The Four-Second Supervisor: Multi-Tasking Supervision and Its Support

Mark F. St. John & Mary Ann King
Pacific Science & Engineering Group, Inc.
San Diego, CA, USA

Supervision is a broad term that covers a wide range of task environments and cognitive activities. Here, we describe one specific type called multi-tasking supervision. It involves rapidly assessing a set of distinct systems or situations and determining where attention is needed moment by moment. Multi-tasking supervision is a key component of many command and control, and other, tasks. This paper focuses on understanding the cognitive needs of multi-tasking supervision and developing support concepts for it. First, we place multi-tasking supervision within the context of other types of supervision. Next, we consider the types of tasks and work environments that are appropriate for multi-tasking supervision. Finally, we discuss computer display concepts for supporting multi-tasking supervision.

Supervision is a broad term that covers a wide range of task environments and cognitive activities from teaching and directing robots or other people to perform tasks, to monitoring complex systems, to actively troubleshooting system displays. Further, an individual user's role may consist of a mixture of activities under different circumstances. There are also different views or theories of supervision that focus on different contexts and activities. This variety makes any discussion of supervision complex and potentially confusing. Productive discussion requires fractioning supervision into its different subtypes and studying their different characteristics and support requirements.

Here, we describe one specific type of supervision, which we call multi-tasking supervision. It is increasingly important in command and control, as well as other tasks, as supervisors are increasingly required to supervise more systems and automation.

Multi-tasking supervision involves rapidly assessing multiple systems or situations and determining where attention is needed moment by moment. The US Navy position of Tactical Action Officer (TAO) provides a useful example. While not the captain of the ship, the TAO has real-time responsibility and decision-making authority. TAOs must continually monitor, albeit at a high level, a wide variety of ship systems including combat, communications, and engineering. They also monitor all mission areas including air, surface, and subsurface warfare, as well as support for littoral operations (occurring on or near land) including strikes, reconnaissance, and mine countermeasures. Consequently, TAOs must make rapid situation assessments and switch tasks frequently to direct their attention where it is most required at any given moment.

First, we place multi-tasking supervision within the context of other types of supervision in order to help us understand its objectives and unique cognitive demands. Then we consider the types of tasks and work environments that are appropriate for multi-tasking supervision. Finally, we discuss computer display concepts for supporting it.

MULTIPLE TYPES OF SUPERVISION

Multi-tasking supervision. The objective of multi-tasking supervision is to maintain a cursory, high level awareness over a set of tasks or systems and be able to direct attention and decision making resources to any task as needed. Multi-tasking supervision is a metacognitive process in the sense that supervisors are scanning and deciding where to direct their attention. In some situations, supervisors might be solely responsible for all tasks. In other situations, there might be primary operators for each task, and supervisors are responsible for presiding over all of them as well as performing their own duties. In either case, supervisors might be engaged in a variety of tasks over time, but periodically scan all areas under their responsibility.

A situation might require attention because of an external threat, such as a threatening aircraft in the context of military airspace monitoring. Or, a situation might require attention because of imminent tasks that will create high workload, such as UAVs approaching their reconnaissance targets. Therefore, situations demanding supervisory attention are not necessarily threatening or dire. We have chosen the term "deviant" to express their out of the ordinary quality that requires attention from a supervisory authority. In terms of Endsley's model of situation awareness (Endsley, 1995), the multi-tasking assessment would focus on the meaning of the deviation (level 2 SA) and perhaps its implications for the future (level 3 SA). The details of the situation (level 1 SA) might be reserved until the supervisor fully attends to the task.

The scanning would necessarily need to be quick so that supervisors would still have the majority of their time available for decision making. The moniker "four-second supervisor" in the title is based on an empirical measurement of assessment time using multi-tasking display tools (St. John & King, 2010), summarized below. The frequency of these rapid scans depends on a number of factors, including the rate of change of information in the supervised tasks, the cost of gathering that information, and the value or consequence of that information (e.g., Moray, 1984).

Following the assessment, supervisors may decide to access a primary task display to fully assess the situation, troubleshoot, and intervene. Because operational situations and displays can be complex, supervisors may need significant time to come up to speed on a situation. This task engagement time is related to the concept of task switching (e.g., Altmann & Trafton, 2007; Monsell, 2003). Each switch to a new task imposes a cost in terms of time and mental effort to engage and "get up to speed" with the next display and task.

Accordingly, multi-tasking supervision should benefit

from displays that provide summary information that can be assessed quickly to minimize information gathering costs. Additionally, multi-tasking supervision should benefit from displays that provide entry into the deeper semantics of the deviant issues in the situation and minimize the switching cost of engaging with the primary task displays. We call this property of the initial supervisory assessment "leading-edge SA." It is related to the more general notions of visual momentum (Woods, 1984) and entry points (Kirsh, 2001), for providing information about the relationships among displays. In this case, the multi-tasking supervision display can ease the transition to primary task displays. Billings (1997, p. 97) described a similar use for "synoptic displays" in aircraft cockpits.

Further, a lead-edge display might itself serve as a context for collaboration and consultation, since it describes the situation at a high level without the cluttering detail that may only be appropriate for primary operators.

Supervisory control. One common view of supervision that also provides an overarching framework is supervisory control. "The term supervisory control indicates that one or more human operators are setting initial conditions for, intermittently adjusting, and receiving information from a computer that itself closes a control loop" (Sheridan, 1987, p.1244). The supervisors' activities are divided into five stages: plan, teach, monitor, intervene, and learn. The monitor-intervene stages constitute an inner loop of control that nest inside the longer-spanning outer stages.

The inner loop of monitoring and intervening bears a resemblance to the monitoring and intervention of multi-tasking supervision, though multi-tasking supervision emphasizes monitoring multiple distinct systems and deciding when and where to intervene, though not specifically on how to intervene. Supervisory control, on the other hand, focuses more on aspects of control of automation or a robot from an indirect, supervisory point of view, for example through high-level "delegation" of objectives (Parasuraman et al., 2005).

Since supervisory control involves both monitoring and control, it should also benefit from summary displays and the leading-edge SA concept for guiding supervisors into a problem, particularly if the user is controlling multiple robots or systems.

Knowledge-driven monitoring. Another view of supervision is knowledge-driven monitoring. Vicente, Mumaw, and Roth (2004) describe the case of nuclear power plant operators supervising the automated operation of a plant, "most of the time, the operators work in a supervisory control mode (Sheridan, 1987), monitoring the status of the plant while it is being controlled by automation. This task is challenging because operators are not actively in the loop controlling the plant. Nevertheless, they must try to be attentive and make sure that the plant and automation are both behaving as intended and as expected" (p. 362). This form of supervision applies to situations where the operator/supervisor is engaged in an "active construction process" with the system, confirming expectations about plant state, pursuing unexpected findings, checking for problems considered to be likely, validating initial indications, and determining an appropriate reference for a specific indication. This form of supervision is very hands-on and engaged, and closer to what might be considered the role of an operator. Yet these operator/supervisors are indeed monitoring and interacting with highly automated systems. They monitor the closed-loop functioning of the automated system. In contrast, multi-tasking supervision has a less constructivist character. Multi-tasking supervision does not involve actively investigating system states or developing new indicators on the fly, as does knowledge-driven monitoring (Vicente, Mumaw, & Roth, 2004, p. 374). Rather, multi-tasking supervision involves periodically, and rapidly, scanning established high-level indicators for signs of deviation.

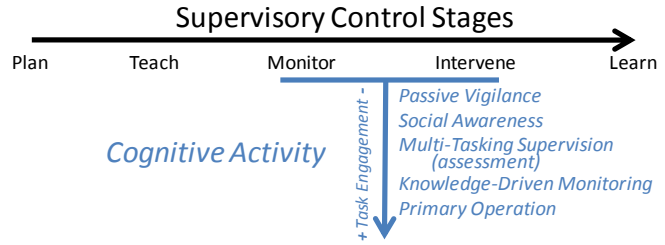

Figure 1. Alternative types of supervision. Within the monitoring stage, different types of supervision vary in their objectives and levels of task engagement.

Knowledge-driven monitoring might still benefit from summary displays, particularly for disparate subsystems. Yet the investigative nature of this form of supervision implies the need for more detailed displays of system variables. It is likely that operator/supervisors perform a mixture of activities over time. Their activities sometimes involve multi-tasking supervision by way of quickly scanning summary displays, and they sometimes involve delving into primary displays to assimilate details of system states and trouble shoot.

Adaptive automation. The concept of adaptive automation or adaptive function allocation is to monitor operators for fluctuations in workload and automate tasks when workload exceeds what the operator can manage. Situational variables, such as the number of pending or upcoming tasks, and operator performance are often measured, as are physiological indicators of workload or stress (e.g., Byrne & Parasuraman, 1996; Kaber & Endsley, 2004; Scerbo, 2005).

From the supervising automation's perspective, the automation may in fact monitor multiple individuals for high workload, but monitoring can be continuous and detailed, so there is no sense of the periodic, high-level scanning or deciding where to attend that is at the core of human multi-tasking supervision. From the user's perspective, the adaptive changes in automation level and task allocation will move the user between primary operation and supervision and perhaps between single and multiple tasking, but the emphasis is on manipulating the user's workload rather than supporting the user's multi-tasking scanning and attending decisions.

Passive vigilance. Passive monitoring of system parameters or even summary displays represents another type of supervision. Passive monitoring is noted for increasing user complacency, resulting in loss of situation awareness and automation mode, and loss of manual skill (e.g., Endsley & Kiris, 1995, Parasuraman & Riley, 1997; Sarter & Woods, 1995). Mumaw et al. (2000) noted a distinction between

passive monitoring and the active engagement of knowledge-driven monitor that they observed among power plant operators. Similarly, multi-tasking supervision does not involve constant passive monitoring of key indicators, or exhaustive scanning of all system indicators, to determine system status. Rather, it involves periodic and rapid scanning of the key indicators for each system, specifically to determine whether more detailed attention is warranted. Nonetheless, support concepts for multi-tasking supervision and vigilance are similar, such as peripheral or ambient high-level status displays (Matthews, Rattenbury, & Carter, 2007; Watson & Sanderson, 2004).

Social awareness. Social awareness is an important topic in collaborative work research. While social awareness is not supervision, per se, it bears some resemblances to multi-tasking supervision. "Social awareness helps to judge how to engage in a cooperative effort. For example, nurses spend time on maintaining a mental picture of where relevant physicians are, their current workload, and their schedule for the day. When a nurse needs to contact a physician, this picture helps her decide who to contact, when, how, and where" (Bardram & Hansen, 2004, p. 192).

While the objective of social awareness is different from the objective of multi-tasking supervision, it shares the activity of monitoring indicators from multiple colleagues or tasks in order to determine when, or if, to intervene. Hence, both would benefit from displays that provide this information readily and in a concise form for quick scanning. However, there is no real need for leading-edge SA, since the goal of the interruption is to introduce a new topic rather than intervene with an existing topic.

Summary. There is a broad variety of supervision types that vary in their objectives and cognitive activities. Multi-tasking supervision is a distinct type, though there is some overlap among types. Multi-tasking is more supervision than control in the sense that it focuses on high-level monitoring. It involves rapid scanning and sampling from multiple systems, and quick assessments to determine where to focus additional attention. Accordingly, multi-tasking supervision should benefit from well-organized, high-level summary displays that support quick assessments and leading edge SA.

WHEN, WHERE, AND HOW

Where is multi-tasking supervision appropriate and feasible? Responsibility over multiple systems or tasks is the first and most defining criterion for multi-tasking supervision. The TAO position represents one type of multi-tasking situation in which there is a supervisor who monitors multiple tasks but rarely takes primary control over any task. Other types include an operator who monitors and takes primary control of multiple tasks, and a team of operators who individually have primary responsibility for one task, but who monitor and share secondary responsibility for teammates' tasks.

The second multi-tasking supervision criterion is that the equipment and physical situation should support task switching. For example, a TAO's multi-tasking would be facilitated by a workstation that presented summary displays for all ship systems and mission areas. If, on the other hand, all workstations were dedicated to single tasks, then multi-tasking would require physically moving among workstations. While such a situation would not preclude multi-tasking, it would make it more difficult and limit the efficiencies that could be obtained from a more flexible display system.

Multi-tasking Supervision Criteria

- **Multi-tasking Situation**
  - Supervisor of multiple tasks
  - Operator of multiple tasks
  - Team of operators of multiple tasks
- **Rate of Change**
  - High rate of (cumulative) change
  - High demand for prompt attention
- **Task Environment**
  - Multi-tasking workstations
  - Convenient task switching

Further, for workstations that do support task switching, the task switching process must be efficient and usable. Calling up and configuring summary displays and possibly the primary task displays must be convenient and timely. Historically, many ship systems have been "stove piped" as dedicated, stand-alone systems and workstations. While this strategy provided a modular design that allowed new systems to be added over time without affecting other systems, the task environment did not support multi-tasking. Future concepts for combat operation centers, however, such as the Consolidated Afloat Networks and Enterprise Services (CANES; Turner, 2007) and the Multi-Modal Watch Station (MMWS; Osga, 2000), highlight the efficiencies of operational teams who share responsibility. These concepts build flexibility into the data streams, workstations, and teaming arrangements that enable multi-tasking and multi-tasking supervision.

The third multi-tasking criterion involves the rate of change of task information and the frequency of supervisory assessments (e.g., Moray, 1984). If task information changes only very slowly, then the need for multi-tasking is reduced. Multi-tasking supervision is required when the pace of changes among multiple systems is rapid, and immediate or relatively prompt attention to those changes is desirable. Further, as more systems are monitored, the cumulative rate of change increases. It is this cumulative rate of change that determines the need for multi-tasking support. Finally, the rate of change within a system or mission area may change dramatically under different operational conditions and make multi-tasking more or less desirable at different times.

As more systems are developed to support multi-tasking and task switching, the number of suitable environments for multi-tasking supervision will increase. Similarly, as better support concepts are developed and implemented, they will afford more multi-tasking supervision in more situations. Given the common need to reduce manpower costs, multi-tasking may become much more prevalent in a variety of situations.

Support concepts. There are a wide variety of support concepts for supervision including alert messages, status indicator boards, and summary system displays of many sorts. Most of these concepts are inappropriate for multi-tasking supervision, and care must be taken to design appropriate support that matches the cognitive activity of multi-tasking supervisors.

Many Navy command and control systems provide alert text messages that describe a wide variety of major and minor system malfunctions. It is our contention, however, that alert messages are inappropriate for multi-tasking supervision for two reasons. First, alerts are text, which requires more time to interpret than simple indicator lights and labels. Second, alerts describe specific issues and arrive separately, and both of these characteristics require the supervisor to integrate information over time and across subsystems. Interpreting and integrating multiple alerts presented in near-random order could require significant time and effort that is inappropriate for a multi-tasking scan.

Aircraft cockpits and factory control rooms are notorious for presenting tens or hundreds of status indicators (Sheridan, 1987, p. 1254). Methods for organizing and designing the indicators are well known, and users are taught to develop scan patterns to review them. Scanning these many indicators, however, still may require significant time, and it may require some interaction from the supervisor to call up various displays. Additionally, as with alert messages, the supervisor is required to integrate the indicators to assess the overall status of the system or systems.

There are also many types of summary displays. The ecological interface design (EID) community has developed a tradition of complex displays that provide a sophisticated graphical summary of a system's state. One guiding design principle of EID is to "make available overview displays by which 'free running' routines can be monitored by fringe consciousness" (Rasmussen & Vicente, 1989, pg. 524). This appeal to monitoring via "fringe consciousness" sounds akin to multi-tasking supervision. However, these displays tend to be quite complex in order to be comprehensive (e.g., Jamieson, 2002), and they may require substantial time to parse and assimilate. In fact, they are often meant to be used more as primary task displays for monitoring and problem solving than the quick scans performed during multi-tasking supervision, "[EID displays are meant to] represent the work domain in the form of an abstraction hierarchy to serve as an externalized mental model for problem solving" (Vicente, 2002, p. 64). The synoptic displays in aircraft cockpits (Billings, 1997, p. 97) are another example of summarized graphical displays that may be closer in content to multi-tasking supervision displays. However, these displays also tend to be used for troubleshooting rather than quick scans to assess where trouble might be brewing.

Of course, there is a fine line between primary system displays and summary displays. Operators need summaries and overviews just as do supervisors. The distinction drawn here involves both the objective of the multi-tasking supervision display, to scan for deviant situations, and the cognitive activity, quick scans across multiple systems.

Design method. We have been developing a systematic approach to the design of multi-tasking supervision displays. The procedure begins with a focused knowledge elicitation interview with subject matter experts to identify key deviant issues for the system or situation. The resulting list of factors is then taxonomized and aggregated to extract a set of six to ten key indicators. The idea is that the number of indicators must be manageable for a quick scan, while providing an aggregate of system or situation deviations. The supervisor can then use this information to quickly assess whether the system or situation requires more attention.

These indicators can be represented simply as a set of labeled lights presented in a small window near the center of the supervisor's workspace. Importantly, the indicators are not summarized into a master alert. Our rationale is that a master alert requires a function to aggregate the indicators into a situation assessment. This function is clearly an expert judgment that we feel is better left to the experts, whereas the indicators themselves are more objective.

To illustrate the content of the indicators and to support leading-edge SA, we supplemented the indicators with a simplified map representation. This map provides additional semantics and context for the indicators. Indicators that cannot be reasonably illustrated on a map may be illustrated in other representations, such as a Gantt chart for timeline problems. Supervisors can start with the indicators, then move to the illustrations. These illustration can be used to gain a richer yet still high level understanding as well as guide supervisors to the locations of deviant issues if they move on to a primary display to intervene or troubleshoot. This sequence provides a progressive deepening into the situation and problems.

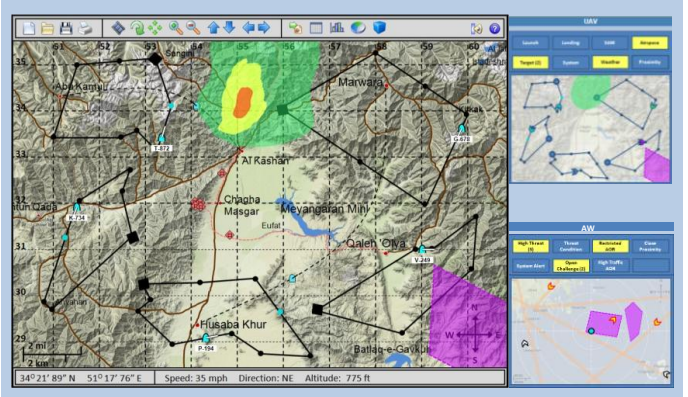

Figure 2. A primary task display (left) and two multi-tasking supervision displays (right). The primary display shows a set of five UAVs flying routes over complex terrain. The top multi-tasking supervision display summarizes the UAV situation with key indicator lights and a simplified map indicating the locations of UAV problems. The bottom multi-tasking supervision display summaries another mission area, air warfare.

An initial evaluation. This account of multi-tasking supervision and how to support it will require a series of experiments to test and refine the concept and displays and document its value as a human factors construct. We recently conducted an initial study to measure the impact of multi-tasking supervision displays on the speed with which supervisors can make situation assessments (St. John & King, 2010). 36 naïve participants were trained to view snapshots of situations from two different military domains, air warfare, in which operators monitor an airspace for threats, and UAV, in which operators monitor the flight of multiple remote aerial vehicles. The participants were taught simple rules for evaluating the snapshots to determine if they were "deviant"

and needed closer attention. In one within-participant condition, participants viewed complex and cluttered primary task displays, while in the other condition, they viewed multi-tasking supervision displays that summarized the situations using indicators and a simplified map (see figure 2).

Participants required 14.5 seconds on average to assess the primary displays but only 4.0 seconds to assess the multi-tasking supervision displays. Participants were 72% faster with the supervision display, a 360% improvement. They were also more accurate when using the supervision displays, 91% correct with the primary displays and 99% correct with the supervision displays. Interestingly, the benefit was more pronounced for negative trials, in which the situation was normal, than for positive trials, in which the situation was deviant. The reason is that the negative-normal trials required an exhaustive search for issues while the positive-deviant trials could be terminated once the situation was deemed deviant. With the supervision displays, participants could quickly "dismiss" a situation as normal and move on. This benefit for negative-normal trials is important since it is likely that a large majority of assessments will be negative-normal. Facilitating these assessments is therefore highly valuable.

CONCLUSIONS

Multi-tasking supervision is characterized by rapid scanning across multiple distinct systems in order to determine which systems require more focused, in depth attention, what we call four-second supervision. Display concepts that rapidly convey system deviations should support this activity. Further, the displays should support a leading edge SA to help supervisors transition to and engage with specific systems and their primary displays.

In an initial experiment, we found that the benefits of such displays can be substantial. While these results are not particularly surprising, since the displays summarized the situation and circumvented the need to search the primary display, the size of the benefit was of high practical value. Bringing the assessment time down from 14 seconds to four seconds gives back to supervisors a significant amount of time during which they can focus on specific tasks. Supervisors may also be willing to scan more frequently and detect problems sooner because the costs of scanning are lower. Further experiments are required to test whether the displays can be used effectively in real-time to monitor multiple tasks, find problems more quickly, and devote attention more efficiently to where it is required.

As automation increases in command and control contexts, as well as many civilian contexts, the opportunity and need for multi-tasking supervision will grow substantially. Synergistically, as support concepts mature, more opportunities may become viable. Within the broad topic of supervision, multi-tasking supervision is a distinct type with distinct objectives, cognitive activities, and support requirements. Our goal here was to call out this type and describe support concepts to suite the demands of this activity.

ACKNOWLEDGEMENTS

The authors thank Maia Cook and Harvey Smallman for their thoughtful comments, and Linda Wu for conducting the study. Support was provided by the Navy SBIR Program Element (PE 0605502N).

REFERENCES

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin and Review, 14,* 1079–1084.

Bardram, J. E. & Hansen, T. R. (2004). The AWARE architecture: Supporting context-mediated social awareness in mobile cooperation. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work.* New York, NY: ACM.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach.* Mahwah, NJ: Lawrence Erlbaum.

Byrne, E. A. & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42,* 249-268.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32–64.

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. *Theor. Issues Ergonom. Sci., 5,* 113–153.

Kirsh, D. (2001). The context of work, *Human computer Interaction, 16,* 305-322

Matthews, T., Rattenbury, T., & Carter, S. (2007). Defining, designing, and evaluating peripheral displays: An analysis using activity theory. *Human-Computer Interaction, 22,* 221–261.

Monsell, S. (2003). Task switching. *Trends in Cognitive Science, 7,* 134–140.

Moray, N. (1984). Attention to dynamic visual displays in man-machine systems. In R. Parasuraman and R. Davies, Eds., *Varieties of attention.* New York: Academic, pp. 485-512.

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors, 42,* 36-55.

Osga, G. (2000). Task managed watchstanding - Concepts for 21st century naval operations. In *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

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 – Part A: Systems and Humans, 35,* 481-493.

Rasmussen, J., & Vicente, K. J. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man-Machine Studies, 31,* 517–534.

Sarter, N. B & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37,* 5-19.

Scerbo, M. W. (2005). Adaptive automation. In R. Parasuraman & M. Rizzo (Eds.), *Neuroergonomics: The brain at work.* Oxford, Great Britain: Oxford University Press.

Sheridan, T. B. (1987). Supervisory control. In G. Salvendy (Ed.), *Handbook of human factors*. New York, NY: Wiley.

St. John, M. & King, M. (2010). Evaluating a display for making multi-tasking supervision assessments. Techical report. San Diego, CA: Pacific Science & Engineering Group.

Turner, P. (2007). The CANES initiative: Bringing the Navy warfighter onto the Global Information Grid. *CHIPS – The Department of the Navy Information Technology Magazine*, Oct-Dec, 2007. Found on the web at http://www.chips.navy.mil/archives/07_Dec/web_pages/CANES.htm

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors, 44, No. 1,* pp. 62-78.

Vicente, K. J., Mumaw, R. J., & Roth, E. M. (2004). Operator monitoring in a complex dynamic work environment: A qualitative cognitive model based on field observations. *Theoretical Issues in Ergonomic Science, 5,* 359-384.

Watson, M., & Sanderson, P. (2004). Sonification supports eyes-free respiratory monitoring and task time-sharing. *Human Factors, 46,* 497–517.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21,* 229-244.