others, for example, shooting down threatening enemy aircraft is higher priority than tracking a commercial air flight. The AWC receives information updates only at discrete intervals as the radar sweeps by an area of interest. Thus the AWC operator expects information to arrive in a certain time interval, called preview time, which could reduce uncertainty. However, time-critical decisions may have to be made without updated information.

A central issue with preview times is how to maintain task priority when additional information is expected in the future and how to assimilate this preview information during emergent situations. Tulga and Sheridan (1980) investigated these aspects in a generic multi-task supervisory control setting. 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 recent work examining the use of intelligent agent predictions for future periods of high workload in order to aid operators controlling multiple UAVs, results revealed that subjects fixated on attempts to globally optimize an uncertain future schedule to the detriment of solving certain, local problems (Cummings & Mitchell, 2008). Another study relative to UAVs demonstrated that efficiency of instrument information presentation significantly affected operator attention scan patterns (Tvaryanas, 2004). These initial efforts demonstrate 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 individuals use to “satisfice” in uncertain situations, i.e., choosing the current best plan that is good enough, but not necessarily best (Simon, et al., 1986). The general problem is as follows: 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 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 communications uncertainty), but they have a time limit in which to act. An individual's stopping rule would determine when the decision was made.

NCO 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. Automation may be able to assist operators in better stopping rule-generation, as well as attention allocation strategies across multiple tasks in general. However, as will be discussed in the section on the role of automation, introducing automation to reduce the workload of humans carries other costs that must be considered in the context of overall human-system performance.

Decision Biases

On July 3rd, 1988, 290 passengers and crew departed from Bandar Abbas airport in Iran on Iran Air Flight 655, bound for Dubai in the United Arab Emirates. Tragically, the Aegis class cruiser USS Vincennes shot down the flight over the Strait of Hormuz just minutes after takeoff, killing all on board. Many factors contributed to this accident, such as the tense atmosphere in the Gulf at that time due to the Iran-Iraq war, however, 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.

However, it was not meant to operate in cluttered littoral regions. Given confusing data presentation, two pieces of wrongly interpreted data 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. As a result, the flight was thought to be on 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).

A defining characteristic of NCO is the expected increased information-sharing tempo over platform-centric forces of the past, which will require rapid decision making with imperfect information. Humans in general, and especially under time pressure, do not make decisions according to rational decision theories. Rather, they act in a naturalistic decision-making (NDM) setting in which experience, intuition, and heuristics play a dominant role (Klein, 1989). Humans generally employ heuristics in order to reduce cognitive load (Tversky & Kahneman, 1974; Wickens & Hollands, 2000), which will likely be the case in NCO settings. Even though humans can be effective in naturalistic decision-making scenarios in which they leverage experience to solve real world ill-structured problems under stress (Zsambok, Beach, & Klein, 1992), they are prone to fallible heuristics and various decision biases that are heavily influenced by experience, framing of cues, and presentation of information. The *Vincennes* accident illustrates how the three general classes of heuristics (representativeness, anchoring, and availability), could be problematic in NCO environments.

In the representative heuristic, probabilities are evaluated by the degree to which an unfamiliar situation resembles a familiar one. This heuristic can provide the illusion of validity because decision makers are generally insensitive to prior probabilities. In the case of the

Vincennes, the operators ignored the more likely prior probability that the plane was a commercial airliner which traveled on a fairly regular schedule and instead believed the much less likely event that the plane was a military fighter. Misconception of chance is a classic sign of the representative heuristic (Tversky & Kahneman, 1974).

The anchoring heuristic, which occurs when an initial guess or hypothesis is not adjusted appropriately given new information, was also evident in the *Vincennes* incident. An initial hypothesis was made that the radar contact was an enemy fighter, and evidence was ignored that refuted this theory, including published commercial airline schedules and the climbing attitude of the plane, which would not be the case if the aircraft was attacking. Time pressure exacerbated this anchoring of the initial hypothesis, which will also likely be a significant problem in NCO.

The availability heuristic in which probabilities of events are judged based on recency and retrievability was also evident in the *Vincennes* accident. One year prior, the *USS Stark* had been hit by two Exocet missiles in the Persian Gulf, fired from a relatively low, fast flying Iraqi Mirage. In addition, on the morning of the shoot down, the *Vincennes* had engaged with Iranian gunboats so tensions both locally and globally were high. Given human memory limitations and the known propensity for humans to resort to often flawed applications of heuristics, a significant challenge in the design of NCO decision support systems will be how to give operators timely, probabilistic, and unbiased information.

These heuristics often lead to decision biases, such as confirmation bias, which takes place when people seek out information to confirm a prior belief and discount information that does not support this belief (Lord, Ross, & Lepper, 1979). Another decision bias, assimilation bias, occurs when a person who is presented with new information that contradicts a preexisting mental model, assimilates the new information to fit into the original mental model (Carroll &

Rosson, 1987). Evidence of these biases was present in the *Vincennes* case, as operators interpreted incoming information in light of their original hypotheses, and interpreted all new information through this filter.

Of particular concern in the design of intelligent decision support systems that will be fundamental to NCO processes is the human tendency toward automation bias, which occurs when a human decision maker disregards or does not search for contradictory information in light of a computer-generated solution which is accepted as correct (Mosier & Skitka, 1996; Parasuraman & Riley, 1997). Related issues of overconfidence and reliance will be discussed in a following section on trust. Human errors that result from automation bias can be further decomposed into errors of commission and omission. Automation bias errors of omission occur when humans fail to notice problems because the automation does not alert them, while errors of commission occur when humans erroneously follow automated directives or recommendations (Mosier & Skitka, 1996). Operators are likely to turn over decision processes to automation as much as possible due to a cognitive conservation phenomenon (Fiske & Taylor, 1991), and teams of people, as well as individuals, are susceptible to automation bias (Skitka & Mosier, 2000).

Many studies have demonstrated evidence of automation bias in laboratory settings. Layton, Smith, and McCoy (1994) examined commercial pilot interaction with automation in an enroute flight planning tool, and found that pilots, when given a computer-generated plan, exhibited significant automation over-reliance causing them to accept flight plans that were significantly sub-optimal. Skitka, Mosier, and Burdick (1999) found that when automated monitoring aids operated reliably, they led to improved human performance and fewer errors as opposed to not having an aid. However, when the automation failed to detect and notify operators of an event, or incorrectly recommended action despite the availability of reliable disconfirming evidence, human error rates increased significantly. More directly related to NCO processes, in a Tomahawk strike planning and execution task, Cummings (2006) found evidence of automation bias when an intelligent decision aid recommended a single course of action for retargeting missiles to emergent targets. Similar results were found for operator control of multiple unmanned aerial vehicles (Cummings & Mitchell, 2007), which resulted in operators dropping weapons on the wrong targets, reminiscent of the *Vincennes* incident.

As these cases demonstrate, heuristics and decision biases are a very real concern in the development of intelligent systems that provide decision support for humans, especially those in time-critical domains with significant uncertainty. Designers of intelligent systems should be aware of the potential negative effects on decision making in terms of inappropriate heuristics and biases as the human is further removed from the decision control loop. Intelligent decision aids are intended to reduce human error and workload but designers must be mindful that automated recommendations combined with unreliable systems can actually cause new errors in system operation if not designed with human cognitive limitations and biases in mind. Design of an intelligent system that provides decision support must consider the human not just as a peripheral device, but also as an integrated system component that in the end will ultimately determine the success or the failure of the system itself.

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 overall performance and providing an additional layer of safety. Frequently, these operators engage in HSC tasks, so it is the job of a supervisor to observe and diagnose HSC issues in one or more team members. Most HSC tasks are primarily cognitive in nature, so the supervisor cannot easily infer accurate performance from the physical actions of operators. In most cases, a supervisor can only evaluate the quality of operator interactions with an automated system from the results of those efforts.

In HSC settings, 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 the supervisor. 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. Clearly a problem for individual operators, 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 scenarios characteristic of NCO 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, then, is how to support supervisors of HSC tasks so that they are better able to understand how their subordinates' performance relates to overall mission success. To quickly observe and diagnose HSC problems, supervisors must have a high level of situation awareness (SA) for both 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 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 is higher, thus increasing situation complexity.

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

1) ***Provide a warning to the operator whose performance is deteriorating***. It may be the case that an operator does not realize that his or her performance has dropped below acceptable levels, and merely need to be reminded of it and/or motivated to improve. An operator's attention may be inappropriately allocated, so a warning provided by the supervisor (who is monitoring performance) could cue the operator 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 introduce.

2) ***Redistribute the task load among 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 could allow other team members to absorb the excess workload. If all team members are equally busy, or if others lack the expertise needed to perform certain tasks, this redistribution will not be possible. 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.*** Poor performance can result from the overloading or underloading of teams or team members, and these team members can be humans, computers, or a combination of both. If a supervisor observes an unacceptable level of performance from a subordinate, he or she could initiate a process to either relieve that operator of the most problematic aspects of the tasks or add to their workload, as required. This change could be manifested in one of two ways: 1) increasing the role of the automation (e.g., letting a UAV land itself instead of the operator manually controlling it), or 2) adjusting the individual workload by 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 and training to be effective. In addition, changing the number of team members may help in the longer term to reduce overall workload and/or improve performance, however, 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 NCO 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.

There are several design interventions that can aid NCO supervisors in amalgamating the numerous sources of incoming information. The use of large screen displays to provide “big picture” situation awareness is a common visualization technique used in command and control settings (Dudfield, Macklin, Fearnley, Simpson, & Hall, 2001; Scott, Rico, Furusho, & Cummings, 2007). Displays that aid supervisors through activity awareness have also been shown to promote effective supervision (Scott, Sasangohar, & Cummings, 2009). Activity awareness is a design approach in the collaborative technologies research field aimed at improved decision making, communication, and coordination in teamwork through intelligent sharing of group activity information (Carroll, Neale, Isenhour, Rosson, & McCrickard, 2003).

Increasing the role of automation in this nested supervisory control problem is another possible intervention. At a lower level, automation can be used to aid the human supervisor in monitoring tasks and detection of anomalous events. Recent research has shown that in NCO settings, probabilistic models can be used to notify supervisors that operators are exhibiting anomalous behaviors, and possibly even predict when such deviations from expected behavior will occur (Boussemart & Cummings, 2008). However, one significant limitation of such models, especially ones that involve machine learning, is that in NCO settings with many degrees of freedom and significant uncertainty, such models can only be used to detect different patterns of behavior than those learned, as opposed to predicting good or bad behavior. In addition, probabilistic model predictions also carry a great deal of uncertainty and are likely difficult for supervisors to understand (Tversky & Kahneman, 1974).

Lastly, whether the supervisor should be a human or a computer should be considered. 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, which are inherent in NCO settings. The computer would also lack the human capability to predict future performance based on subjective judgments from visual observations and face-to-face interactions. Furthermore, as will be discussed in the final section, issues with accountability and responsibility arise when computers are given authority for critical, life-threatening decisions.

Distributed Decision Making and Team Coordination

In 1994, Operation Provide Comfort provided humanitarian aid to over one million Kurdish refugees in northern Iraq in the wake of the first Gulf War. As part of this, the United Stated 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 system (AWACs) 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 lack of teamwork 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 pilot'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.*

Platform-centric command and control in past military operations often avoided distributed decision making and minimized team coordination in favor of a clear hierarchy for both information flow and decision making. 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 units that had little direct contact between one another. This has begun to change in recent times, and a fully realized vision of NCO will require that both local and global teams of distributed decision makers, not a few people at the top of a hierarchy, make decisions under time-pressure. Therefore, understanding the issues unique to team-based coordination and decision making takes on new importance in the context of NCO. 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 NCO scenarios.

A fundamental building block of good decision making is a high level of situation awareness (Endsley, 1995; Tenney & Pew, 2006), and in the case of distributed decision making a high level of team SA or shared situation awareness (SSA). The three levels of individual SA are: 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 2.2), 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. Given the need to support team SA in dynamic NCO environments, more research is needed in three areas: 1) technologies that can best support NCO team decision making and performance, 2) NCO team processes, both locally and distributed, that are most effective, and 3) the impact of different team architectures for network-centric operations.

Table 2.2. Team SA Requirements for Shared Information (Endsley & Jones, 2001)

Level I SA - Data	
	System
	Environment
	Other team members
Level 2 SA - Comprehension	
	Status relevant to own goals
	Status relevant to other's goals/requirements
	Effect of own actions/changes on others
	Effect of other's actions on self and overall goal
Level 3 SA - Projection	
	Actions of team members

There has been extensive work in the computer-supported cooperative work (CSCW) community that examines how different types of technologies, to include both software and hardware, support effective team decision making. Of particular interest to NCO is the relationship between collaborative technologies over space and time as depicted in Figure 2.4 (adapted from Johansen, 1988). Developing technologies that promote collaboration both locally and remotely in space as well as synchronous and asynchronous in time is highly relevant to NCO. For example, Navy personnel assigned to different watch sections for a ship's power plant share a space but communicate across different times. One collaborative technology they use to pass knowledge is a log book. Logs are used by watchstanders to record significant events, times of occurrence, who was notified, what actions were taken, etc. Existing on paper for hundreds of years on ships, written logs are now giving way to electronic logs, with the added benefits of an easier search space and automated reminders.

Many technologies developed for corporate settings show promise for NCO applications (e.g., electronic whiteboards (Price, Miller, Entin, & Rubineau, 2001), and table top displays (Scott, Carpendale, & Habelski, 2005)), however, more research is needed into both the promised benefits, as well as unintended consequences. For example, chat, a popular and ubiquitous synchronous communication tool for remotely distributed military individuals and teams, can have unintended consequences in terms of degraded human and system performance ((Cummings, 2004b) and see the Attention Allocation discussion.) In another study focusing on the benefits of shared displays between co-located team members, the shared displays unexpectedly contributed to degraded performance due to an increase in workload (Bolstad & Endsley, 2000). Given the high risk and uncertainty of time-sensitive operations, more investigation into the impact of new collaborative technologies is warranted, particularly in regards to how they support or detract from team SA.

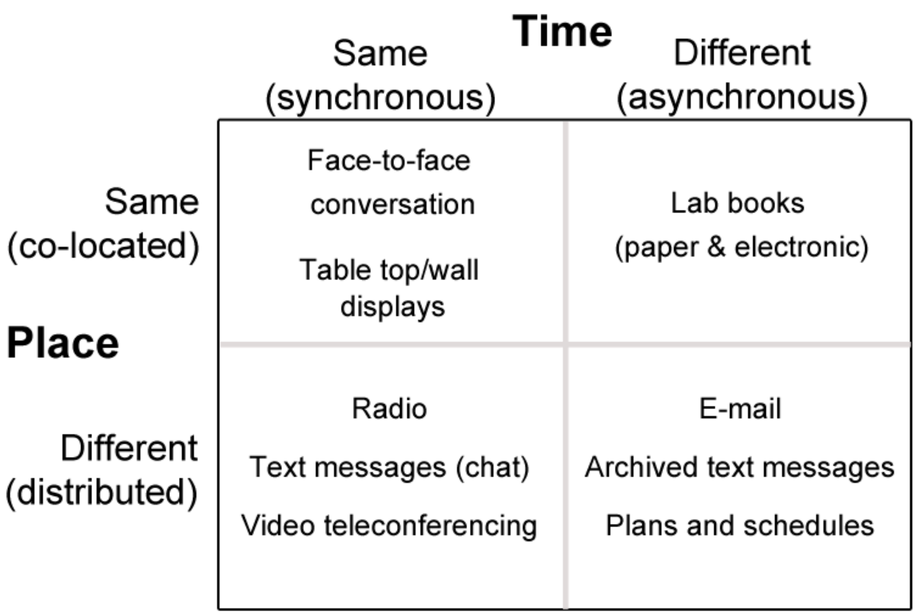

Figure 2.4. NCO time and place communications tools, adapted from Johansen, 1988.

Understanding team processes, the second area of import for distributed decision making in NCO settings, is yet another area that can benefit from research already underway in the CSCW community. However, the caveat remains that NCO settings carry unique constraints not experienced in business settings. Processing capabilities of humans are limited, especially in time-critical domains, so distribution of information, resources and tasks among decision-makers must be done in such a way that the cognitive load of each person on a team is not exceeded. However, as pointed out by Hutchins in aviation and ship navigation (Hutchins, 1995), cognition is distributed across persons and artifacts in a system such that teams often possess shared capabilities beyond the sum of individual contributions.

While difficult to capture, Cooke et al. (2003) measured distributed cognition in a command and control setting through team knowledge and succeeded in predicting subsequent team performance. However, distributed cognition across large-scale time-sensitive operations with multiple entities, both human and automated, is not well understood. For example, the misperception of system feedback may lead to catastrophic events and degraded performance in complex, tightly coupled, economic systems, as demonstrated in laboratory experiments (Sterman, 1989). Similarly, in the domain of multi-operator, multi-automation systems, experimental applications have shown that inappropriate reliance on automation hinders cooperation between human operators (Gao, Lee, & Zhang, 2006). This same bullwhip effect, i.e., the failure to perceive appropriate feedback cues and time delays resulting in oscillating behavior which may lead to highly degraded system performance, is an intrinsic risk of NCO systems. Indeed, research by Perrow (1984) has demonstrated that degrees of coupling between system elements are of particular importance to characterize and predict future system behavior. In particular, complex, tightly coupled technological systems (such as power plants) may invariably produce disasters due to this complex nature.

Lastly, the design of team architectures to include role allocation, team geographic distributions, communication, and team sizes is critical to successful NCO distributed decision making and team coordination. Research has shown that organizations operate most efficiently when their structures and processes match their mission environments (Levchuk, Levchuk, Luo, Pattipati, & Kleinman, 2002), but it is not always clear whether traditional top-down hierarchies or more lateral heterarchies provide the most efficient structure (Baker, et al., 2004). Moreover, sensor quality and operational tempo can drive the need for distributed versus centralized control

(Dekker, 2002), thus further complicating the team architecture problem by adding technological artifacts.

The problem of uncertainty, driven both by the lack of sensor and human information, will likely be a significant driver of NCO team decision-making success or failure. Price et al. (2001) hypothesize that teams organized by functional structure (a team that specializes in performing a task) perform better in terms of time and task accuracy when there is a high degree of certainty in the environment. When an environment is characterized by uncertainty and unpredictability, Price et al (2001) assert that instead of a functional structure, teams should be organized through a divisional structure that is relatively autonomous and independent. Unfortunately the very nature of NCO is high uncertainty, but since information sharing between teams is critical, they cannot be either independent or autonomous, thus the characterization of teams either as functional or divisional in NCO is problematic. Levchuk et al. (2004) propose that the most effective team architecture will be a hybrid organization which utilizes the beneficial characteristics of hierarchies as well as heterarchies such that control is an emergent property which is a function of the environment, scenario initial conditions, and adversary behavior.

Given these three areas of distributed decision making that have particular relevance in NCO settings, one unifying theme is the need for better team performance metrics and models. Assessing team performance in a repeatable, objective fashion that is not strictly contextually-based and developing predictive models of resultant behavior is a significant challenge. For example, efficient team performance has been shown to be related to the degree that team members agree on, or are aware of task, role, and problem characteristics (Fiore & Schooler, 2004), which is essentially team SA. However, there is no common team SA measurement method or metric, and even overall mission performance measures cannot effectively measure team SA. Moreover, better performance does not necessarily mean all team members share a common picture (Gorman, Cooke, Pederson, Connor, & DeJoode, 2005). Thus, in order to describe, diagnose and predict team behavior, more effective measures of team awareness are needed.

The emerging field of social network analysis may provide new insight into more objective team assessment methods, particularly in NCO settings. Theories of network analysis have emerged from various fields like sociology, anthropology, physics, mathematics, biology, and economics (see Watts, 2004, for a broad review), and have been successful at modeling uncertainty in complex, interconnected networks, like those in supply-chain management (Blackhurst, Wu, & O'Grady, 2004). It has also been argued that a network-based approach using non-linear dynamics is efficient at modeling and studying complex, coupled systems such as food webs, neurons, or power grids (Strogatz, 2001). Using social network analysis methods for NCO systems might therefore prove successful at identifying high-traffic, central nodes, providing a theoretical basis for describing team characteristics and dynamics that make NCO supervisory control different than typical supervisory control.

Trust and Reliability

A survey of Air Force and Air National Guard pilot attitudes regarding the role of UAVs as wingmen for manned aircraft revealed an inherent distrust in highly autonomous systems. Pilots generally thought UAVs were unsuited for a variety of missions, including close air support, search and rescue, and most strike missions. The pilots asserted that only humans are capable of operating in the "free flowing environment" of an offensive combat mission, which requires experience and knowledge to accurately assess the situation and determine a *course of action. In addition, pilots did not think that a group of UAVs should be allowed to self-organize because they thought UAVs could not make informed decisions about both their individual states and how their capabilities served the current mission. Furthermore, most pilots did not want UAVs operating anywhere near friendly forces on the ground and some did not think UAVs should operate in the same airspace with tactical manned aircraft. Fighter pilots, in particular, were hesitant to accept a role for UAVs in offensive combat operations, believing that a UAV could never replace a human wingman. An A-10 pilot described his relationship with his wingmen as one of trust and loyalty in that they trained, worked, and fought together, and that a UAV could never replace a human wingman (Cummings & Morales, 2005).*

Trust is an issue central to appropriate reliance on automation. Distrusting automation when it is perfectly capable has been shown to lead to disuse or misinterpretation of results in order to fit an operator's mental model, even if the subsequent workload caused by the distrust is very demanding and/or time consuming (de Vries, Midden, & Bouwhuis, 2003; Lee & Moray, 1994; Muir, 1987; Parasuraman & Riley, 1997). In contrast, other studies have found that users tend to trust automation even when it was failing (Conejo & Wickens, 1997; Gempler & Wickens, 1998). This is a particular problem with unfamiliar, time-pressured scenarios most critical to NCO, which can cause operators to become complacent and overly trust unreliable automated systems to the detriment of the overall mission. These results have also been mirrored in the domain of highly automated process control settings (Manzey, 2008). Thus designers of NCO systems are faced with a conundrum – how to design a system that is trusted and utilized to its fullest extent, yet not overly trusted such that operators become complacent.

Calibrating an operator's trust level to an automated system's actual trustworthiness or reliability is one solution to the problem of too little or too much trust (Moray, Inagaki, & Itoh, 2000; Muir, 1987). However, this is not an easy task because trust is dynamic in that it changes with exposure to, and time between, system failures. People initially exhibit a positive bias towards automation, expecting it to be perfect or at least better than humans. After an initial system failure, there is a sharp decrease in trust unless there is some level of transparency provided to the user as to why the system might fail (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003). This is particularly important in NCO settings due to the potential for bad information to propagate through the network, possibly due to information warfare or deception tactics.

Trust typically rebounds with consistently correct automation, and if the automation fails subsequent to this initial failure, trust decreases once again but it is regained more quickly (Lee & Moray, 1992; Moray, et al., 2000; Muir & Moray, 1996). Even though the loss of trust for a specific failure can be overcome, the danger is that such failures can cause long-term distrust in the automated system to the extent that it falls into disuse without further opportunity for the user to interact with it and properly re-calibrate trust (Ruff, Narayanan, & Draper, 2002). In NCO settings, research in trust and automation reliability in command and control tasks has demonstrated that reliable automation aids operators, and unreliable automation can significantly degrade performance (Dixon, Wickens, & McCarley, 2007; Rovira, McGarry, & Parasuraman, 2007). Dixon et al. (2007) recommended that any automated decision support system that functions below 70% accuracy would result in unacceptable performance degradation.

Although the reliability of automation is a primary factor affecting trust, how information is presented also affects a user's perception of reliability (Lee & See, 2004; Sarter & Schroeder, 2001). Therefore, feedback in terms of self-evaluation and interaction with automated decision aids has been suggested as potential strategies for appropriate trust calibration (Lee & Moray, 1994; Muir, 1987). Furthermore, Lee and See (2004) suggest that trust is more easily calibrated when the system is designed with the additional requirements of high resolution and specificity. Resolution measures how system reliability linearly maps to user trust, and specificity measures the degree to which a user is able to differentiate levels of trust between different parts of the system (functional) and current/past performance (temporal). Displaying and clearly differentiating the automation's confidence in its components may facilitate better calibration of trust. However, displaying confidence information about an automated recommendation or solution is not a straightforward matter. Humans are not intuitive statisticians and tend to introduce biases in probabilistic reasoning (Tversky & Kahneman, 1974), which makes presenting probabilistic confidence information to operators so they can make unbiased decisions a challenging design problem.

In NCO situations, confidence in information is a function of meta-information characteristics such as uncertainty, age of the data, and the source(s), which should be incorporated as feedback to the user in some manner. For example, if a vital network element such as a sensor stops updating in a timely manner, it is important to show the user that automated decisions made using the outdated information are less reliable, and therefore to be accorded less trust.

Several techniques have shown promise for displaying confidence information to users. McGuirl and Sarter (2003) demonstrated that a categorical trend display of a computer's confidence in its recommendations was superior to a probabilistic static representation for in-flight icing interventions. Uncertainty in information has also been successfully conveyed through degraded images using blended color icons. In addition, the addition of numeric probabilities provided no additional advantage (Finger & Bisantz, 2002). Bisantz et al. 2006) also showed that users could distinguish and rank display elements by differences in graphic saturation, transparency, and brightness.

In terms of automation transparency, making probabilistic rules and algorithms used by the automation transparent to the user, particularly if they are context dependent also may help humans to better understand and interpret complex system decisions (Atoyan, Duquet, & Robert, 2006). In a time-pressured navigation setting, Buchin (2009) demonstrated that operators can develop appropriate trust in an imperfect automated path planning tool when given the ability to adjust an algorithm's level of risk in obstacle avoidance.

Once an operator's trust has been properly calibrated, Xu et al. (2004) have demonstrated that even with imperfect automation, human operators can still properly execute their tasking, which is an important design criteria in NCO settings with significant embedded autonomy that can never be perfect. Given the significant uncertainty that will exist within actual NCO environments, as well as the uncertainty introduced by imperfect automation, more research is needed in trust calibration techniques, especially as they apply to time-pressured decisions.

Previous trust literature has generally not focused specifically on trust in and across networks, so there is a need for further study of trust as it relates to human interaction, large-scale network information uncertainty, system security, information verification, possible sources and impact of information deception, and sabotage. One exception is a study that focused on the influence of network reliability on human performance for the National Oceanographic and Atmospheric Administration (NOAA) Continuous Operational Real-Time Monitoring System (CORMS) (Bayrak & Grabowski, 2006). This study showed that decreased network reliability resulted in degraded operator satisfaction, reduced operator confidence, and increased operator workload. What remains an open question is how to design systems (including the actual networks, levels of reliability, interfaces, etc.) to mitigate these effects on confidence an workload.

TECHNOLOGICAL ASPECTS OF SUPERVISORY CONTROL SYSTEMS

Role of Automation

The Patriot missile system has a history of friendly fire incidents that can be partially attributed to a lack of designer understanding of human limitations in supervisory control. On March 23rd, 2003, a RAF Tornado GR4 was shot down in friendly airspace by a Patriot missile. 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. The "hostile radar" was, in fact, a Patriot missile battery aiming at him. On April 2nd, 2003, just nine days later, another Patriot missile shot down a US Navy F/A-18 returning to base from a mission. The Patriot missile has two modes: semi-automatic (management by consent – an operator must approve a launch) and automatic (management by exception – the operator is given a period of time to veto the computer's decision). However, in practice the Patriot is typically left in the automatic mode and the friendly fire incidents are believed to be a result of problems in the automatic mode.

There are known "ghosting" problems with the Patriot radar, which result in the appearance of false targets on a Patriot operator's screen, caused by close proximity to other Patriot missile batteries. Under the automatic model, operators are given approximately 15 seconds to reject the computer's decision, which is insufficient for *operators to detect both false targeting problems as well as adequately address friend or foe concerns through other means of communication. After the accident investigations, the US Army admitted that there is no standard for Patriot training, that autonomous operations procedures (automatic mode) are not clear, and that operators commonly lose situational awareness of air tracks. Despite all the known technical and operational problems for Patriot, the US Army states, "Soldiers [are] 100% reliant on the Patriot weapon system" (32nd Army Air and Missile Defense Command, 2003).*

Automating significant aspects of NCO is necessary so that information sharing can be both quick and comprehensive. However, what to automate and to what degree to automate a process/system are central questions in the design of NCO systems. Sheridan and Verplank (1978) outlined a scale from 1-10 where each level represented the machine (i.e., computer) performing progressively more tasks than the previous one, as shown in Table 2.3. 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 workload, and the opposite is true for low levels of automation (Kaber, Endsley, & Onal, 2000).

Table 2.3: Levels of Automation (Sheridan & 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.

It is possible to have a LOA that is too high or too low, each with its own distinct set of problems (Billings, 1997). In supervisory control applications, a LOA that is too high can result in manual or mental skill degradation, loss of situational awareness due to lack of automation transparency and inadequate feedback, brittleness, and literalism. Automated systems may not be able to handle novel or unexpected events or operate effectively in conditions near, or at, the edge of the intended operating envelope (Smith, McCoy, & Layton, 1997). On the contrary, a LOA that is too low can results in problematic decision biases and heuristics discussed previously, complacency and boredom, and greater operator confusion when something fails, unless safeguards are in place.

A significant design question in the context of developing HSC decision support systems for NCO is determining the appropriate LOA. 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 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 (2007), each information-processing task can be further divided into simple sub-tasks with differential levels of automation. For NCO, examples of generalized cognitive tasks under these proposed information processing stages are given in Table 2.4.

Table 2.4. Generalized NCO Cognitive Tasks

Information acquisition
Monitoring resources (such as friendly forces)
Monitoring systems (such as surveillance networks)
Communications
Information analysis
Data fusion and 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. As discussed in a previous section (information overload), the potential for cognitive saturation in these phases 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. LOAs for data fusion can vary such that 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).

Two different levels of automation for the same data fusion task are represented in Figure 2.5, which represents the allocation of resources (weapons) to appropriate missions. At the top, operators are given detailed information about weapon capabilities as well as mission requirements and constraints and the automation provides low level filtering (LOA 2, Table 2.3). At the bottom of Figure 2.5, operators are given the same information at a higher, more graphical level using a configural display, but a “smart” search algorithm narrows the solution space to nominally assist the operators (LOA 3, Table 2.3). The graphical display may seem more intuitive, however, results with military personnel highlight a significant general problem with more advanced LOAs. While operators generally agreed that low level display caused excessive workload, users felt that the high level graphical display did not provide sufficient detail to completely understand the problem (Cummings & Bruni, 2009). Designers of NCO systems struggle with this same conundrum of selecting automation levels that both relieve operator workload, yet provide enough information for operators to make informed decisions.

Figure 2.5. Low vs. high level displays for data fusion.

In addition, the dynamic and stochastic nature of NCO often requires flexibility in human-automation task allocation. For example, whereas a high level of automation may be desired during monotonous extended periods of time of monitoring, when an off-nominal problem occurs that requires human judgment, lower levels of automation are likely preferable. One way to achieve this flexibility in NCO systems is through adaptable automation, which differs from adaptive automation in that the operator can dictate when to change some level of automation, as opposed to some automated agent changing the system state. Adaptive automation will be discussed in detail in the next section, and an additional discussion of human-automation interaction issues can be found in Klein, Woods, Bradshaw, Hoffman, and Feltovich (2004).

An example of adaptable automation is pictured in Figure 2.6. Similarly to those in Figure 2.5, this interface supports the data fusion and planning task of allocating weapons to missions. However, this interface operates on a range of levels of automation, from LOA 2 to LOA 4, rather than at one specific LOA. Indeed, operators using the interface of Figure 2.6 may choose between the following three levels of automation: 1) LOA 2 – the operator solves the data fusion task manually, similarly to the leftmost interface of Figure 2.5, where automation only provides low-level filtering and sorting functionalities, 2) LOA 3 – a multi-criteria, heuristic search algorithm (Automatch) may be used by the operator to produce feasible solutions, which can be saved, compared, and manually completed, updated or modified by the operator, and 3) LOA 4 – the automated search algorithm creates one feasible solution upon direct request by the operator.

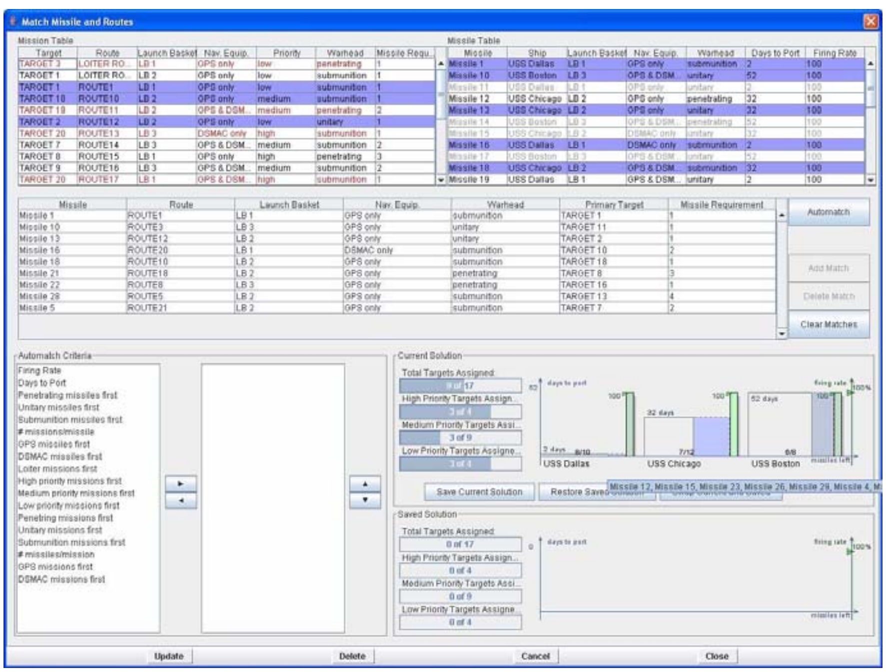

Figure 2.6. Multi-LOA display for data fusion

This multi-LOA interface's adaptable automation therefore allows the operator to select the extent to which automation supports the data fusion task, from a mostly manual setting (LOA 2) to intermediate collaboration between the operator and the automated algorithm for building solutions (LOA 3), to fully automated solution construction (LOA 4). Such flexible approaches to planning problems have also been used in the single operator control of multiple unmanned ground vehicles (Parasuraman, Galster, Squire, Furukawa, & Miller, 2005), so their use across different NCO cognitive tasks will likely be widespread. Even though such approaches are powerful in improving human understanding and possibly improving system performance, as time pressure and incoming information increases, providing operators with too much flexibility could lead to attention allocation problems as previously discussed.

Adaptive Automation

Research is underway to develop the Cognitive Cockpit which mitigates excessive pilot workload through augmented cognition (Dickson, 2005; Diethe, Dickson, Schmorrow, & Raley, 2004). The Cognitive Cockpit contains a Cognition Monitor that hypothetically provides a real-time analysis of the cognitive–affective state of the pilot through psychophysiological measures as well as inferences about current attentional focus, ongoing cognition, and intentions. These inferences become inputs to a task manager that changes the levels of automation (as described in Table 2.3) for relevant tasking and workload. This process is termed Adaptive Dynamic Function Allocation, which increases automation levels in periods of high workload and vice versa. Adaptive automation is thought to benefit operators during periods of high workload, however, using a psychophysiological state as a trigger for adaptation is problematic due to the *large amount of noise present in these signals and extreme individual variability. Although preliminary research in lab settings suggests psychophysiologic adaptive automation is possible, the current state of-the-art technology does not support operational use.*

Military operations are 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 temporary information overload placed on the operator and the need for the operator to cognitively reorient to the new situation. With NCO, the amount of information available to personnel at all levels will be greater than in previous operations. Therefore, the problem of information overload, particularly during bursts of actions, will become much more common. In addition, vigilance must be sustained between high-demand cognitive states, when detrimental boredom and lack of attention may loom (Wiener, 1984, 1987).

One method to alleviate these problems is the use of adaptive automation (AA), where changes in the level of automation are made based on some sensed operator state change. Automated-driven changes in the level of automation (Table 2.3) can be driven by specific events in the task environment, models of operator performance and task load, physiological responses, or by changes in operator performance (Parasuraman, Bahri, Deaton, Morrison, & Barnes, 1992). AA has been shown to improve task performance (Hilburn, Jorna, Byrne, & Parasuraman, 1997; Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2003), to enhance situational awareness (Kaber & Endsley, 2004), and reduce workload (Prinzel, et al., 2003), including single operator control of multiple unmanned vehicles (Parasuraman, Cosenzo, & de Visser, 2009). Two important questions to address when developing an adaptive automation decision support system are: 1) When should adaptive automation be used? (i.e., under what circumstances should LOAs change?), and 2) Should the computer or the human should change the LOA?

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 appropriate levels of vigilance, automation will be turned off during periods of low workload. Alternatively, it may be known that particular events in the environment cause higher levels of workload than desired for a short time, during which the automation could increase to compensate. An example of this would be an operator responsible for multiple UAVs performing a bombing mission. Cruise and loiter phases of flight are relatively low workload, but approach and bombing phases significantly increase operator workload. As a preemptive measure, adaptive automation could increase the LOA for some aspect of the approach and bombing phases. However, this approach is problematic since it is scenario specific. It will likely not handle unexpected situations correctly because desired changes in LOA must be pre-determined. In addition, this method 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.

AA cueing can also be accomplished through models of operator performance which can predict the effectiveness of humans during particular processes and behaviors. Thus, such a model's forecasted level of operator mental workload or performance on any number of specified tasks can be used to change the LOA. What defines an acceptable level of a predicted measure 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 models and first principles models. 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.

One modeling approach to adaptive automation cueing in supervisory control systems is through mental workload models that correlate a level of workload with some undesired performance level, and then change the LOA accordingly. The key in this AA approach is correctly identifying an undesired mental workload state. Psychophysiological measures such as the electroencephalogram (EEG), event-related brain potentials (ERPs), eye movements and electroculography, electrodermal activity, heart rate and heart rate variability, 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.

Analysis of brain waves has historically been the primary method of investigating neural indexes of cognition (Fabiani, Gratton, & Coles, 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, Bogart, & Bartolome, 1995). 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. The P300 component of ERPs has been associated with the availability of information processing resources, (Polich, 1991; Rugg & Coles, 1995), and as a measure of mental workload (Donchin, Kramer, & Wickens, 1986; Kramer, 1991). However, use of ERPs in non-laboratory settings to measure workload in real-time has proven difficult, as they are obtained from averaging EEG signals over a number of trials.

Eye movements have also been used extensively in a variety of studies, particularly in distraction studies to identify areas of attention. 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 & Wilson, 1998; Orden, Limbert, Makeig, & Jung, 2001; Veltman & Gaillard, 1998). Simon et al. (1993) demonstrated that the more general visual behavior of scanning patterns becomes more organized when task difficulty increases. 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 & Polt, 1964; Orden, et al., 2001). However, 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.

One significant problem for psychophysiologic devices is that they can be obtrusive and physically uncomfortable for subjects, creating a possible anxiety effect as well as interfering with external validity. Other significant problems include the large amount of noise present in readings, and extreme individual variability. One way to lessen these effects is to use combinations of measurements taken in concert for an individual, 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 set of 43 physiological features.

Unfortunately, the engineering obstacles in combining measures from psychophysiologic devices such as EEG and eye trackers to adapt automation in actual NCO settings are substantial.

For example, how to measure EEG signals in a dynamic, noisy environment, which is critical to the success of field technologies, is still illusive. One major hurdle is the development of a wireless EEG device that is unobtrusive, does not require the use of conducting gel (otherwise known as a dry EEG), and is able to process on-board signals, all while personnel are in motion, often under hostile environmental conditions. Even though some advances have been made in wireless EEGs, the signals from these devices are substantially weaker than more traditional EEG devices. Eye tracking devices suffer from similar problems in that they currently require sophisticated head tracking devices in addition to the eye tracking devices for field operations, so encapsulating this technology into an unobtrusive device that can be worn in the field is a significant engineering challenge.

In addition to the hardware limitations for psychophysiologic adaptive automation technologies, all adaptive automation approaches that rely on predictive probabilistic algorithms to change automation tasking suffer from the inherent uncertainty typical in NCO settings, which limit their ultimate usefulness. Actual battlefield conditions and the large number and degrees of freedom of information states will mean that much more precision will be needed in predictions. This means that not only will the sensors and signal processing algorithms have to improve substantially as has been noted by others (Kramer & Weber, 2007), but also significant advances are needed in decision-theoretic modeling. In addition, the impact of information visualization and decision authority in adaptive automation command and control environments is not well understood and deserves further scrutiny (Parasuraman, Barnes, & Cosenzo, 2007).

Lastly, in order to change LOAs, such predictive systems assume that they can determine an optimal cognitive load for an individual in dynamic, highly uncertain settings. The problem with this assumption is that optimality is dynamic and contextual, and precise individual workload models are extremely difficult to build. So unless dramatic leaps are made in the miniaturization of these technologies as well as improved signal processing algorithms in the near-future, the realization of programs like the Cognitive Cockpit or many others envisioned under the augmented cognition umbrella (Stanney, et al., 2006) are in the very distant future.

Multimodal Technologies

Since UAV operators are primarily tasked through their visual channel (i.e., computer screens are the primary conduit for communication in remote UAV control), the military branches are interested in extending human information processing capabilities through other channels, i.e., auditory and touch. Laboratory research has shown clear performance improvements in command and control tasks when operators are given multimodal displays such as spatial audio in concert with visual displays. Although multimodal displays have been beneficial in laboratory settings, their success in operational settings is not as clear. This disconnect primarily stems from a lack of designer understanding of the environmental context of such displays. In one case, despite careful research and development, the introduction of spatial audio target alerting for UAV operators failed when actual operators would wear only one of two ear cups (both are required for proper spatial audio effects). Operators insisted on wearing only one ear cup since they wanted the other free to hear conversation in the command center. In a related case, the introduction of touch displays for operators in command centers was not as successful as hoped, primarily because operators would aggregate around a single operator’s screen and point close to the screen in an explanatory gesture, which would cause inadvertent changes in displays, leading to high levels of user frustration.

Given the likely overload of information in a NCO supervisory control system, one possible avenue for mitigation of high mental workload is the use of multimodal displays, which allow operators the ability to utilize multiple sensory channels to increase the amount of information processed (Wickens, Lee, Liu, & Becker, 2004). Multiple resource theory (MRT), a theory of attention and workload, asserts that people can process some tasks in parallel without sacrificing performance because there are distinct attentional and cognitive resources that differ along three dimensions (Wickens, 2002). The first dimension, processing stages, encompasses the three major information processing stages of perception, cognition and response. Processing codes, the second dimension, refers to verbal (i.e., listening and speaking) and spatial processes (i.e., navigation and shape rotation). The last dimension, perceptual modalities, includes visual and auditory modalities as the two major channels where information is perceived.

MRT predicts that tasks will less likely interfere with each other if they occur during different information processing stages (i.e., perception and cognition vs. response), with different cognitive coding (i.e., spatially-coded tasks vs. verbally-coded tasks), and requiring different response outputs (i.e., manual output like toggling a switch vs. vocal output like saying a command). Multimodal displays essentially try to maximize this use of multiple, but orthogonal channel processing in order to improve human performance. This form of dual-coding of information has been shown to reduce workload and speed information processing, particularly for visual and auditory, or visual and tactile modalities (Burke, et al., 2006; Miller, 1991). However, it should be noted that even though humans can theoretically integrate information from multiple channels, functional MRI research has demonstrated that when humans attempt to accomplish two concurrent cognitively demanding tasks, responses are inferior to those of the individual tasks had they been performed separately (Just, et al., 2001).

The use of audio displays to augment visual ones for alerts and warnings is standard practice in many time-pressured, supervisory control systems such as process control and aviation. For example, audio displays in cockpits allow pilots the ability to successfully divide their attention so that they can listen for relevant audio signals, while still attending to visual information. Audio displays work best when the inputs are short, simple, do not have to be remembered for long periods, and relate to an action required immediately (McCormick, 1976), so they are primarily used for alerts and warnings in the cockpit. Pilots have been shown to respond faster to auditory alerts than visual ones by up to 30-40ms. In current UAV operations, discrete auditory alerts are used to supplement visual information, and research shows that providing audio alerts to UAV operators enhances their performance in monitoring automated tasks, as well as their overall mission performance (Dixon, Wickens, & Chang, 2005).

Recent research has shown that more advanced forms of auditory displays can also improve performance. In terms of operators working in NCO environments that are inherently visually intensive, one example of a multimodal auditory intervention shown to be effective in laboratory environments is spatial (sometimes referred to as 3D) audio, which presents information spatially in three dimensions so operators can determine the longitude and latitude of auditory signals. This kind of information presentation, when combined with visual cues, has been shown in command and control settings to improve target acquisition, enhance alert and warning detection, and lower overall perceived workload (see Nehme & Cummings, 2006, for a review).

Another type of advanced auditory display that shows promise in reducing workload and improving overall operator performance in supervisory control settings is the use of sonifications. Leveraging preattentive processing, sonifications are a form of continuous audio that map alerts to the state of a monitored task, allowing operators to quickly determine a task's current and projected state (Kramer, 1994). Research has shown sonifications to be beneficial for anesthesiology time-shared tasks, especially when coupled with visual displays (Loeb & Fitch, 2002; Watson & Sanderson, 2004). In terms of NCO settings, a study examining the effect of sonifications for different task types in the single operator control of multiple UAVs demonstrated that sonifications can lead to superior performance in this demanding task, as compared to the current state-of-the-art discrete alerts. However, sonifications masked other auditory alerts in some cases, and interfered with other monitoring tasks that require divided attention (Donmez, Cummings, & Graham, 2009). These results highlight the importance of understanding the systems implications of inserting a new technology. Sonifications can be generally helpful, but they may not always be an effective alerting scheme in the presence of other aural cues.

Often command and control settings are very noisy, resulting in masked auditory alerts even with simple discrete alerts. One other sensory channel that can be utilized to provide feedback to operators is the haptic channel. Haptic feedback has been shown to support operator performance in the domains of driving (Schumann, Godthelp, Farber, & Wontorra, 1993; Suzuki & Jansson, 2003), and aviation (Sklar & Sarter, 1999). A common haptic alert in many cockpits is the use of “stick shakers”. As an aircraft detects the onset of a stall, the stick/yoke column will shake, which simulates the movement that a plane would experience if approaching a fully-developed stall. The shaking of the stick alerts the pilot through vibrations and is an extremely salient cue, which immediately informs the pilot that preventative action must be taken. In everyday life, similar haptic alerts can be seen in actual or virtual rumble strips on the sides of highways that alert drivers to a possible departure from the roadway.

Sklar & Sarter (1999) showed that when compared to visual cues, tactile and redundant visual and tactile cues resulted in faster responses to higher detections rates for mode transitions in cockpits. Employing a meta-analysis of previous literature, Burke et al. (2006) demonstrated that, in general, an additional modality enhances operator performance overall, with visual-auditory feedback as the most effective in moderate workload conditions, and visual-tactile feedback more effective for high workload conditions. Furthermore, other natural language multimodal design interventions can potentially be useful in supervisory control settings such as voice recognition (Draper, Calhoun, Ruff, Williamson, & Barry, 2003), speech-to-text applications (and vice versa) and gesture recognition (Perzanowski, Schultz, Adams, & Marsh, 2000). How to best design adaptive multimodal systems, as well as cross-modal systems that leverage multiple concurrent modalities is still an area of burgeoning research (Sarter, 2006).

Although multimodal designs can promote reduced workload and increased performance in laboratory settings, the more important considerations in designing field applications, particularly in high risk supervisory control settings, are what characteristics and constraints of the environment, both physical and social, affect the system. As demonstrated in the vignette at the beginning of this section, in the case of spatial audio applications which provide benefit for some laboratory tasks, not understanding the social infrastructure caused a design intervention to fail. This was also the case for the touch displays, which are generally effective for single user applications, but in group settings where people gesture as part of their natural interactions, such applications may cause high levels of frustration, which is contrary to the intended effect.

Such potential application issues highlight the critical need to take a more holistic systems engineering approach in the design and introduction of multimodal technologies, so that all stakeholders’ requirements are considered. A systems approach assumes that thorough requirements are identified, which in human-in-the-loop settings generally means that a comprehensive cognitive task/work analysis is conducted to understand users, the intended use of the technology, and the context of the use. Failing to address even one of these three critical elements will likely result in deficient technology.

Accountability

Figure 2.7 (courtesy of Dr. John Pye of Exponent) illustrates the potentially tragic, albeit unintended, consequences of autonomous system design. The purpose of the unmanned ground vehicle (UGV) in the center of the picture is to enter a hostile area and autonomously kill enemy soldiers using its twin double-barreled shotguns. In the scenario in Figure 2.7, the UGV entered a village thought to be inhabited by insurgents, only to be greeted by curious children, obviously fascinated by the seemingly innocent toy. Fortunately the guns were not loaded that day and the automatic firing capability was not enabled. However, research is currently underway in industry and academia to develop unmanned systems that can independently engage targets without human approval.

Figure 2.7. Unintended consequences of automation, courtesy of Dr. John Pye.

In addition to the myriad of technical issues that surround NCO human supervisory control problems, there are also social and ethical considerations, especially for weapon systems that impact humans in such a dramatic fashion. What might seem to be the most effective design from a technical viewpoint may not be the most responsible. In one of the few references in the technical literature on humans and automation that considers the relationship between automation and moral responsibility, Sheridan (1996) is wary of individuals "blissfully trusting the technology and abandoning responsibility for one's own actions."

Many technical design issues can be resolved through modeling and testing, however, degradation of accountability and abandonment of responsibility when using automated systems is a much more difficult question to address. Automated tools are designed to improve decision effectiveness and reduce human error, but they can cause operators to relinquish responsibility and accountability because of a perception that the automation is in charge. Sheridan (1983) maintains that even in the information processing role, "individuals using the system may feel that the machine is in complete control, disclaiming personal accountability for any error or performance degradation."

This possible abandonment of accountability will likely also be exacerbated by the remote distances inherent in NCO. Remote warfare is thought to significantly reduce human resistance to killing (if not eradicate it) (Grossman, 1998), so the combination of distance and autonomous computer agents, expressed through human-computer interfaces, could erode operators' senses of accountability and create moral buffers between operators and the world they are influencing (Cummings, 2004a).

In theory, increased accountability motivates people to employ more self-critical and cognitively complex decision-making strategies (Tetlock & Boettger, 1989). In one of the few studies attempting to examine the effects of automation on accountability, Skitka, Mosier, and Burdick (2000) performed an experiment in which subjects were required to justify strategies and outcomes in computerized flight simulation trials. The results showed that not only did increased accountability lead to fewer instances of automation bias through decreased errors of omission and commission, but also improved overall task performance.

How then could systems be designed to promote accountability, especially in the context of NCO? One tangible system architecture consideration for accountability is the number of people required to interact with a given decision support system. Research indicates that responsibility for tasks is diffused when people work in collective groups as opposed to working alone. This concept is known as "social loafing" (see Karau & Williams (1993) for a review). This is of particular concern in distributed systems like those expected in NCO systems since task responsibility will often be delegated to many. Although research indicates that people experience degraded task responsibility through collective action, the potential loss of a sense of moral responsibility and agency for operators interacting collectively through human-computer interfaces is not as clearly understood. It is likely that the computer interface becomes another entity in the collective group so that responsibility, and hence accountability, can be cognitively offloaded not only to the group, but also to the computer (Cummings, 2006).

Recognizing that such complex systems embedded with autonomy could make it difficult for humans to make correct ethical judgments, another proposed solution is to embed ethics in autonomous agents such as in robots (Arkin, 2007), which are expected to play a large role in the military’s instantiation of NCO. Arkin (2007) believes that robots can actually behave more ethically than humans, particularly on the battlefield, since they are not predisposed to the decision biases as previously discussed, and can act more rationally than humans under stress. However, one significant caveat to this approach is the assumption that the correct rule-based algorithms can be encoded in autonomous agents in engineering and design phases. In addition, another major hurdle is ensuring that the appropriate sensors can effectively distinguish friend from foe.

It is well-established that algorithms are notoriously brittle, particularly in uncertain dynamic settings like command and control (Smith, et al., 1997), so reliance on an *a priori* set of rules to guide robot ethical behavior, especially in areas like weapons release, seems risky. Moreover, reliance on brittle algorithms in the presence of highly imperfect sensors mounted on robots brings into question the ability of a robot embedded with an ethical governor to make good “judgments.” Ultimately, even if robots can be embedded with some degree of ethical decision making, when something goes wrong and a robot accidentally kills a human, who will be held accountable? Thus, it seems that even if we artificially embed ethics in machines, the problem of accountability does not disappear.

CONCLUSION

Both the military and the commercial sectors will be able to capitalize on the benefits from network-centric operations such as rapid access to information across the network and overall increased productivity. However, synthesizing the voluminous data intrinsic to the network and executing decisions in real-time, often with high-risk consequences under significant uncertainty will likely be a major bottleneck and possible point of failure for these futuristic systems. One important source of uncertainty often not considered in these systems are the interactions between known entities in the network and possibly unknown agents, such as general aviation aircraft with no electronic signature operating near commercial aircraft corridors. In the presence of such sources of uncertainty, the adoption of NCO principles will be problematic for human decision makers who need to execute supervisory control across complex, distributed networked systems under time pressure.

When considering the human aspect of dynamic, time-pressured NCO settings, the dominant human performance concern is the addition of the number of available information sources as well as the volume of information flow. Without measures to mediate this volume, *information overload* will be problematic, and operators will struggle with appropriate *attention allocation* strategies. Because NCO systems inherently contain webs of people working together to achieve some overarching goal, these concerns also affect personnel responsible for *supervisory monitoring of operators*, as well as those people engaged in *distributed decision making and team coordination*.

To manage the increase in information across the network, increasing *the role of automation*, including *adaptive automation* will be needed. Moreover, using *multimodal technologies* may be able to reduce operators' workload. However, technological interventions should be considered in light of *decision biases* both for individuals and groups. Often these decision biases will result in complacent behavior such that operators overly *trust* a complex automated system, but significant distrust of automated systems can also result, which is particularly linked to a system's *reliability*. Lastly, this potentially displaced trust in automation and complacency can lead to a loss of *accountability* and erosion of moral responsibility.

Despite the fact that networked systems are envisioned to support human intentions, technological determinism is pervasive in that the primary thrust of NCO research is directed toward improvements and innovations in technology (Alberts, Garstka, & Stein, 2000), and not towards human interaction with these complex systems. The typical, but naïve, assumption is that advancements in automated systems will naturally improve both system and human performance. Without dedicated focus on the impact of NCO technologies for both individual and team cognitive processes, the vision of a network with self-synchronization, shared situational awareness, and increased speed and effectiveness of operations will be replaced with a problematic, sub-optimal, and reactive network with significantly increased risk for errors.

Unfortunately, there is no set of design guidelines that can always be true for all NCO systems because, as is true in *all* human-centered systems, context is critical. More importantly, government agencies and industrial system integrators should ensure they have qualified human –system engineering personnel on staff to conduct critical cognitive task analyses, provide human-centered design guidance, and develop effective assessment strategies to evaluate the impact of technologies on human and system performance.

ACKNOWLEDGMENTS

This work was an amalgamation of research conducted in the Humans and Automation Laboratory (HAL) at MIT for a number of agencies such as Boeing Phantom Works, the Office of Naval Research, the Air Force Office of Scientific Research, the U.S. Army, the FAA, Lincoln Laboratory, Charles River Analytics, and NASA. The reviewers also provided critical insight into the shaping of this effort.

REFERENCES

32nd Army Air and Missile Defense Command (2003). Patriot Missile Defense Operations during Operation Iraqi Freedom. Washington DC: U.S. Army.

Alberts, D. S., Garstka, J. J., & Stein, F. P. (2000). *Network Centric Warfare: Developing and Leveraging Information Superiority* (2nd ed.). Washington, DC: Command and Control Research Program (CCRP).

Altmann, E. M., & Trafton, J. G. (2004). *Task Interruption: Resumption Lag and the Role of Cues.* Paper presented at the 26th Annual Conference of the Cognitive Science Society.

Andreassi, J. L. (2000). *Psychophysiology: Human Behavior & Physiological Response* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Arkin, R. (2007). *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture** (No. GIT-GVU-07-11). Atlanta, GA: Georgia Tech Mobile Robot Laboratory.

Atoyan, H., Duquet, J.-R., & Robert, J.-M. (2006). *Trust in new decision aid systems.* Paper presented at the 18th International Conference of the Association Francophone d'Interaction Homme-Machine, Montreal, Canada.

Baker, K., Entin, E., See, K., Baker, B. S., Downes-Martin, S., & Cecchetti, J. (2004). *Organizational Structure and Dynamic Information Awareness In Command Teams.* Paper presented at the 2004 International Command and Control Research and Technology Symposium, San Diego.

Bayrak, T., & Grabowski, M. R. (2006). Network Performance Impacts On Operators In Safety-Critical Systems. *International Journal of Information Technology & Decision Making, 5*(1), 173-194.

Bennett, K. B. (1992). Graphical Displays: Implications for Divided Attention, Focused Attention, and Problem Solving. *Human Factors, 34*(5), 513-533.

Berka, C., Levendowski, D. J., Cvetinovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., et al. (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*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bisantz, A. M., Pfautz, J., Stone, R., Roth, E. M., Thomas-Meyers, G., & Fouse, A. (2006, October 16-20). *Assessment of Display Attributes for Displaying Meta-information on Maps.* Paper presented at the Human Factors and Ergonomics Society 50th Annual Meeting, San Francisco, CA.

Blackhurst, J., Wu, T., & O'Grady, P. (2004). Network-based approach to modeling uncertainty in a supply chain. *International Journal of Production Research,, 42*(8), 1639-1658.

Bolstad, C. A., & Endsley, M. R. (2000). *The Effect of Task Load and Shared Displays on Team Situation Awareness.* Paper presented at the 44th Annual Meeting of the Human Factors and Ergonomics Society.

Boussemart, Y., & Cummings, M. L. (2008, 3-4 September). *Behavioral Recognition and Prediction of an Operator Supervising Multiple Heterogeneous Unmanned Vehicles.* Paper presented at the Humans Operating Unmanned Systems, HUMOUS'08, Brest, France.

Buchin, M. (2009). *Assessing the Impact of Automated Path Planning Aids in the Maritime Community.* Massachusetts Institute of Technology, Cambridge, MA.

Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R. B., Coovert, M. D., et al. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis *Proceedings of the 8th international conference on multimodal interfaces* (pp. 108-117).

Buttigieg, M. A., & Sanderson, P. M. (1991). Emergent Features in Visual Display Design for Two Types of Failure Detection Tasks. *Human Factors, 33*(6), 631-651.

Byrne, M. D., & Pew, R. W. (2009). A History and Primer of Human Performance Modeling (Vol. 5, pp. 225-263): Human Factors and Ergonomics Society.

Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., & McCrickard, D. S. (2003). Notification and Awareness: Synchronizing Task-Oriented Collaborative Activity. *International Journal on Human-Computer Studies*, *58*, 605-632.

Carroll, J. M., & Rosson, M. B. (1987). Paradox of the Active User. In J. M. Carroll (Ed.), *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction* (pp. 80-111). Cambridge, MA: MIT Press.

Caterinicchia, D. (2003, June 23). DoD Chat Use Exploded in Iraq. *Federal Computer Week*.

Cellier, J., & Eyrolle, H. (1992). Interference Between Switched Tasks. *Ergonomics, 35*, 25-36.

Conejo, R., & Wickens, C. D. (1997). *The effects of highlighting validity and feature type on air-to-ground target acquisition performance* (No. ARL-97-11/NAWCONR-97-1). Savoy, IL: Institute of Aviation.

Cooke, N., Kiekel, P. A., Salas, E., Stout, R. J., Bowers, C., & Cannon-Bowers, J. A. (2003). Measuring Team Knowledge: A Window to the Cognitive Underpinnings of Team Performance. *Group Dynamics: Theory, Research and Practice, 7*, 179-199.

Crandall, J. W., & Cummings, M. L. (2007). Identifying predictive metrics for supervisory control of multiple robots. *IEEE Transactions on Robotics - Special Issue on Human-Robot Interaction*, *23*(5), 942-951.

Crandall, J. W., Goodrich, M. A., Dan R. Olsen, J., & Nielsen, C. W. (2005). Validating Human-Robot Interaction Schemes in Multitasking Environments. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 35*(4).

Cummings, M. L. (2004a). Creating Moral Buffers in Weapon Control Interface Design. *IEEE Technology and Society*, *23*(3), 28-33.

Cummings, M. L. (2004b). The Need for Command and Control Instant Message Adaptive Interfaces: Lessons Learned from Tactical Tomahawk Human-in-the-Loop Simulations. *Cyberpsychology and Behavior, 7*(6).

Cummings, M. L. (2006). Automation and Accountability in Decision Support System Interface Design. *Journal of Technology Studies, 32*(1), 23-31.

Cummings, M. L., & Bruni, S. (2009). Collaborative Human Computer Decision Making. In S. Nof (Ed.), *Handbook of Automation*. New York: Springer.

Cummings, M. L., Brzezinski, A. S., & Lee, J. D. (2007). The Impact of Intelligent Aiding for Multiple Unmanned Aerial Vehicle Schedule Management. *IEEE Intelligent Systems: Special issue on Interacting with Autonomy, 22*(2), 52-59.

Cummings, M. L., & Guerlain, S. (2004). *Using a Chat Interface as an Embedded Secondary Tasking Tool.* Paper presented at the Human Performance, Situation Awareness and Automation Technology II Conference, Daytona Beach, FL.

Cummings, M. L., & Guerlain, S. (2007). Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles. *Human Factors, 49*(1), 1-15.

Cummings, M. L., & Mitchell P.J. (2008). Predicting Controller Capacity in Remote Supervision of Multiple Unmanned Vehicles. *IEEE Systems, Man, and Cybernetics, Part A Systems and Humans, 38*(2), 451-460.

Cummings, M. L., & Mitchell, P.J. (2007). Operator Scheduling Strategies in Supervisory Control of Multiple UAVs. *Aerospace Science and Technology, 11,* 339-348.

Cummings, M. L., & Morales, D. (2005). UAVs as Tactical Wingmen: Control Methods and Pilots' Perceptions. *Unmanned Systems, February*.

Cummings, M. L., & Nehme, C. E. (2009). *Modeling the Impact of Workload in Network Centric Supervisory Control Settings*. Paper presented at the 2nd Annual Sustaining Performance Under Stress Symposium.

Cummings, M. L., Nehme, C. E., & Crandall, J. W. (2007). Predicting operator capacity for supervisory control of multiple UAVs. In J. S. Chahk, L. C. Jain, A. Mizutani & M. Sato-Ilic (Eds.), *Studies in computational intelligence*. New York: Springer.

de Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human-Computer Studies, 58*, 719 -735.

Dekker, A. H. (2002). *C4ISR Architectures, Social Network Analysis and the FINC Methodology: An Experiment in Military Organisational Structure*. Edinburgh, South Australia: Defence Sciences and Technology Organization.

Dickson, B. T. (2005). The Cognitive Cockpit - a test-bed for Augmented Cognition. In D. Schmorrow (Ed.), *Foundations of Augmented Cognition* (pp. 479-485). Mahwah, NJ: Lawrence Erlbaum Associates.

Diethe, T. R., Dickson, B. T., Schmorrow, D., & Raley, C. (2004). *Toward an Augmented Cockpit.* Paper presented at the 2nd Annual Human Performance, Situation Awareness, and Automation Conference, Daytona Beach.

Dixon, S., Wickens, C., & Chang, D. (2005). Mission Control of Multiple Unmanned Aerial Vehicles: A Workload Analysis. *Human Factors, 47*, 479-487.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors, 49*(4), 564-572.

DoD (2001). *Network Centric Warfare: Department of Defense Report to Congress*. Washington DC: Office of the Secretary of Defense.

Donchin, E., Kramer, A. F., & Wickens, C. D. (1986). Applications of Brain Event-Related Potentials to Problems in Engineering Psychology. In M. G. H. Coles, E. Donchin & S. Porges (Eds.), *Psychophysiology: Systems, Processes, and Applications* (pp. 702-718). Middletown, NJ: Till & Till.

Donmez, B., Cummings, M. L., & Graham, H. D. (2009). Auditory decision aiding in supervisory control of multiple unmanned aerial vehicles. Human Factors, 51(5), 718-729.

Draper, M., Calhoun, G., Ruff, H., Williamson, D., & Barry, T. (2003). *Manual Versus Speech Input for Unmanned Aerial Vehicle Control Station Operations* (No. A842634). Wright-Patterson AFB, Ohio: Air Force Research Laboratory.

Dudfield, H. J., Macklin, C., Fearnley, R., Simpson, A., & Hall, P. (2001). *Big is better? Human factors issues of large screen displays with military command teams.* Paper presented at the Proceedings of People in Control (PIC'01): International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres Manchester, UK.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal Human-Computer Studies 58,* 697-718.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R., & Jones, W. M. (2001). A Model of Inter- and Intrateam Situation Awareness: Implications for Design, Training, and Measurement. In M. McNeese, E. Salas & M. R. Endsley (Eds.), *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*. Santa Monica, CA: Human Factors and Ergonomics Society (HFES).

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-Related Brain Potentials. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology*. Cambridge, England: Cambridge University Press.

Finger, R., & Bisantz, A. M. (2002). Utilizing graphical formats to convey uncertainty in a decisionmaking task. *Theoretical Issues in Ergonomic Science, 1*, 1-25.

Fiore, S. M., & Schooler, J. W. (2004). Process Mapping and Shared Cognition: Teamwork and the Development of Shared Problem Models. In E. Salas, S. M. Fiore & J. A. Cannon-Bowers (Eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance*. Washington, DC: American Psychological Association.

Fisher, C. S. (2008). *Real-Time, Multi-Agent System Graphical User Interface*. Massachusetts Institute of Technology, Cambridge, MA.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Gao, J., Lee, J. D., & Zhang, Y. (2006). A dynamic model of interaction between reliance on automation and cooperation in multi-operator multi-automation situations. *International Journal of Industrial Ergonomics, 36*, 51-526.

Gempler, K. S., & Wickens, C. D. (1998). *Display of Predictor Reliability on a Cockpit Display of Traffic Information* (No. ARL-98-6/ROCKWELL-98-1). Savoy, IL: Institute of Aviation.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gopher, D., Armony, L., & Greenspan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General, 129*(3), 308-339.

Gorman, J., Cooke, N., Pederson, H., Connor, O., & DeJoode, J. (2005). *Awareness of Situation by Teams (CAST): Measuring Team Situation Awareness of a Communication Glitch.* Paper presented at the Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL.

Grossman, D. (1998). *The Morality of Bombing: Psychological Responses to "Distant Punishment"*. Paper presented at the Center for Strategic and International Studies, Dueling Doctrines and the New American Way of War Symposium, Washington DC.

Guerlain, S., & Bullemer, P. (1996). *User-initiated notification: A concept for aiding the monitoring activities of process control operators.* Paper presented at the Human Factors and Ergonomics Society 40th Annual Meeting, Santa Monica, CA.

Guerlain, S. (2000). *Interactive advisory systems.* Paper presented at the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium, Savannah, Georgia.

Hankins, T. C., & 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., & 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., & Parasuraman, R. (1997). The Effect of Adaptive Air Traffic Control (ATC) Decision Aiding on Controller Mental Workload *Human-automation Interaction: Research and Practice* (pp. 84-91). Mahwah, NJ: Lawrence Erlbaum.

Histon, J. M., Hansman, R. J., Gottlieb, B., Kleinwaks, H., Yenson, S., Delahaye, D., et al. (2002). *Structural Considerations and Cognitive Complexity in Air Traffic Control.* Paper presented at the 19th IEEE/AIAA Digital Avionics Systems Conference, Irvine, CA.

Hutchins, E. (1995). How a cockpit remembers its speed. *Cognitive Science, 19*, 265-288.

Johansen, R. (1988). *Groupware: Computer Support for Business Teams*. NY: Free Press.

Just, M. A., Carpenter, P. A., Keller, T. A., Emery, L., Zajac, H., & Thulborn, K. R. (2001). Interdependence of Nonoverlapping Cortical Systems in Dual Cognitive Tasks. *NeuroImage, 14*( ), 417-426.

Kaber, D. B., & 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., & 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.

Karau, S. J., & Williams, K. D. (1993). Social loafing: a meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*(4), 681-706.

Klein, G. (1989). Recognition-Primed Decisions. *Advances in Man-Machine Research, 5*, 47-92.

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten Challenges for Making Automation a ‘Team Player’ in Joint Human-Agent Activity. *IEEE Intelligent Systems, 19*(6), 91-95.

Kramer, A., & Weber, T. (2007). Application of psychophysiology to human factors. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed.): Springer.

Kramer, A. (1991). Physiological Metrics of Mental Workload: A Review of Recent Progress. In D. L. Damos (Ed.), *Multiple Task Performance* (pp. 279-328). Washington, DC: Taylor & Francis.

Kramer, G. (1994). Some organizing principles for representing data with sound. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity* (Vol. *XVIII* pp. 185-222). Reading, MA: Addison Wesley.

Laughery, K. R., & Corker, K. (1997). Computer Modeling and Simulation of Human/System Performance. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 1375-1408). New York, NY: John Wiley & Sons.

Laughery, K. R., & Wogalter, M. S. (2006). Designing effective warnings. In R. C. Williges (Ed.), *Reviews of human factors and ergonomics* (Vol. 2, pp. 241-271). Santa Monica, CA: Human Factors and Ergonomics Society.

Lee, J.D, & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243 - 1270.

Lee, J.D, & Moray, N. (1994). Trust, self confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*, 153-184.

Lee, J. D., & See, K. A. (2004). Trust in technology: Designing for Appropriate Reliance. *Human Factors, 46*(1), 50-80.

Levchuk, G. M., Levchuk, Y. N., Luo, J., Pattipati, K. R., & 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.

Levchuk, G. M., Yu, F., Levchuk, Y., & Pattipati, K. R. (2004). *Networks of Decision-Making and Communicating Agents: A New Methodology for Design and Evaluation of Organizational Strategies and Heterarchical Structures.* Paper presented at the 2004 International Command and Control Research and Technology Symposium, San Diego.

Liu, Y. (1997). Queuing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems, Man, and Cybernetics, 27*(2), 195-207.

Loeb, R. G., & Fitch, W. T. (2002). A laboratory evaluation of an auditory display designed to enhance intraoperative monitoring. *Anesthesia and Analgesia, 94*, 362-368.

Lord, C. G., Ross, L., & Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *The Journal of Personality and Social Psychology, 47*, 1231-1243.

Maglio, P. P., & Campbell, C. S. (2000). *Tradeoffs in Displaying Peripheral Information.* Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI 2000).

Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies, 66*, 688-699.

McCormick, E. J. (1976). *Human Factors in Engineering and Design* (4th ed.). New York: McGraw-Hill Book Co.

McFarlane, D. C. (2002). Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction, 17*(1), 63-139.

McGuirl, J. M., & Sarter, N. B. (2003). *How are we doing?: Presenting System Confidence Information to Support Trust Calibration and Adaptive Function Allocation.* Paper presented at the Human Factors and Ergonomics Society 47th Annual Meeting, Denver, CO.

Miller, C. A., & Parasuraman, R. (2007). Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*(1), 57-75.

Miller, J. (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 160-169.

Miyata, Y., & Norman, D. A. (1986). Psychological Issues in Support of Multiple Activities. In D. A. Norman & S. W. Draper (Eds.), *User Centered System Design: New Perspectives on Human Computer Interaction* (pp. 265-284). Hillsdale, NJ: Lawrence Erlbaum.

Moray, N., Dessouky, M. I., & Kijowski, B. A. (1991). Strategic Behavior, Workload, and Performance in Task Scheduling. *Human Factors, 33*(6), 607-629.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied, 6*(1), 44 - 58.

Mosier, K. L., & Skitka, L. J. (1996). Human Decision Makers and Automated Decision Aids: Made for Each Other? In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 201-220). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

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.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics, 39*(3), 429-460.

Nehme, C. E., & Cummings, M. L. (2006). Audio Decision Support for Supervisory Control of Unmanned Vehicles HAL2006-06.

Norman, D. A. (1988). *The Design of Everyday Things*. New York: Doubleday.

Orden, K. F., Limbert, W., Makeig, S., & 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., & Barnes, M. (1992). *Theory and Design of Adaptive Automation in Adaptive Systems* (No. Progress Report No. NAWCADWAR-92033-60). Warminster, PA: Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., Barnes, M., & Cosenzo, K. (2007). Adaptive automation for human-robot teaming in future command and control systems. *International Journal of Command and Control, 1*(2), 43-68.

Parasuraman, R., Cosenzo, K., & de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology, 21*(2), 270-297.

Parasuraman, R., Galster, S., Squire, P., Furukawa, H., & Miller, C. (2005). A Flexible Delegation-Type Interface Enhances System Performance in Human Supervision of Multiple Robots: Empirical Studies with RoboFlag. *IEEE Transactions on Systems, Man, and Cybernetics, 35*(4), 481-493.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, *39*(2), 230-253.

Parasuraman, R., Sheridan, T. B., & 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.

Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. Princeton, N.J: Princeton University Press.

Perzanowski, D., Schultz, A. C., Adams, W., & Marsh, E. (2000). *Using a Natural Language and Gesture Interface for Unmanned Vehicles*. Washington DC: Navy Center for Applied Research in Artificial Intelligence.

Pfautz, J., Roth, E., Bisantz, A. M., Thomas-Myers, G., Llinas, J., & Fouse, A. (2006). *The Role of Meta-Information in C2 Decision-Support Systems.* Paper presented at the Command and Control Research and Technology Symposium., San Diego, CA.

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., & Bartolome, D. (1995). Biocybernetic System Evaluates Indices of Operator Engagement. *Biological Psychology, 40*, 187-196.

Price, J., Miller, D., Entin, E., & Rubineau, B. (2001). *Collaborative Planning and Coordinated Team Performance.* Paper presented at the 6th International Command and Control Research and Technology Symposium, Annapolis, MD.

Prinzel, L. J., III, Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & 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.

Rovira, E., McGarry, K., & Parasuraman, R. (2007). Effects of imperfect automation on decision making in a simulated command and control task. *Human Factors*, *49*(1), 76-87.

Ruff, H., Narayanan, S., & Draper, M. H. (2002). Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. *Presence, 11*(4), 335-351.

Rugg, M. D., & Coles, M. G. H. (Eds.). (1995). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. New York, NY: Oxford University Press.

Sanderson, P. M., Flach, J. M., Buttigieg, M. A., & Casey, E. J. (1989). Object Displays Do Not Always Support Better Integrated Task Performance. *Human Factors, 31*(2), 183-198.

Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges *International Journal of Industrial Ergonomics, 36*(5), 439-445

Sarter, N. B., & Schroeder, B. (2001). Supporting decision making and action selection under time pressure and uncertainty: The case of in-flight icing. *Human Factors, 43,* 573-583.

Savla, K., Nehme, C. E., Frazzoli, E., & Temple, T. (2008). *On Efficient Cooperative Strategies between UAVs and Humans in a Dynamic Environment*. Paper presented at the AIAA Conference on Guidance, Navigation, and Control.

Schmidt, D. K. (1978). A Queuing Analysis of the Air Traffic Controller's Workload. *IEEE Transactions on Systems, Man, and Cybernetics, 8*(6), 492-498.

Schumann, J., Godthelp, H., Farber, B., & Wontorra, H. (1993). Breaking up open-loop steering control actions the steering wheel as an active control device. In A. G. Gale (Ed.), *Vision in Vehicles IV* (pp. 321-332). North-Holland, Amsterdam.

Scott, S. D., Carpendale, M. S. T., & Habelski, S. (2005). Storage Bins: Mobile Storage for Collaborative Tabletop Displays. *IEEE Computer Graphics & Applications: Special Issue on Large Displays, July/August*.

Scott, S. D., Rico, A. E., Furusho, C. Y., & Cummings, M. L. (2007, March 19-21). *Aiding Team Supervision in Command and Control Operations with Large-Screen Displays.* Paper presented at the HSIS: ASNE Human Systems Integration Symposium, Annapolis, MD, USA.

Scott, S. D., Sasangohar, F., & Cummings, M. L. (2009, March 17-19). *Investigating Supervisory-level Activity Awareness Displays for Command and Control Operations*

Paper presented at the HSIS 2009: Human Systems Integration Symposium, Annapolis, MD.

Shapiro, C., & Varian, H. R. (1999). *Information Rules: A Strategic Guide to the Network Economy*. Boston, MA: Harvard Business School Press.

Sheridan, T., & Parasuraman, R. (2006). Human-automation interaction *Reviews of Human Factors and Ergonomics* (Vol. 1, pp. 89-129): Human Factor and Ergonomics Society.

Sheridan, T. B. (1992). *Telerobotics, Automation and Human Supervisory Control*. Cambridge, MA: The MIT Press.

Sheridan, T. B. (1996). Speculations on Future Relations Between Humans and Automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance* (pp. 449-460). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Sheridan, T. B., Vamos, T., & Aida, S. (1983). Adapting Automation to Man, Culture and Society. *Automatica, 19*(6), 605-612.

Sheridan, T. B., & Verplank, W. L. (1978). *Human and Computer Control of Undersea Teleoperators* (Man-Machine Systems Laboratory Report). Cambridge: MIT.

Simon, H. A., Hogarth, R., Piott, C. R., Raiffa, H., Schelling, K. A., Thaier, R., et al. (1986). *Decision Making and Problem Solving*. Washington D.C.: National Academy Press.

Simon, P., Rousseau, F., & Angue, J.C. (1993, October 17-20). *Quantitative Analysis of Mental Workload Influence on Eye Scanning Movements.* Paper presented at the IEEE International Conference on Systems, Man and Cybernetics, Le Touqet, France.

Skitka, L. J., & Mosier, K. L. (2000). Automation Bias and Errors: Are Crews Better Than Individuals? *International Journal of Aviation Psychology, 10*(1), 85-97.

Skitka, L. J., Mosier, K. L., & Burdick, M. D. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*(5), 991-1006.

Skitka, L. J., Mosier, K. L., & Burdick, M. D. (2000). Accountability and automation bias. *International Journal of Human-Computer Studies, 52*, 701-717.

Sklar, A. E., & Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human Factors, 41*(4), 543-552.

Smith, P., McCoy, E., & Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: The effects on user performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 27,* 360-371.

Somervell, J., Srinivasan, R., Vasnaik, O., & Woods (2001). *Measuring Distraction and Awareness Caused by Graphical and Textual Displays in the Periphery*. Paper presented at the 39th Annual ACM Southeast Conference, Athens, GA.

Squire, P., Trafton, G., & Parasuraman, R. (2006). *Human Control of Multiple Unmanned Vehicles: Effects of Interface Type on Execution and Task Switching Times.* Paper presented at the HRI, Salt Lake City, Utah.

Stanney, K. M., Schmorrow, D., Johnston, M., Fuchs, S., Jones, D., Hale, K., et al. (2006). Augmented Cognition: An Overview. In F. T. Durso (Ed.), *Reviews of Human Factors and Ergonomics* (Vol. 4, pp. 89-129). Santa Monica, CA: Human Factors and Ergonomics Society.

Sterman, J. D. (1989). Misperceptions of Feedback in Dynamic Decision-Making. *Organizational Behavior and Human Decision Processes, 43*(3), 301-335.

Strogatz, S. H. (2001). Exploring complex networks. *Nature, 410*(6825), 268-276.

Suzuki, K., & Jansson, H. (2003). An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *JSAE Review, 24*, 65-70.

Tenney, Y. J., & Pew, R. W. (2006). Situation awareness catches on. What? So what? Now what? In R. C. Williges (Ed.), *Reviews of Human Factors and Ergonomics* (Vol. 2, pp. 89-129). Santa Monica,

CA: Human Factors and Ergonomics Society.

Tetlock, P. E., & Boettger, R. (1989). Accountability: A Social Magnifier of the Dilution Effect. *Journal of Personality and Social Psychology, 57*(3), 388-398.

Trafton, J. G., & Monk, C. A. (2007). Task Interruptions *Reviews of Human Factors and Ergonomics* (Vol. 3, pp. 111-126): Human Factors and Ergonomics Society.

Tulga, M. K., & Sheridan, T. B. (1980). Dynamic Decisions and Work Load in Multitask Supervisory Control. *IEEE Transactions on Systems, Man, and Cybernetics, 10*(5), 217-232.

Tvaryanas, A. P. (2004). Visual scan patterns during simulated control of an uninhabited aerial vehicle (UAV). *Aviation and Space Environmental Medicine, 75*(6), 531-538.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological Workload Reactions to Increasing Levels of Task Difficulty. *Ergonomics, 41*(5), 656-669.

Vicente, K., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics, 22*, 589-606.

Watson, M., & Sanderson, P. (2004). Sonification Supports Eyes-Free Respiratory Monitoring and Task Time-Sharing. *Human Factors, 46*(3), 497-517.

Watts, D. J. (2004). The "new" science of networks. *Annual Review of Sociology, 30*, 243-270.

Wickens, C. D. (2002). Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, N.J.: Prentice Hall.

Wickens, C. D., Lee, J. D., Liu, Y., & Becker, S. E. G. (2004). *An Introduction To Human Factors Engineering* (2nd ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.

Wiener, E. L. (1984). Vigilance and inspection. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 207-246). Chichester, UK: Wiley.

Wiener, E. L. (1987). Application of vigilance research: Rare, medium or well done. *Human Factors, 87*, 725-736.

Wilson, G. F., & Russell, C. A. (2003). Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. *Human Factors, 45*(4), 635-643.

Xu, X., Wickens, C. D., & Rantanen, E. (2004). *Imperfect Conflict Alerting Systems for the Cockpit Display of Traffic Information* (No. AHFD-04-8/NASA-04-2). Savoy, IL: Aviation Human Factors Division Institute of Aviation.

Zsambok, C. E., Beach, R. B., & Klein, G. (1992). *Literature Review of Analytical and Naturalistic Decision Making*. San Diego: Naval Command, Control and Ocean Surveillance Center.