SPECIAL ISSUE: Innovations in Cognitive Engineering and Decision Making, Part II

Considering Etiquette in the Design of an Adaptive System

Michael C. Dorneich
Patricia May Ververs
Santosh Mathan
Stephen Whitlow
Honeywell Laboratories
Caroline C. Hayes
University of Minnesota

ABSTRACT: In this article, the authors empirically assess the costs and benefits of designing an adaptive system to follow social conventions regarding the appropriateness of interruptions. Interruption management is one area within the larger topic of automation etiquette. The authors tested these concepts in an outdoor environment using the Communications Scheduler, a wearable adaptive system that classifies users' cognitive state via brain and heart sensors and adapts its interactions. Designed to help dismounted soldiers, it manages communications in much the same way as a good administrative assistant. Depending on a combination of message priority, user workload, and system state, it decides whether to interrupt the user's current tasks. The system supports decision makers in two innovative ways: It reliably measures a mobile user's cognitive workload to adapt its behavior, and it implements rules of etiquette adapted from human-human interactions to improve human-computer interactions. Results indicate costs and benefits to both interrupting and refraining from interrupting. When users were overloaded, primary task performance was improved by managing interruptions. However, overall situation awareness on secondary tasks suffered. This work empirically quantifies costs and benefits of "appropriate" interruption behaviors, demonstrating the value of designing adaptive agents that follow social conventions for interactions with humans.

KEYWORDS: adaptive automation, augmented cognition, communication, decision-aiding, human computer interaction, information overload

---

ADDRESS CORRESPONDENCE TO: Michael C. Dorneich, Human-Centered Systems, Honeywell Laboratories, 1985 Douglas Dr. North, Golden Valley, MN 55422, michael.dorneich@honeywell.com.

Journal of Cognitive Engineering and Decision Making, Volume 6, Number 2, June 2012, pp. 243-265.
DOI: 10.1177/1555343412441001.

Introduction

In the past decade, researchers have begun to consider the role of etiquette in design of intelligent systems that interact with people, with the goal of enhancing joint human-computer performance in complex tasks (Bickmore, 2002; Hayes & Miller, 2010a; Horvitz, 1999; Nass, 2004; Parasuraman & Miller, 2004). Etiquette between people is not just a matter of being polite but more so about being *appropriate*—behaving in ways that others will understand and perceive to be correct in context. Hayes and Miller (2010b) define *human-computer etiquette* as a similar set of conventions that facilitate smooth and effective interactions between people and computers. The themes underlying this body of work are that if etiquette rules are guidelines for effective social interactions in a given cultural and situational context, then it is expedient for software designers to design computer assistants so that they conform to the etiquette expectations of their user population insomuch as much as it is possible (Hayes & Miller, 2010b). The goal of doing so is to reduce miscommunications between computers and their users, reduce the frustration people feel when interacting with a computer agent, and increase effectiveness in a family of tasks. In this article, we focus on one important area of etiquette, interruption management.

Although there is a large body of empirical and theoretical work focusing on etiquette and its role in interactions between humans (Brown & Levinson, 1987; Grice, 1975), there is relatively less work empirically validating the role and impact of etiquette in interactions between humans and computers (e.g., Miller & Funk, 2001). Specifically, in the area of human-computer interaction (HCI), etiquette research has been primarily focused on alarm management. For instance, Woods (1995) recognized that many alarms are poorly designed and become a nuisance when they distract the person from the task at hand to warn him or her of a condition that is irrelevant to the current situation. Often this distraction occurs during high-workload periods, as nuisance alarms require a shift of attention and reasoning, expending valuable resources during cognitively complex tasks. Accordingly, general guidelines and alarm philosophies have been developed for allocating attention, supporting decision making, and avoiding nuisance interruptions (e.g., Bye et al., 1992; Dorneich, Olofinboba, Ververs, Winchester, & Krishnamurthi, 2002; Errington, Reising, & Burns, 2009; Riley et al., 1999; Woods, 1995). Work has also been done to study the influences on alarm perception and response behavior, for instance, the influence of alarm relevance and reliability on human perceptions of the alarm relevance, urgency, and importance (Newlin, Bustamante, Bliss, Spain, & Fallon, 2006). The goals of this work were to explore the role of etiquette in the design of a system to manage interruptions in a complex military-combat monitoring task. This article makes two contributions: (a) It will describe how to design a system from an etiquette perspective to manage interruptions in a contextually appropriate manner, and (b) it will describe an empirical assessment of the positive and negative impact of interruption management on overall task performance.

It is not generally considered polite to interrupt. There are many practical reasons underlying this and other etiquette guidelines; interruptions can disrupt the flow of thought in a complex task and reduce overall productivity. Not only is it counterproductive to interrupt a critical task requiring great concentration, but it can be dangerous or even life threatening if that task is safety critical. Conversely, there are situations in which it is essential and appropriate to interrupt. For example, it is inappropriate to interrupt an airline pilot during landing with the baseball score but entirely appropriate to interrupt with the information that the landing gear has malfunctioned.

A good human assistant can be invaluable in helping a decision maker to manage interruptions. Similarly, computer assistants hold the potential to improve decision makers' performance by either (a) minimizing interruptions for noncritical information or requests or (b) redirecting attention (e.g., interrupting) at critical junctures to necessary information. Unfortunately, computer assistants typically lack the savvy of human assistants to know when it is appropriate to interrupt. This lack of savvy stems from two factors: (a) the computer's inability to understand what constitutes an appropriate interruption and (b) the inability to understand when the user is "busy." What type of interruption is appropriate may depend on factors such as task relevance, urgency, and interruption salience. The second factor involved in calculating when to interrupt has traditionally been the user's current workload and task needs (on the basis of the user's current actions) and the computer's understanding of the task and environment. Such approaches are often prone to error when the user's current cognitive state and workload are incorrectly assessed and end up annoying more than assisting by interrupting at the wrong time.

Researchers have been exploring more direct means for assessing humans' cognitive state using physical and neurophysiological sensors to assess brain activity, heart rate, skin conductance, and other indicators. In earlier work, we developed and tested a cognitive-state classifier that could assess when mobile users were busy by monitoring a combination of indicators of brain activity and heart rate (Dorneich et al., 2006; Dorneich, Whitlow, et al., 2007). *Busy*, in this case, means users' cognitive workload was so high that adding another task would cause overall performance to suffer. In this work, it was our goal to determine whether we could incorporate the cognitive-state classifier into an adaptive system aimed at enhancing task performance for mobile dismounted soldiers by managing interruptions.

Traditional human-automation systems fixed the roles of the human and the automation at the design stage. In these systems, adaptations during task execution, if needed, were always initiated by the humans, for example, adapting to the automation or commanding the automation to adapt. In contrast, the class of human-automation systems known as *adaptive systems* breaks from this traditional approach in also allowing the system to invoke varying levels of automation support in real time during task execution, often on the basis of its assessment of the current context. Typically, the system will turn on and turn off the adaptations, invoking them only as needed.

To help soldiers manage interruptions and maximize performance during very busy high-workload periods, we designed and built the Communications

Scheduler, which assists soldiers by deciding when *not* to interrupt. The decision of when to interrupt, and when not to interrupt, is a critical etiquette issue that is equally relevant in interactions between humans as it is in interactions between humans and computers. Researchers have examined etiquette in terms of Brown and Levinson's model (1987), which focuses on etiquette in spoken and text communications as a means to mitigate *face threat*, or the imposition placed on the listener by a speaker through the act of communicating (Wu & Miller, 2010). Brown and Levinson additionally state that "politeness . . . makes possible communication between potentially aggressive parties" (Brown & Levinson, 1987, p. 1).

In this work, our focus is not on face threats or verbal communications per se but on the impact interruptions have on task performance, including the complex trade-offs between the costs and benefits that result in various conditions. We posit that etiquette guidelines associated with interruptions in human social interactions are based on a cooperative desire to maximize the performance of a group of actors who share a common set of tasks and goals. We found that maximizing performance is complex and highly situation dependent, requiring constant monitoring of the human's cognitive state and the ability to adapt as circumstances change. This nuanced understanding, which may allow either computers or humans to judge when to interrupt or hold back, is an important part of etiquette. In this view, etiquette is not just a way of mitigating potential hostilities in a competitive group; it is also a way of maximizing effectiveness in a cooperative group, regardless of whether it is a group of humans or a group of humans and computers working together.

Related Work

The Need to Manage Interruptions

Computer and communications systems are more often than not designed with little consideration of social conventions or the social impact of their actions (Fogarty, Hudson, & Lai, 2004). Constant interruptions are not only irritating, but the very tools introduced to increase productivity can actually rob workers of it. For instance, the common office worker spends up to 25% of his or her workday recovering from interruptions or switching between tasks (Spira, 2005), is interrupted or distracted on average once every 3 min (Zeldes, Sward, & Louchheim 2007), and is often interrupted at inopportune times (Chen, Hart, & Vertegaal, 2007). Inappropriate interruptions can increase errors, frustration, and stress and can reduce efficiency and decision quality (Chen & Vertegaal, 2004; Gillie & Broadbent, 1989; Iqbal, Adamczyk, Zheng, & Bailey, 2005; McFarlane & Latorella, 2002). There is a general desire to create automated systems that are less frustrating to users by endowing them with the same courtesies exhibited by human colleagues. Improved interruption management will only grow as an area of interest and concern in the field of HCI (Adamczyk, Iqbal, & Bailey, 2005; Hudson et al., 2003; McFarlane & Latorella, 2002).

The Need to Understand User Context and Workload

There has been research on ways to make systems more aware of the user's context to better time and tailor interruptions (Chen & Vertegaal, 2004; Fogarty et al., 2004; Mathan, Whitlow, Dorneich, Ververs, & Davis, 2007; Wickens & Hollands, 2000). For an adaptive system to decide when to interrupt, it would ideally have a sense of the value of the interruption relative to its costs. Understanding or having some sense of a person's cognitive workload (i.e., how busy or overloaded someone is) can be a reasonable basis for such a judgment and is often used (both alone and in conjunction with other methods) by humans. The ability to assess someone's cognitive workload would allow the designers of automated systems to take advantage of periods of high interruptibility to display incoming alerts, suggest that the user attend to different information, or switch tasks and to protect periods of low interruptibility by minimizing disruptions (Feigh, Dorneich, & Hayes, 2012). This knowledge can be derived in many ways. Computer systems often assess users' cognitive workload through indirect measures, such as their activities (Horvitz, 1999), focus of attention (Duchowski, 2003), performance (Hancock & Chignell, 1987; Scerbo, 1996), and interactions with devices (Shell, Vertegaal, & Skaburskis, 2003). System designers have historically found it difficult to accurately model user tasks, workload, and interruptibility, thereby limiting how effectively users could manage interruptions. Such systems often failed to gain user acceptance because they either filtered too much or too little information (Parasuraman & Miller, 2004). Improved interruption management continues to be an area of interest and concern in the field of HCI (Adamczyk et al., 2005; Hudson et al., 2003). However, in task domains in which sources that contribute to workload are unpredictable and difficult to track, users vary widely in experience, and the task environment is complex, a more direct measure of the user's cognitive workload may be needed. This assessment is especially important in high-workload domains in which the cost of inappropriate interruptions can be extreme (Mathan, Dorneich, & Whitlow, 2005).

Research has also focused on methods to directly measure cognitive workload and attention via physiological and neurophysiological sensors (Dorneich, Mathan, et al., 2007; Kramer, 1991; Pope, Bogart, & Bartolome, 1995; Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2000). Neurophysiologically and physiologically based assessment of cognitive state has been captured in several different ways, including measures of heart activity via electrocardiogram (ECG) and brain activity via electroencephalogram (EEG), functional near-infrared (fNIR) imaging, and galvanic skin response (GSR). Although direct measurement of cognitive workload may not be the way humans detect interruptibility, this method allows a computer to accomplish the same function with a different means. One of the most promising approaches is the use of EEG to measure cognitive state. EEG measures the electrical activity on the scalp associated with neuronal firing in the brain. It is a fast signal, with excellent temporal resolution, and therefore is a moment-to-moment measure. Research has shown that EEG activity can be used to assess a variety of cognitive states that affect complex task performance, such as working memory (Gevins & Smith, 2000), alertness (Makeig & Jung, 1995), engagement (Pope et al., 1995), executive control (Garavan, Ross, Li, & Stein, 2000), visual information processing (Thorpe, Fize, & Marlot, 1996), and target identification (Mathan et al., 2006). These findings point to the potential for using EEG measurements as the basis for driving adaptive systems that demonstrate a high degree of sensitivity and adaptability to human operators in complex task environments.

The Need to Optimize Performance

Several examples of neurophysiologically triggered adaptive automation have been developed to help optimize the performance of stationary and mobile operators. Such systems include an fNIR-based information display system for an unmanned air vehicle ground-control station operator (Snow, Barker, O'Neill, Offer, & Edwards, 2006) and an EEG-, ECG-, and GSR-based weapon control system for a Tactical Tomahawk operator (Tremoulet et al., 2006). In earlier work, we demonstrated that adaptive scheduling of communications based on EEG-derived assessment of the readiness to process information resulted in a twofold increase in message comprehension and situation awareness (SA; Dorneich et al., 2006). These results highlight the potential benefits of a neurophysiologically triggered adaptive automation. Many of the limitations of model-based approaches are alleviated with direct estimates of the operator's cognitive state. Complex task models based on indirect measures are not required to estimate cognitive workload. In fact, with reliable, real-time measures of cognitive state, adaptive systems can be much more proactive and aggressive in managing user interruptions. However, such systems must be designed with consideration of human-computer etiquette, lest they fail to gain user acceptance because the design fails to recognize the costs associated with any automated intervention, despite the benefits.

The Communications Scheduler

Using lessons learned from human-human interruption etiquette, we designed our Communications Scheduler to help soldiers manage their workload and maintain performance on high-priority tasks. It does so by detecting when their workload becomes very high (with information from an EEG cap and ECG patch) and changing the way their radio messages are presented to them. If the soldier has a high workload, the system will escalate the alert of an incoming high-priority radio message or divert low-priority radio messages to his or her PDA as text messages.

The Communications Scheduler operated on a computer system that was portable and mounted on the body so that mobile soldiers could carry the system with them in the field. The computer was carried in a backpack, and sensors and other equipment were mounted on the head and body. A cognitive state classification system monitored the soldier's cognitive state using EEG and ECG sensors. All sensors were connected wirelessly to the computer in the backpack.

EEG signals were preprocessed to remove eye blinks, power drifts, muscle artifacts, and all activity in frequency bands not generally associated with cognitive activity. The power spectral density (PSD) of the EEG signals was estimated with the Welch method (Welch, 1967) and was integrated across five well-established frequency bands: 4–8 Hz (theta), 8–12 Hz (alpha), 12–16 Hz (low beta), 16–30 Hz (high beta), and 30–44 Hz (gamma) (e.g., Gevins, Smith, McEvoy, & Yu, 1997). Estimates of spectral power formed the input features for a pattern classification system to assess the likely cognitive state (high or low workload). The classification process relied on probability density estimates derived from a set of spectral samples gathered during training with tasks that were as close as possible to the eventual task environment. Participants were trained prior to each day's trials to establish a baseline for the classifier. We have tested this approach for monitoring cognitive state for users in simulated environments and for mobile users in field environments. We were able to achieve overall classification accuracy in the 70% to 95% range for mobile participants. For more details on the specifics of the classification approach, see Dorneich, Whitlow, et al. (2007).

In military field operations, commanders are expected to carry out multiple tasks simultaneously and to perform them all well. Their lives, and the lives of those whom they command, depend on it. The commander must keep track of the positions of his or her own troops and other friendly troops, plan the next move, watch for civilians and enemy soldiers, and report those locations to others. Additionally, the commander may be interrupted sporadically by anything from enemy fire to requests from central command for information. Communication about most of these activities will occur through the radio. Some of the radio messages may be directed to a particular commander, whereas others may concern other units. However, it is still important for the commander to monitor such messages to maintain awareness of the location and actions of those friendly units to avoid firing on them.

In this domain, there is typically a cost associated with interruptions, but that cost may be repaid many times in some circumstances, for example, if that interruption is a warning that a sniper has been spotted in the vicinity. Conversely, there is also a cost associated with minimizing interruptions, although doing so may be necessary at critical times to achieve focus. The cost may be a loss of SA about other tasks or reduced performance on those tasks. The cost and benefits of computer decision-support tools that either create or minimize interruptions must be carefully weighed when designing such tools and when deciding when to and where to apply them.

System Design

The design of the Communications Scheduler was inspired by common etiquette rules in human interactions. As a third party in radio communications, it is usually most appropriate for the Communications Scheduler to do nothing—simply pass messages through and not interfere in synchronous human-to-human radio communications. However, much like a human assistant who can see that someone is becoming overwhelmed by his or her task load and proactively offers help, the

Figure 1. The Communications Scheduler includes an electroencephalogram-based cognitive monitoring system to assess workload, a handheld radio for aural messages, and a handheld PDA for text messages.

Communications Scheduler steps in to offer assistance when the person's workload reaches a critical level. People generally pick up on various verbal and nonverbal cues to decide when to step in or remain silent. The Communications Scheduler relies on the real-time assessment of operator workload to make this judgment.

The Communications Scheduler, shown in Figure 1, is designed to help soldiers manage their workload and maintain performance on high-priority tasks. It does so by detecting when their workload becomes very high (through the EEG cap and ECG patch) and by diverting low-priority radio messages to their PDA as text messages. The intention is to lower workload by temporarily removing low-priority messages to the PDA to allow soldiers to concentrate on more important messages and tasks. High-priority messages continue to be delivered as radio messages. For high-priority messages, a tone plays once before they are presented. If the system detects that the soldier is highly engaged in a task, it will play the tone more loudly and more saliently once before the message is presented. Low-priority messages do not have a tone associated with them. The text messages sent to the PDA are available immediately to the soldiers if they want to look at them, but they do not have to do so. Most important is the fact that text messages are less salient or "attention grabbing" than radio messages. By converting the low-priority messages to a less salient form, it allows soldiers to focus on the high-priority messages during high-workload periods.

To enable the Communication Scheduler to determine which messages to present as radio messages and which to divert to the PDA, all messages had a priority assigned to them a priori by the scenario designer. The priority (or criticality) of the message was determined by many factors, such as the message source (e.g., soldier's direct commander, another commander, a fellow soldier) and the

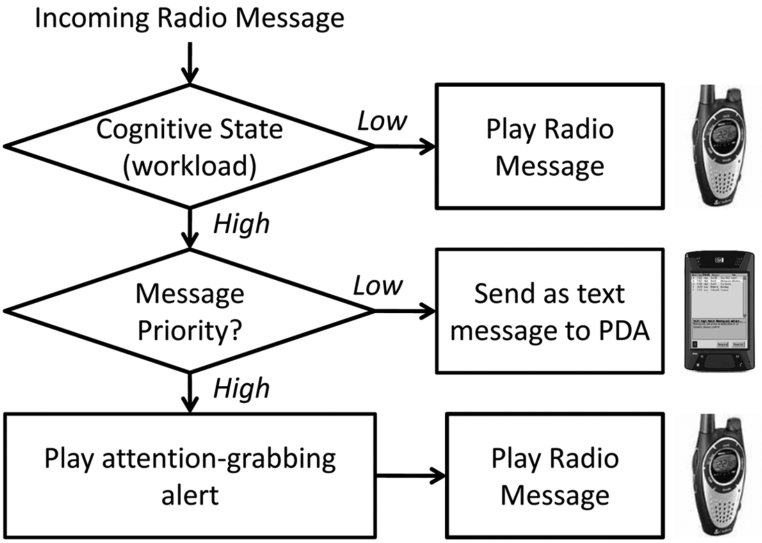


Figure 2. If user's workload is high, the Communications Scheduler interrupts with an alert if message priority is high or defers the message as text to the PDA if the priority is low.

relevance of the content. High-priority items were typically mission critical and time critical. It was imperative that they should receive the soldier's attention as soon as they arrived. Low-priority messages were not critical, although they still may have been important. The logic of the system is illustrated in Figure 2.

The Communications Scheduler mitigation was invoked ("turned on") when the participant's workload was measured as high. However, the question of when to "turn off" the mitigation takes into account more than just the moment-to-moment workload. If the Communications Scheduler was invoked, and subsequently the workload was then lowered below the threshold used to trigger the initial message deferral, the Communications Scheduler continued to defer messages (i.e., it does not turn off). Rather, the mitigation was turned off only when workload was lowered, the soldier had "caught up" by reading all the deferred messages, and the soldier had clicked a *Messages Read* button (Mathan et al., 2005). Deferring communications on the basis of only moment-to-moment fluctuations in workload can be confusing and disruptive. Again, automation etiquette played a role, since the Communications Scheduler needed to be invoked (and gracefully withdrawn) in such a manner that it would not cause confusion or induce unwanted oscillations in workload because of unpredictability. Messages could be misinterpreted without surrounding context if they were to be played in audio modality after their predecessor messages had been deferred to the PDA (and remain unread for a period of time). Thus the logic of when to turn off the mitigation included both workload and the reading of deferred messages.

The Communications Scheduler supports decision makers in two innovative ways: First, it reliably measures a mobile user's cognitive workload to adapt its behavior, and second, it implements rules of etiquette intended for human-human interactions to improve HCI.

Evaluation

The effectiveness of rules of etiquette is backed up by evidence showing that the rules of interruption are cost beneficial (McFarlane & Latorella, 2002)—that they maximize the information-processing resources of the group (in this case, the group is the soldier and the Communications Scheduler) and minimize the costs (e.g., loss of SA). Thus, we evaluated the Communications Scheduler in a field experiment to assess its appropriate use and potential cost and benefits.

In this experiment, we investigated the costs and benefits of *minimizing* interruptions during field operations in a variety of task load conditions. We did so by evaluating the Communication Scheduler's impact on multiple aspects of the mobile soldier's task performance during both high-and low-task-load conditions. Our hypotheses were the following:

1. During high-workload times, use of the Communications Scheduler would enhance performance on high-priority tasks while not degrading performance on other tasks.
2. During low-workload times, there are costs associated with the Communications Scheduler that make it inappropriate to leave it on all the time.
3. If the etiquette of the Communications Scheduler was appropriately designed, the participants would find their interaction with it acceptable.

Operational Scenario

The evaluation was held outdoors in a field of roughly 6,500 m^2, which consisted of primarily open grassy areas with some tree cover and forest in other areas. Each participant played the role of a platoon leader and was part of a larger simulated military company. A company is composed of roughly 62 to 190 soldiers. A platoon is composed of roughly 16 to 44 soldiers, grouped into squads of 9 to 10 soldiers each (Powers, 2010). The commander listened to messages from "soldiers" that were actually prerecorded audio messages, each with a different voice, rather than live confederates in the experiment. The computer played each message to the participant according to a script.

Each participant was the leader of the Red platoon, and his call sign was "Red-6." Each participant was responsible for managing three squads while reporting to his company commander. The squad leaders' call signs were "Red-6-1," "Red-6-2," and "Red-6-3." Figure 3 illustrates the command hierarchy and shows examples of some of the task-related incoming messages heard by the participant and outgoing messages from the participant.

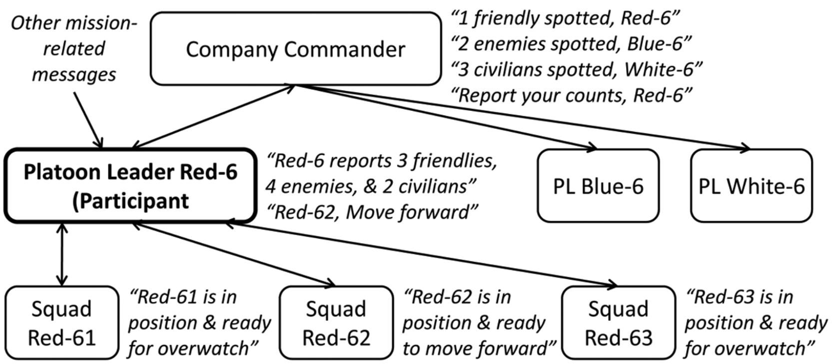


Figure 3. The command structure and messages originating at each level.

Tasks

We asked soldiers to perform five tasks simultaneously, because prior studies revealed that people can in fact manage many tasks simultaneously without becoming cognitively overloaded, especially if the tasks are highly practiced (Dorneich, Whitlow, Ververs, Carciofini, & Creaser, 2004; Dorneich, Ververs, Mathan, & Whitlow, 2005). Additionally, we have observed that the more artificial the situation, the more difficult it becomes to "overload" a person interacting in that environment, possibly because artificial environments lack the detail, richness, and stimulation of a real environment. For example, we tested early versions of the Communications Scheduler in a simulated environment. It was very difficult to overload participants in this environment; there had to be 30 or more characters shooting at participants before it interfered with their ability to perform other tasks (Dorneich et al., 2004). Additionally, we have observed that the richer and more realistic the environment, the easier it becomes to overload participants. For example, in later studies, actual soldiers participating in a field experiment moved through a semiurban environment shooting soap bullets at each other, and it took relatively few (often fewer than three) simultaneous tasks to overload them enough to decrease task performance (Dorneich, Mathan, Ververs, & Whitlow, 2007).

The five tasks given to participants were as follows:

1. Navigation along a route
2. Keeping count of the number of civilians and soldiers sighted
3. Maneuver monitoring
4. Math interruption task
5. Maintaining awareness of the overall situation

All the tasks were designed to resemble military relevant tasks to create experimental conditions as realistic as possible, while still permitting easy collection of performance metrics assessing the speed and accuracy in completing the tasks. Finally, these tasks allowed us to vary the participant's cognitive workload by varying the rate at which messages were delivered or by increasing the number and complexity of the tasks. The following is a more detailed description of the five tasks used in the experiment.

Navigation. Participants were asked to walk along a simple, familiar, and circular route during each trial. The participants were mobile during the task so as to approximate real field conditions; moving in the richness of a real environment requires far more attention and cognitive resources than sitting at a computer and moving through a simulated world. Additionally, we wanted to demonstrate that a neurophysiologically driven system, such as the Communications Scheduler, could provide useful support to a mobile field soldier. We had already demonstrated that a combination of ECG (heart) and EEG (brain waves) could successfully measure the workload of a mobile soldier (Erdogmus et al., 2005); we were now putting the whole system together, and our aim was to assess whether useful support, based on physiological and neurophysiological workload measurements, could be provided to mobile soldiers in a field environment. There were no performance measures associated with this task.

Count maintenance. A simulated Company Commander broadcast radio messages to the three platoon leaders about the number of civilians, enemy soldiers, and friendly soldiers sighted. Two of the platoon leaders (White-6 and Blue-6) were simulated only as radio voices. The participant in the experiment played the role of platoon leader Red-6. The participant was asked to maintain a running total of civilians, enemies, and friendlies reported to Red-6 while ignoring the counts reported to the other two platoon leaders. This task required participants to attend to all messages in sufficient detail to determine whether it was directed at them or other platoon leaders and to keep counts in working memory until asked by the commander to report the counts. After they had reported a count, they could start again at zero. Performance on this task was measured by the number of correct counts reported by the participant for civilians, enemy soldiers, and friendly soldiers.

Maneuver monitoring. A common duty for a platoon leader is to orchestrate a series of maneuvers among the squads in his or her command. For example, in a *bounded overwatch*, the platoon moves toward its objective through a series of steps in which one squad moves while the other two squads protect the moving squad. In our study, participants were asked to keep track of the status of all three (simulated) squads in their command. Each squad would radio a message to Red-6 after it was in position for the maneuver. Two squads would report, "Ready for overwatch," and one squad, "Ready to move." Their reports would arrive in random order. When all three were in position, the participant radioed an order commanding the ready-to-move squad to move forward. This task required the participant to keep track of each of the three squads' status in working memory until all the squads were in position. Performance on this task was measured by the participant's accuracy in sending the correct team forward (e.g., the squad that reported itself as ready to move).

Math interruption. Participants were asked at random intervals to complete a simple math problem. Although this task is somewhat artificial, its cognitive demands are representative of the constant urgent requests platoon leaders receive in such circumstances, in which they may be asked to suddenly turn their attention away from the current tasks to address a specific urgent need requiring their problem-solving skills. Although all radio messages involved in the tasks were interruptions, the math interruption task was designed to be especially disruptive, requiring much concentration and focus. The math task interferes with rehearsal of information kept in working memory (such as counts of civilians and overwatch status) more so than interruptions from other types of radio messages. In this task, the PDA would beep to alert participants that they had a problem awaiting their solution. The participants had to acknowledge the alert, and then the problem would appear on the PDA. Participants had to solve the problem and then submit the solution. Performance on the math interruption task was measured in terms of the participant's speed in responding with the answer to the math problem and the accuracy of the answer.

Mission SA. During the high-task-load conditions, participants received "mission messages" pertaining to their current location, the status of various teams and personnel, the overall situation, and their surroundings. These messages were all low-priority messages. Performance on this task was measured by a three-question test administered at the end of each trial; participants were asked about the content of the messages that they received.

The high-priority tasks in this scenario were to maintain radio counts, to monitor maneuvers, and to complete the math interruption task. These tasks were all assigned a high priority to prevent participants from favoring one to the other. The low-priority tasks were navigation and mission SA. When the EEG and ECG determined that the participant's workload was high, the Communication Scheduler diverted all radio messages associated with mission SA to the PDA as text messages. There were no messages associated with the navigation task.

Experimental Design

Independent variables. We used a 2 × 2 within-subjects design. There were two independent variables: mitigation (mitigated, unmitigated) and task load (high, low). *Mitigated* means that the Communications Scheduler was available but not necessarily "on." It would turn on only during times when the EEG and ECG sensors indicated that the participant's cognitive workload was high. *Unmitigated* means that the Communications Scheduler was "off" and no changes were made to messages. Task load was varied by changing the pace at which tasks had to be completed and the number of simultaneous tasks that had to be carried out. In this experiment, participants in the low-task-load condition received on average 3.8 messages per minute and were asked to carry out four tasks simultaneously

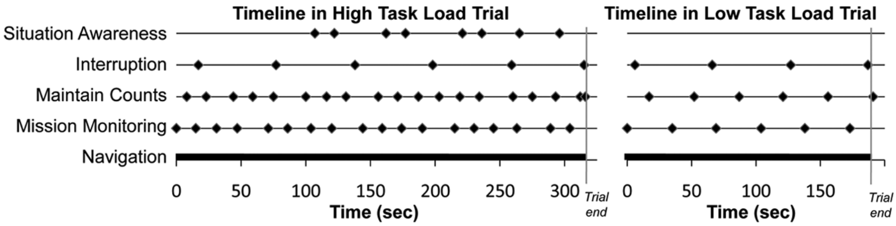


Figure 4. Timeline of message occurrence for four tasks. The navigation task is continuous.

(all but "mission SA"). In the high-task-load condition, participants received on average 8.7 messages per minute and were asked to carry out all five tasks simultaneously. In all conditions, participants were interrupted twice per minute with math interruption tasks. Figure 4 illustrates the occurrence of each task stimuli on a common timeline.

Note that we have used task load, rather than workload, as an independent variable, because one cannot directly manipulate the participant's cognitive workload. One can only manipulate the task load (e.g., number, complexity, and pace of tasks) with the expectation that it will affect cognitive workload. We used our previous experience to set the high and low task loads to levels that we anticipated would result in high and low workloads. A high workload is one that typically degrades task performance. Additionally, in our analysis, we experimentally confirmed that the task loads resulted in the predicted cognitive workloads.

Participants. There were 8 male volunteers who participated in this evaluation. They ranged in age from 21 to 42 years of age with an average age of 29.5. No participants had military experience.

Procedure. After signing a consent form, participants put on the equipment: backpack, EEG cap, ECG patch, radio (clipped to shoulder), and handheld PDA. They received information through the radio and PDA, and they responded through the radio. Participants were then trained in the field. They first practiced each of the five tasks independently and then all tasks together, all without the Communications Scheduler. When performing the tasks together, participants were told that all high-priority tasks were of equal importance and not to favor one task versus the others. They trained until they reached an acceptable level of performance (better than 90% performance on all tasks). Next, the cognitive-state classifier was calibrated for each participant, as each person exhibits workload differently. First, the participant was put in a low-task-load condition, and the system recorded the EEG and ECG patterns he produced. The same was done for the high-task-load condition.

After the training period, the participants completed four trials, as shown in Table 1, for every combination of mitigation and task load: mitigation off, task

TABLE 1. Presentation Order of Mitigation in Experimental Trials

Task Load	Participant 1	2	3	4	5	6	7	8
High	U	U	M	M	U	M	U	M
Low	U	U	M	M	U	M	U	M
High	M	M	U	U	M	U	M	U
Low	M	M	U	U	M	U	M	U

Note. U = unmitigated; M = mitigated.

load low; mitigation on, task load low; mitigation off, task load low; and mitigation on, task load low. High-task-load periods were always followed by low-task-load periods because of the geometry of the routes in the field, but the order of the mitigated and unmitigated trials was randomized.

Dependent variables. Performance metrics in the form of percentage correct were collected for the count maintenance, maneuver monitoring, math interruption, and SA tasks. In addition, the time to respond to the interruption alert was measured. (See Table 2.)

After each trial, participants were asked to fill out a NASA Task Load Index survey (NASA-TLX; Hart & Staveland, 1988). Participants in the high-task-load condition additionally filled out a questionnaire to assess their mission SA. Participants took a postexperimental questionnaire after completing all trials to gather their opinions on usefulness and challenges in using the system.

Results

Given the small sample size, the data were non-normally distributed; therefore the nonparametric Wilcoxon matched-pairs signed ranks test was used to test for significance. Results are reported as significant for $\alpha = .05$ and trending significant for $\alpha = .10$.

Recall that participants' primary tasks were to maintain radio counts, monitor squads' maneuvering, and manage the math interruption task. The main manipulation of task load was significant, indicating that increasing the pace of the messages was effective in driving participant workload and performance. As messages were added to the high-task-load condition, performance on the maneuver monitoring, which required participants to closely monitor the radio communications of their squad leaders, was negatively affected by the increased load ($Z = -2.52$, $p = .01$). The effectiveness of the task load manipulation was confirmed by the NASA-TLX findings. Findings indicated that the high-task-load blocks resulted in higher perceived workload than the low-task-load blocks on five of the six indices: mental demand ($Z = -2.38$, $p = .02$), temporal demand ($Z = -2.52$, $p = .01$), performance ($Z = -2.38$, $p = .02$), effort ($Z = -2.52$, $p = .01$), and frustration ($Z =$

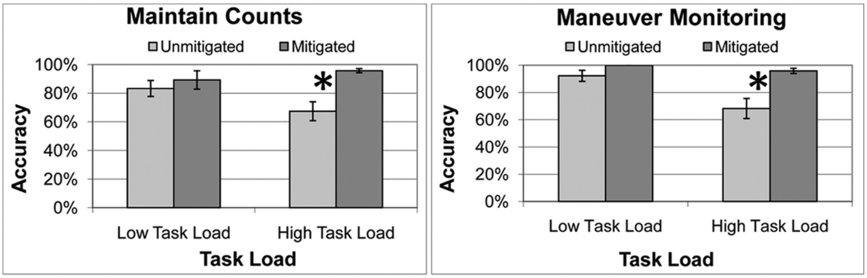
TABLE 2. Data Collected for the Experiment

Task	Data Collected	Results Reported	Frequency Collected
Maintaining counts	Counts of enemies, friendlies, and civilians	% correct	Multiple per trial
Maneuver monitoring	Number of times participants correctly sent the appropriate squad forward	% correct	Multiple per trial
Interruption task	Response time to interruption alert	Time	Multiple per trial
Interruption task	Number of correct answers to math problem	% correct	Multiple per trial
Mission situation awareness	Number correct to a three-question questionnaire	% correct	Once per high-task-load trial

–2.37, p = .02). Physical demand was the only workload measure not found to be significant, which was expected since the physical demand of navigating the course was not different in the high- and low-task-load blocks.

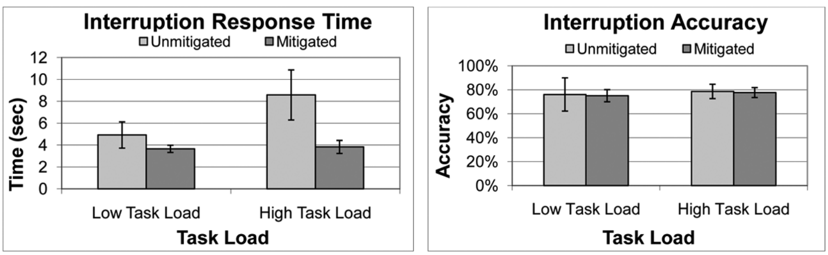
We analyzed the effectiveness of the communications scheduling mitigation by comparing the performance in the low- and high-task-load conditions. Results indicated that the availability of the Communications Scheduler significantly improved performance on each of the three primary tasks during high-task-load trials. Participants were more accurate when the Communications Scheduler mitigation was available, compared with the unmitigated trials for maintaining counts (Z = –2.38, p = .02) and maneuvering squads (Z = –2.54, p = .01) in the high-task-load condition. In the low-task-load condition, mitigated performance did not differ from unmitigated performance for the radio counts, although the mitigation approached significance on maneuver monitoring (Z = –1.63, p = .10). Figure 5 illustrates the means and standard errors for the 8 participants' responses on the count maintenance and maneuver monitoring tasks by task load and mitigation.

On the third primary task, involving the math interruption, participants responded almost 5 s faster in the mitigated condition than in the unmitigated condition, improving their response time from 8.6 s to 3.8 s in the high-task-load condition. This difference was large but not significant. In the low-task-load condition, the response time was approximately 1 s faster when mitigated by the Communications Scheduler, whereby this difference trended toward significant (Z = –1.83, p = .07). Figure 6 illustrates the means and standard errors for the math interruption task. The accuracy of the math task averaged greater than 75%, indicating that participants were devoting resources to the interruption. The accuracy was not affected by the presence of the mitigation in the low-task-load or high-task-load conditions. Note that because of data logging issues, only 4 of 8 participants' data were recorded for the low-task-load condition and 7 of 8 participants' data were used in the analysis of the high-task-load condition.



Tasks	Low Task Load Mean (Range) Unmitigated	Low Task Load Mean (Range) Mitigated	High Task Load Mean (Range) Unmitigated	High Task Load Mean (Range) Mitigated
Maintain counts	83.3% (61.4-100%)	89.2% (50-100%)	67.4% (34.4 – 97.2%)	95.7% (90-100%)
Monitor Maneuver	92.2% (75-100%)	100% (100-100%)	68.2% (29.2-91.7%)	95.8% (87.5-100%)

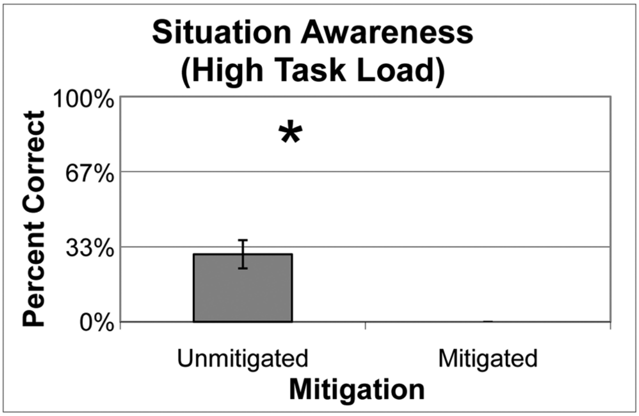
Figure 5. Accuracy for count maintenance and maneuver monitoring tasks. Stars indicate significant differences.



Tasks	Low Task Load Mean (Range) Unmitigated	Low Task Load Mean (Range) Mitigated	High Task Load Mean (Range) Unmitigated	High Task Load Mean (Range) Mitigated
Interruption RT	4.9 (2.9-8.4) sec	3.7 (2.8-4.4) sec	8.6 (2.0-18.9) sec	3.8 sec (1.7-6.5)
Interruption Acc	76.0% (75-100%)	75.0% (62.5-100%)	78.6% (50-91.7%)	77.6% (60-91.7%)

Figure 6. Interruption response time and accuracy.

Analyses of the secondary task of maintaining SA of the overall mission were conducted. Results indicated that the response to messages suffered in the mitigated versus the unmitigated condition with high task load. Low-priority messages were deferred to the PDA until the participants chose to review them as time allowed. In the unmitigated condition, participants scored only 30% correct on the SA questions. In the mitigated condition when the messages were deferred to the PDA, participants chose to ignore the messages in favor of responding to higher-priority tasks and therefore scored 0% when answering the SA questions. The difference between the high-task-load mitigated and unmitigated conditions was significant ($Z = -2.04$, $p = .04$). Figure 7 illustrates the means and standard



Tasks	High Task Load Mean (Range) Unmitigated	Mitigated
Mission SA	30% (16.8-33%)	0% (0 -0%)

Figure 7. Average number of questions correct in the mission situation awareness task. Stars indicate significant differences.

errors for the maintaining-SA task. Note that because of data logging issues, only 5 of 8 participants' data were recorded.

Discussion and Conclusions

The evaluation demonstrated a successful manipulation of workload with the modulation of the number of incoming messages in a fairly realistic military relevant environment. The postevaluation debriefing revealed that the "soldiers" developed a number of strategies to deal with the increased workload, such as visualization, finger counting, and chunking. However, nothing was as effective as the Communications Scheduler. The effectiveness of using a real-time Communications Scheduler in deferring interruptions enabled participants' ability to monitor multiple squads' movements while attending to the influx messages. The interruptions competed for the same working memory resources as the ongoing primary tasks. One participant noted that the "addition task [was] very effective at wiping out count memory." When the classifier detected that the participant's workload increased, the Communications Scheduler deferred the low-priority messages, enabling more attentional resources to react to interruptions. The Communications Scheduler successfully improved performance of the high-priority tasks during high-task-load periods to attend to the interruptions, which may be critical if those interruptions are life-threatening emergencies.

The advantages of the Communications Scheduler did not come without a cost. Participants' SA of low-priority messages suffered when they were deferred. A system designer simply cannot build a scheduler to defer all interruptions in the hopes to improve overall performance. In the end, SA for those deferred tasks suffers. Therefore, a more sophisticated interaction, based on the workload capacity of the human, can be much more effective and acceptable if tasks are deferred only when the humans are overloaded. It is in these narrowly defined, temporary situations that the costs are outweighed by the benefits, and an adaptation of the message scheduling is acceptable. When the workload returns to manageable levels (and the soldier catches up on deferred messages), the system ceases its interventions, as the costs are no longer acceptable.

To live and work in the modern world is to be interrupted. Interruptions are typically thought of as undesirable; however, they have both negative and positive consequences. Similarly, minimizing interruptions can also have both positive and negative consequences, as these experiments have demonstrated. We examined an adaptive system, the Communications Scheduler, aimed at assisting mobile dismounted field soldiers during crisis times by minimizing interruptions. We found that minimizing interruptions at critical times enables decision makers to maintain performance on the most critical tasks but degrades SA on issues that may be currently less pressing but equally important. Designers of intelligent systems need to be aware of both the cost and the benefits of systems that interrupt or manage interruptions.

The evaluation provided insights into interactions between humans; there are practical reasons underlying the social etiquette guidelines that surround interruptions. When a speaker interrupts or fails to interrupt, he or she may have competing consequences on the listener's effectiveness and SA, and these consequences must be kept in the proper balance. Furthermore, understanding when to interrupt and when to hold back is a complex, highly situation-dependent judgment requiring constant monitoring and attention. By better understanding the principles that enable humans to live and work together effectively, we may better design computers that can work effectively with people.

Acknowledgments

This research was supported by a contract with the Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Natick Soldier Center, DAAD16-03-C-0054, for which Capt. Dylan Schmorrow served as the program manager of the DARPA Improving Warfighter Information Intake Under Stress (AugCog) program and Henry Girolamo was the DARPA agent. The opinions expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Army Natick Soldier Center. Additionally, the authors would like to acknowledge the efforts of the Honeywell research and development team that supported the development of the prototypes, the running of the experiments, and the analysis of the data, including Danni Bayn, Jim Carciofini,

Janet Creaser, Trent Reusser, and Jeff Rye. Finally, the authors would like to thank the reviewers for their thoughtful comments and suggestions.

References

Adamczyk, P. D., Iqbal, S. T., & Bailey, B. P. (2005, September). *A method, system, and tools for intelligent interruption management*. Paper presented at TAMODIA, Gdansk, Poland.

Bickmore, T. (2002, November). *When etiquette really matters: Relational agents and behavior change*. Workshop held at the American Association of Artificial Intelligence Fall Symposia, North Falmouth, MA.

Brown, P., & Levinson, S. (1987) *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.

Bye, A., Kårstad, T., Nilsen, S., Barmsnes, K.A., Førdestrømmen, N., Valseth, A., . . . Follesø, K. (1992). *An Integrated Alarm System—A Concept Study.* OECD Halden Reactor Project (Technical Report, HWR-308). Halden, Norway: Institutt for Energiteknikk.

Chen, D., Hart, J., & Vertegaal, R. (2007). Towards a physiological model of user interruptability. In C. Baranauska et al. (Eds.), *INTERACT 2007* (Part 1, pp. 439–451). Berlin: Springer.

Chen, D., & Vertegaal, R. (2004). Using mental load for managing interruptions in physiologically attentive user interfaces. In *Proceedings of CHI 2004* (pp. 1513–1516). New York, NY: ACM Press.

Dorneich, M. C., Mathan, S., Ververs, P. M, & Whitlow, S. D. (2007). An evaluation of real-time cognitive state classification in a harsh operational environment. In *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting* (pp. 146–150). Santa Monica, CA: Human Factors and Ergonomics Society.

Dorneich, M. C., Ververs, P. M., Mathan, S., & Whitlow, S. D. (2005). A joint human-automation cognitive system to support rapid decision-making in hostile environments. In *Proceedings of the International Conference on Systems, Man and Cybernetics* (Vol. 3, pp. 2390–2395). Piscataway, NJ: IEEE.

Dorneich, M. C., Ververs, P. M., Mathan, S., Whitlow, S. D., Carciofini, J., & Reusser, T. (2006). Neurophysiologically-driven adaptive automation to improve decision making under stress. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 410–414). Santa Monica, CA: Human Factors and Ergonomics Society.

Dorneich, M. C., Whitlow, S. D., Mathan, S., Ververs, P. M., Erdogmus, D., Pavel, M., Adami, A., Pavel, & Lan, M. (2007). Supporting real-time cognitive state classification on a mobile participant. *Journal of Cognitive Engineering and Decision Making, 1*, 240–270.

Dorneich, M., Whitlow, S., Ververs, P. M., Carciofini, J., & Creaser, J. (2004). Closing the loop of an adaptive system with cognitive state. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 590–594). Santa Monica, CA: Human Factors and Ergonomics Society.

Dorneich, M. C., Olofinboba, O., Ververs, P. M., Winchester, W., & Krishnamurthi, S. (2002). *Alerting and notification of conditions outside the aircraft (ANCOA): Integrated alerting guidelines and conceptual design*. Technical Report prepared for NASA Langley Research Center under Contract NAS1-00107, Honeywell Laboratories, Minneapolis, MN.

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. London, UK: Springer Verlag.

Erdogmus, D., Adami, A., Pavel, M., Lan, T., Mathan, S., Whitlow, S., & Dorneich, M. (2005, March). *Cognitive state estimation based on EEG for augmented cognition*. Paper presented at the second IEEE EMBS International Conference on Neural Engineering, Arlington, VA.

Errington, J., Reising, D. V., & Burns, C. (2009). *ASM Consortium guidelines: Effective alarm management practices*. Minneapolis, MN: ASM Consortium.

Feigh, K., Dorneich, M. C., & Hayes, C. C. (2012). *A taxonomy of adaptive systems*. Manuscript under review.

Fogarty, J., Hudson, S., &Lai, J. (2004). Examining the robustness of sensor-based statistical models of human interruptibility. In *Proceedings of CHI 2004* (pp. 207–214). New York, NY: ACM Press.

Garavan, H., Ross, T. J., Li, S.-J., & Stein, E. A. (2000). A parametric manipulation of central executive functioning using fMRI. *Cerebral Cortex, 10*, 585–592.

Gevins, A., & Smith, M. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex, 10*, 829–839.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex, 7*, 374–385.

Gillie, T., & Broadbent, D. E. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research, 50*, 243–250.

Grice, P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). New York, NY: Academic Press.

Hancock, P. A., & Chignell, M. H. (1987). Adaptive control in human-machine systems. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 305–345). North Holland, Netherlands: Elsevier.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North Holland, Netherlands: Elsevier.

Hayes, C. C., & Miller, C. A. (2010a). *Human-computer etiquette: Cultural expectations and the design implications they place on computers and technology*. Boca Raton, FL: Taylor & Francis.

Hayes, C. C. & Miller, C. A. (2010b). Should computers be polite? In C. C. Hayes & C. A. Miller (Eds.), *Human-computer etiquette: Cultural expectations and the design implications they place on computers and technology* (pp. 1–12). Boca Raton, FL: Taylor & Francis.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In *Proceedings of the ACM CHI SIGGRAPH Conference on Human Factors in Computing Systems* (pp. 159–166). New York: ACM Press.

Hudson, S. E., Fogarty, J., Atkeson, C. G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J. C., & Yang, J. (2003). Predicting human interruptibility with sensors: A wizard of oz feasibility study. *Proceedings of CHI 2003* (pp. 257–264). New York, NY: ACM Press.

Iqbal, S. T., Adamczyk, P. D., Zheng, X. S., & Bailey, B. P. (2005). Towards an index of opportunity: Understanding changes in mental workload during task execution. *Proceedings of CHI 2005* (pp. 311–320). New York, NY: ACM Press.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple task performance* (pp. 279–328). Bristol, PA: Taylor & Francis.

Makeig, S., & Jung, T.-P. (1995). Changes in alertness are a principal component of variance in the EEG spectrum. *NeuroReport, 7*, 213–216.

Mathan, S., Dorneich, M., & Whitlow, S. (2005, July). *Automation etiquette in the augmented cognition context*. Paper presented at the 11th International Conference on Human-Computer Interaction (Augmented Cognition International), Las Vegas, NV.

Mathan, S., Whitlow, S., Dorneich, M., Ververs, P., & Davis, G. (2007, October). *Neurophysiological estimation of interruptibility: Demonstrating feasibility in a field context*. Paper presented at the 4th International Conference of the Augmented Cognition Society, Baltimore, MD.

Mathan, S., Whitlow, S., Erdogmus, D., Pavel, M., Ververs, P., & Dorneich, M. (2006). Neurophysiologically driven image triage: a pilot study. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 1085–1090). New York, NY: Association for Computing Machinery.

McFarlane, D. C., & Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction, 17*, 1–61.

Miller, C. A., & Funk, H. B. (2001, September). *Associates with etiquette: Meta-communication to make human-automation interaction more natural, productive and polite*. Paper presented at the 8th European Conference on Cognitive Science Approaches to Process Control, Munich, Germany.

Nass, C. (2004). Etiquette equality: Exhibitions and expectations of computer politeness. *Communications of the Association for Computing Machinery, 47*(4), 35–37.

Newlin, E. T., Bustamante, E. A., Bliss, J. P., Spain, R. D., & Fallon, C. K. (2006). The effects of relative system reliability and prioritization on alarm reaction patterns. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1675–1679). Santa Monica, CA: Human Factors and Ergonomics Society.

Parasuraman, R., & Miller, C. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the Association for Computing Machinery*, 47(4), 51–55.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology, 40*, 187–195.

Powers, R. (2010). *United States Army: Chain of command (organization)*. Available from http://usmilitary.about.com/od/army/l/blchancommand.htm?p=1

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology, 10*, 393–410.

Riley, V., DeMers, R., Good, M., Krishnan, K., Miller, C., & Misiak, C. (1999). *Crew-centered flight deck alerting* (NASA Tech. Rep. Contract No. NAS1-20219). Hampton, VA: NASA Langley Research Center.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37–63). Mahwah, NJ: Lawrence Erlbaum.

Shell, J., Vertegaal, R., & Skaburskis, A. (2003). EyePliances: Attention-seeking devices that respond to visual attention. In *Ext. Abstracts CHI'03* (pp. 770–771). New York, NY: ACM Press.

Snow, M. P., Barker, R. A., O'Neill, K. R., Offer, B. W., & Edwards, R. E. (2006). Augmented cognition in a prototype uninhabited combat air vehicle operator console. In D. Schmorrow, K. Stanney, & L. Reeves (Eds.), *Foundations of augmented cognition* (2nd ed., pp. 279–288). Arlington, VA: Strategic Analysis.

Spira, J. B. (2005). The high cost of interruptions. *KMWorld, 14*(1), 1, 32.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Tremoulet, P., Barton, J., Craven, R., Gifford, A., Morizio, N., Belov, N., Stibler, K., Regli, S. H., & Thomas, M. (2006). Augmented cognition for Tactical Tomahawk Weapons Control System operators. In D. Schmorrow, K. Stanney, & L. Reeves (Eds.), *Foundations of augmented cognition* (2nd ed., pp. 313–318). Arlington, VA: Strategic Analysis.

Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short modified periodograms. *IEEE Transactions on Audio and Electroacoustics, 15*(2), 70–73.

Wickens, C. D., & Hollands, J. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Woods, D. D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics, 38*, 2371–2393.

Wu, P., & Miller, C. (2010). The current bottleneck for computer-based culture training: Who cares about etiquette? In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2289–2293). Santa Monica, CA: Human Factors and Ergonomics Society.

Zeldes, N., Sward, D., & Louchheim, S. (2007). Infomania: Why we can't afford to ignore it any longer. *First Monday, 12*(8).

Michael C. Dorneich is a principal research scientist in the Human Centered Systems Group at Honeywell Laboratories. His research focuses on creating joint human-machine systems that enable people to be effective in complex and often stressful environments found in aviation, military, and space applications. He specializes in the design of adaptive systems, including human-machine interaction and interfaces, distributed systems, collaborative interactive learning environments, and decision-support systems.

Santosh Mathan is a principal research scientist in the Human Centered Systems group at Honeywell Laboratories. He is principal investigator on a broad range of multi-disciplinary research efforts that focus on applying neurotechnology to enhance human performance, including neural interfaces for robotic control, sensor-based workload measures for pilots and soldiers, and neural measures for diagnosing and rehabilitating individuals with injury related changes in cognitive function.

Patricia May Ververs is a Technology Fellow in the Human Centered Systems Group at Honeywell Laboratories. She has led programs in the areas of high speed research, flight critical systems research, aviation safety program's system wide accident prevention, advanced primary flight displays, improving warfighter information intake under stress/augmented cognition, neurotechnology for information analysts, and enhanced vision systems.

Stephen Whitlow is a principal research scientist in the Human Centered Systems group at Honeywell Laboratories. He has demonstrated technical expertise in cognitive science, cognitive neuroscience, experimental design, adaptive systems, human-automation interaction, user interface design, product requirements definition, and data analysis. He has worked in several domains including flight operations, military operations, unmanned vehicle control, aviation operations support, medical systems, cyber security, and virtual presence through remote sensing.

Caroline C. Hayes is a professor in the Mechanical Engineering Department and the director of graduate studies for the Human Factors graduate program at the University of Minnesota, in Minneapolis, MN, USA. Her research focuses on understanding how to design effective systems of people and technology using an approach that combines elements of engineering, computer science, psychology and management science.