Mathan, S., Dorneich, M., & Whitlow, S. (2005). "Automation Etiquette in the Augmented Cognition Context", Foundations of Augmented Cognition (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 560-569.

Automation Etiquette in the Augmented Cognition Context

Santosh Mathan, Michael Dorneich, Stephen Whitlow

Honeywell Laboratories
3660 Technology Dr, Minneapolis, MN 55410
santosh.mathan@honeywell.com

Abstract

Neurophysiologically triggered adaptive assistance offers many advantages over traditional approaches to automation. These systems offer the promise of leveraging the strengths of humans and machines -- augmenting human performance with automation specifically when human abilities fall short of the demands imposed by task environments. However, by delegating critical aspects of complex tasks to autonomous automation components, these systems run the risk of introducing many of the problems observed in traditional human-automation interaction contexts. This paper describes steps taken to mitigate the possibility of introducing these problems in the context of an augmented cognition system to help dismounted soldiers perform taxing cognitive tasks. It describes specific design decisions that address many of the concerns raised by automation researchers, and presents a summary of performance results that support the overall efficacy of implemented measures.

1 Introduction

The pros and cons of automating complex systems have been widely discussed in the literature (e.g. Parasuraman and Miller, 2004; Sarter, Woods and Billings, 1997). Automated systems bring precision and consistency to tasks, relieve operator monotony and fatigue, and contribute to economic efficiency. However, as widely noted, poorly designed automation can impose several undesirable consequences. Automation can relegate the operator to the status of a passive observer – serving to limit situational awareness, and induce cognitive overload when a user may be forced to inherit control from an automated system.

Automation technologies that have emerged under the Augmented Cognition program address several problems associated highly automated systems. They offer the potential to engage the user in a mixed initiative interaction -- leveraging the strengths of both machines and their human operators. Based on real-time assessments of cognitive state, these systems dynamically provide assistance to users when they are likely to be overwhelmed by task demands. However, there are several features of neurophysiologically triggered automation that can have a detrimental impact on performance. First, many neurophysiological indices fluctuate rapidly over short time windows. Triggering automation on the basis of an index with a high degree of inherent non-stationarity can severely disrupt task performance. Second, adaptive assistance can alter the task demand that the controller is subject to. As a consequence, when adaptive assistance is engaged, neurophysiological measures alone may not effectively reflect the overall task demand imposed by the environment. Unless the task context is also assessed and considered, a neurophysiologically triggered adaptive system could potentially return control to the user under circumstances that may be beyond the capability of a user to handle. Third, despite the fact that systems developed under the Augmented Cognition program display high degree of sensitivity to a user's cognitive state, as automated systems they stand to inherit many of the problems commonly observed with highly automated human-in-the-loop systems.

This paper describes the considerations that went into the design of an Augmented Cognition system being developed by Honeywell. The system is designed to help soldiers perform effectively under extremely demanding task conditions. We describe specific design decisions that address many of the concerns raised by automation researchers, and present a summary of performance results that support the overall efficacy of implemented measures.

2 Task Context and Mitigation Strategies

Before we describe the design considerations associated with automation in the Honeywell augmented cognition context, we present a brief overview of the task domain.

2.1 Task Context

The Honeywell Augmented Cognition program focuses on the dismounted Future Force Warrior (FFW). A critical element of the FFW program is a reliance on networked communications and high density information exchange. These capabilities are expected to increase situational awareness at every level of the operational hierarchy. It is hoped that an information technology based transformation of the military will facilitate better individual and collaborative decision making at every level. However, effective use of these information sources is constrained by the limitations of the human attentional system. The goal of the Honeywell Augmented Cognition team is to use physiological and neurophysiological sensors to detect states where attentional resources may be inadequate to cope with mission relevant demands. Efforts of the Honeywell team have focused on ways to leverage automation to manage information when human attentional resources may be inadequate for the tasks at hand.

The Honeywell AugCog prototype currently takes the form of a wireless, body-worn system. Information from a variety of sensors, including EEG, EKG, fNIR, are used to provide real time assessments of cognitive state. The Honeywell system is being evaluated in field contexts with mobile users. However, many of the mitigation strategies used to assist users were developed and tested in the context of a virtual environment simulated using the Quake III Team Arena game engine. Much of the discussion about the design of automation will be situated in the context of this virtual environment.

The tasks subjects were asked to perform in the virtual environment imposed extreme demands on attentional resources. Tasks ranged from a long and tedious vigilance task prior to a mission, to the execution of a raid on an enemy compound. During the vigilance task, subjects were asked to maintain sustained attention to detect infrequent targets within images transmitted from a surveillance camera. During raids, subjects had to split their attention among a variety of challenging tasks. These tasks included handling critical communications, navigating, and engaging hostile enemy forces.

2.2 Mitigation Strategies

A variety of mitigation strategies were developed to help users perform each of the tasks mentioned above effectively under extreme conditions. Each of these mitigation strategies leveraged automation to handle aspects of a user's task load when attentional resources were likely to be inadequate. We describe these mitigation strategies below.

2.2.1 Mixed Initiative Search

During the vigilance task subjects were asked to detect infrequent targets in scenes transmitted from a surveillance camera. Maintaining attention over long periods of time in a target deficient environment is a difficult task for subjects. Our pilot studies showed that, on average, alert users were able to detect targets with an accuracy of about 80%. However, following periods ranging from 20 to 40 minutes of sparse targets, performance fell to approximately 40% accuracy.

The system described here was designed to help users during periods of low sustained attention. The system helps users by providing advice on the likely location of targets in a scene. Potential enemies in the scene are tagged with a yellow box (Figure 1). Computer systems trained to detect target stimuli may not perform as well as an alert human. Consequently, they may not be able to completely replace the human operator in operational contexts. However, such a system could aid a human who may not be adequately alert. These systems could play a helpful role if they can be triggered when cognitive state classifiers detect a vigilance decrement. This system was modelled after enabling technology currently under development (Schneiderman & Kanade, 2000). The equipment necessary for such a system would include a display integrated helmet with multi-spectral vision capabilities. Mixed initiative search is a part of the FFW vision. For more information see (US Army, 2003).

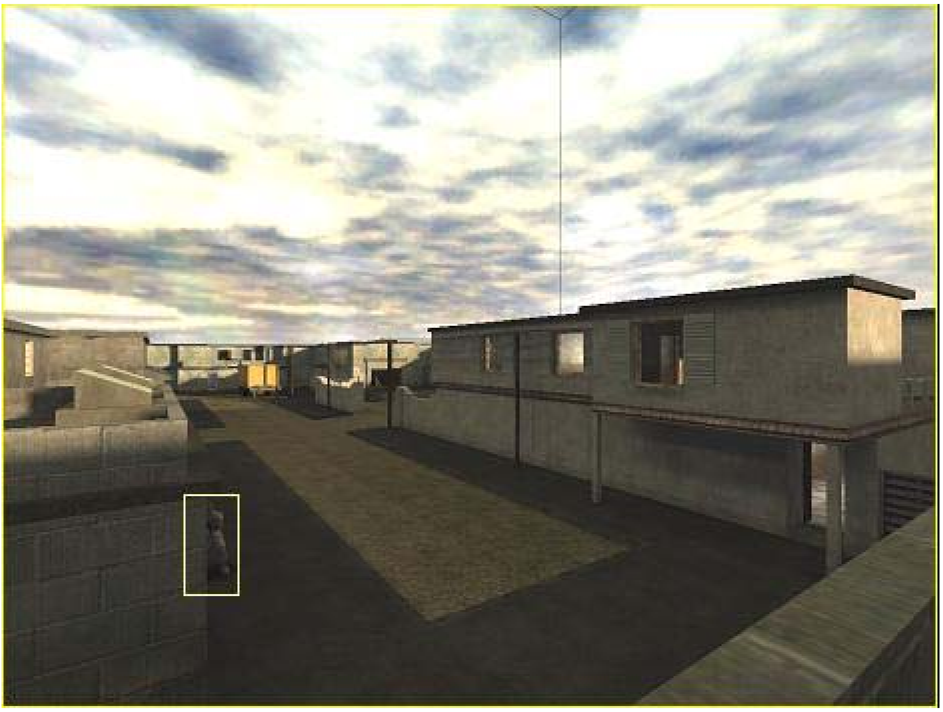


Figure 1: *Mixed initiative target identification system tags likely targets with a yellow box*

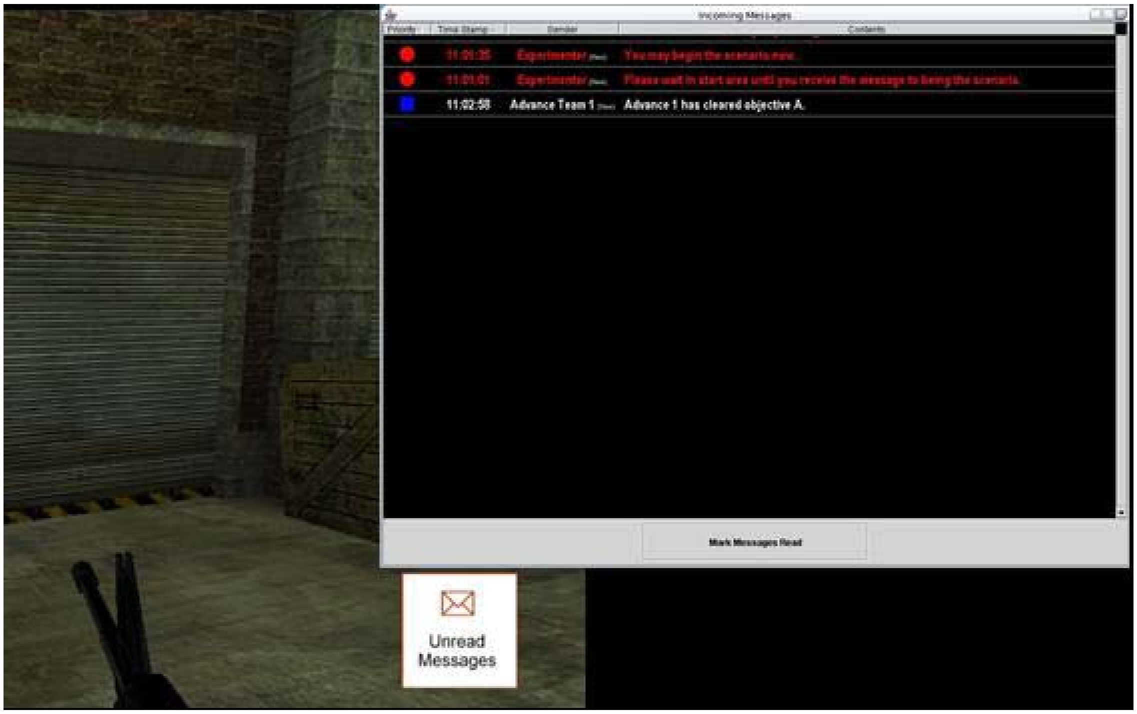
2.2.2 *Communications Scheduling*

During scenarios that involved raids on an enemy compound, subjects had to balance demands associated with navigating a difficult course, managing complex communications, and engaging enemy forces. Quite often subjects would encounter situations where concurrent execution of these three tasks would become near impossible. As Ververs et. al. (2005) have noted, battlefield communications are often the most vulnerable task in the chaos of battle. Critical, mission relevant information may go unheeded and overall situational awareness may be compromised in these situations.

An important feature of the Honeywell Augmented Cognition prototype is the Communication Scheduler. The communication scheduler prioritizes information based on mission relevance. The system adapts presentation modality and message sequencing so that only critical mission relevant information is presented to the user. All messages are stored on a Tablet PC that the user carries (see Figure 2). Low and medium priority messages are deferred for later review, high priority messages are presented immediately to the user in audio format. Redundant visual cues summarize the content of high priority messages.

2.2.3 *Task Off-Loading*

Many mission critical communication transactions follow a well defined format. The transactions associated with a medical evacuation (MEDEVAC) are an example of such a task. The soldier coordinating a MEDEVAC on the field has to accurately relay information about injuries and casualties and communicate precisely where the evacuation can take place. However, the ability



Priority	Time Stamp	Sender	Contents
●	11:01:35	Experimenter	You may begin the scenario now.
●	11:01:01	Experimenter	Please wait in start area until you receive the message to being the scenario.
■	11:02:58	Advance Team 1	Advance 1 has cleared objective A.

Figure 2: *Icon on the HUD informs that communications scheduler is actively managing messages (left). Deferred messages appear on a Tablet PC, organized by priority (right).*

of a soldier to conduct such a lengthy transaction over voice channel, while in a hostile engagement, may be seriously compromised. Personnel may omit important information or make errors in the information transmitted. Additionally, attention devoted to the MEDEVAC exchange may detract from the performance of other critical tasks.

The MEDEVAC Negotiation Agent was triggered on the basis of task context and cognitive State Profile. If workload was high, and a MEDEVAC had to be coordinated, the MEDEVAC Negotiation Agent was triggered. An icon on the HUD notified the user about the need to coordinate an evacuation using the agent. The user could review the MEDEVAC information on the Tablet PC and transmit information using the interactive form. Figure 3 illustrates a MEDEVAC Negotiation Application. Information available on the FFW Netted Communications network would be automatically entered into the form. The system would present this information to the user for inspection. By off-loading the demands associated with a cognitively demanding and lengthy voice transaction to the MEDEVAC agent, the user is able to direct cognitive resources to other tasks at hand.

2.2.4 Navigation Cues

While in hostile territory, soldiers have to be able to adapt their navigation plans to evolving tactical threats. However, during engagements in hostile territory, the cognitive resources necessary to generate a safe route, while engaging the enemy and handling communications, may simply not be available. To address these concerns, the Honeywell Augmented Cognition prototype incorporated functionality to assist users with navigation tasks. In hostile areas the system would generate navigation plans based on knowledge of the mission's geographical

Evacuees:

P Lee	[CRIT]	head	US
S Rogers	[CRIT]	chest	US
Civilian	[CRIT]	chest	non-US
P Lee		head	US
S Rogers		chest	US
V Aleven		chest	US
A Powers		leg	US

Evacuation Coordinates:

N 45° 05.057'

W 93° 23.536'

Evacuation Zone Security:

Medium

Delegate further negotiation:

Yes

Figure 3: *MEDEVAC icon on HUD informs the user that the MEDEVAC agent is ready to coordinate an evacuation (left). User reviews parameters and engages agents through the negotiation application (right).*

objective and information about enemy locations gathered from the FFW communications networks. The system would provide users with a graphical plan in conjunction with tactile cues to guide them through relatively safe zones. Navigational assistance was invoked when the cognitive state profile indicated workload was high and the subject needed to navigate through an unfamiliar route.

Tactile cues were provided to the user by means of a tactor belt worn around the waist. Tactors were fired to direct subjects toward the bearing of their next waypoint The rate of firing the tactors increased from 1 to 2 to 8 Hz as the subject approached each waypoint. When a waypoint was reached the system provided navigation cues relative to the next waypoint until the subject reached the appropriate location. Tactile cues have been shown to be effective in improving performance of spatial tasks even in the presence of competing secondary workload tasks (Raj, Kass, & Perry, 2000).

3 Automation Etiquette

While the mitigation strategies just described promise to help users perform critical tasks under extreme task contexts, like any complex automated system, they have the potential to hurt task performance in a variety of ways. The system described here was designed on the basis of a close consideration of several problems with human-automation interaction highlighted by Sarter, Woods, and Billings (1997). We describe each of these potential problems and summarize the design features they motivated.

3.1 Uneven Distribution of Workload

As Sarter and colleagues point out, many automated systems actually hinder performance in high workload conditions. Many systems require the user to play the role of a translator or mediator – communicating aspects of the task environment to the system. Operators have to take on the responsibility of explicitly specifying task parameters for the automation to execute. In many cases these demands come during the busiest phases of work.

Automation in the context described here was designed to be invoked and parameterized with minimal involvement from the user. Most of the mitigation strategies described here were triggered based on assessments of cognitive state, and task context. As a result, users received assistance automatically in difficult task contexts -- users were not distracted from the task at hand to configure automation. Parameterization of the automated system was supported by the assumed netted communications infrastructure that is a central component of the FFW program. For example, in the navigation task, likely ambush locations were assumed to be continually assessed using information from human and electronic surveillance assets. Real time access to this information by the mitigation engine would allow the system to come up with route plans without explicit intervention from the user. Similarly, the medical evacuation task could rely on the availability of continually updated health status information over communication networks. As a result, many of the parameters associated with the medical evacuation task could be completed with minimal user intervention.

3.2 Breakdowns in Mode Awareness

Sarter and Woods define mode awareness as the ability of a system user to anticipate the behavior of automated systems. They suggest that breakdowns in mode awareness, so-called automation surprises, can lead to errors of omission in which the operator fails to observe and respond to uncommented or undesirable system behavior.

There are several sources of potential automation surprises in the context described here. First neurophysiological and physiological indices that serve to invoke automation exhibit a great deal of inherent non-stationarity. Triggering mitigations on the basis of signals that vary a great deal over short windows of time can be extremely disruptive for the user. To address this problem, cognitive state classification was based on joint consideration of several indices. Some of the indices employed in our system included EEG, galvanic skin response, heart beat variability and pupilometry. These redundant sources of information combine to provide a more stable indication of cognitive load than any single index would. Our current efforts add additional robustness to cognitive state classification by picking the modal classification output over specified time windows. The tradeoff between mitigation latency and required classification robustness determines the size of the window employed.

Second, once effective mitigation strategies are triggered, they effectively reduce the cognitive load on the user. Consequently, neurophysiological and physiological indices loose their value as indicators of task load. Disengaging mitigations solely on the basis of indices associated with cognitive load can return control to users under very difficult task conditions. To address this issue, mitigations were turned off on the basis of context related information. That is, task related information that is independent of cognitive state assessment. For example, communications scheduling was turned off after the user had indicated that they had caught up on all the deferred messages. Navigation cues were terminated only after a user had arrived at the destination

Third, the system described here provides a range of different types of assistance to users in different tasks contexts. Each of these mitigations assumes control over a certain aspects of a user’s task. Unless a user is clearly aware of the status of the adaptive system, the user could encounter a range of automation surprises. To avoid these problems, the system was always explicit concerning the automation mode. For instance, when the communications scheduler was engaged, the user would see an icon on the heads up display (HUD) indicating that messages were being deferred. When system was ready to transact a Medevac, a notification icon would pop up on the HUD; the user would have to explicitly authorize a Medevac based on information populated by the system. Authorization was communicated by clicking a button. Once navigation aiding was turned on, user would feel pulses on a belt that unambiguously conveyed the navigation mode to the user.

3.3 New Coordination Demands

Sarter and colleagues suggest that autonomous automation components effectively become like crew members by taking over aspects of critical tasks. However, unlike good crew members, poorly designed automation may fail to keep users informed about task status. These systems may perform tasks autonomously and silently, but return control to users abruptly when things fail. This serves to raise the coordination requirements and could add to cognitive load.

Elements of the system described here were designed with the assumption that they were fallible. Several mitigations were designed to allow a human to intervene if the system was unable to handle a situation effectively. For example, we recognized that the process of organizing a MEDEVAC may often require more knowledge of the mission context than the MEDEVAC Negotiation Agent might be able to infer from netted data sources. In case further clarifications were required, the MEDEVAC dialog had an option for the user to delegate further negotiation to a platoon member with fewer task demands. The system provided a way to minimize coordination demands by off-loading clarifications to another human in a graceful manner.

3.4 Complacency and Trust in Automation

Sarter and colleagues suggest that complacency induced by automation may be a critical factor in many accidents. They suggest that users may come to relay on automation. Not realizing that these systems, though largely reliable, may be fallible.

Issues of complacency were also a matter of concern in the context described here. By delegating critical tasks such as communications and visual monitoring to an automated system, users stood the risk of missing critical task relevant information. Our approach to reducing possible complacency relied on training to emphasize the fallibility of the system and to provide users with procedures for monitoring the system and resuming control of delegated tasks as soon as practical. For instance, in training associated with the vigilance task, subjects were told that the system was highly fallible, with a performance accuracy of approximately 68%. They were asked to consider the system's assessment, but to remain vigilant. With tasks involving the communications scheduler, the system would only allow high priority messages to be played to users during periods of high workload. However, training emphasized the need to catch up with medium and low priority messages at the earliest possible moment. Results from experiments suggest that these issues of complacency were successfully addressed. Subjects detected targets with an accuracy of over 80%; a level substantially over the 68% base accuracy of the system. Subjects also had a much higher level of situational awareness of the status of the mission relative to conditions where they did not get automatic communications-related assistance from the system.

3.5 Training

Automation of complex tasks often introduces the need for additional training. Besides learning to master the performance of inherently complex tasks, users have to learn about the use of complex automation components to support the execution of these tasks. Sater and colleagues, argue that sophisticated automation components interact with the task environment in complex ways. They argue that training has to occur in the context of use for users to be able to acquire accurate mental models of the system.

All subjects who used the system discussed here received extensive training in the use of automation components in the actual contexts of use. Subjects progressed on to task scenarios only after they were able to successfully demonstrate use of each automation component in training scenarios. Subjects also had to answer a broad range of questions about each automation component and it's interaction with the task environment.

4 Results

Experimental results support the overall efficacy of the automation components incorporated in the system. The Honeywell Augmented cognition system helped subjects perform significantly better on communication, fighting, navigation and vigilance tasks with neurophysiologically triggered automation. Additionally, most subjects felt that the automation components described here actually made task execution easier.

The experiment was conducted with 13 experienced video game players who performed military relevant tasks in the Quake virtual environment. Subjects performed tasks within experimental scenarios that required users to balance communication, fighting and navigation tasks. Subjects performed each scenario with automation and without. Such an experimental design allowed for a within subjects assessment of the effects of neurophysiologically triggered automation on performance. Subjects performing tasks with assistance from the communication scheduler responded to double the number of messages as they did in scenarios without assistance from the system. They performed almost twice as well on assessments of situational awareness in scenarios with automation. Additionally, navigational assistance allowed users to run into a third of the number of ambushes as they did in scenarios without assistance. Subjects also performed significantly better with the help of automation on the vigilance task. Subjects who received neurophysiologically triggered assistance detected 30% more targets than subjects who did not. A between subjects comparison was employed to assess system effectiveness on the vigilance task. The length of time required to induce a vigilance decrement precluded the use of a within subject design in this context.

Subjective assessments of the automation suggest that users found the automation components in the system to be helpful in performing tasks. For instance, 76.9% of the subjects thought it was easier to perform tasks in the mitigated condition. Subjects also found mitigated condition easier for fighting (61.5%), communicating (84.6%) and navigating (76.9%).

5 Conclusion

This paper reviewed several problems with human automation interaction that were relevant to the design of a neurophysiologically triggered adaptive system. Specific strategies to minimize the negative impact of automation were discussed. Experimental evaluations suggest that automation components designed with the considerations specified here were indeed successful in helping users. However, there are several limitations of this work that are important to mention. The automation strategies described here were assessed in a military relevant desktop training environment. It is unclear whether these techniques will remain effective in actual military contexts. Second, it is hard to tell which of the specific automation design strategies implemented here had the most impact on users. The design of the experiment only attests to the relative efficacy of these strategies as a package. Third, the system evaluated here only incorporated support for a small subset of tasks that soldiers are likely to perform. As automation components get added and layered, the overall complexity of the system could grow far beyond the capacity of the strategies described here alone to address.

Acknowledgments

This research was supported by a contract with DARPA and funded through the U.S. Army Natick Soldier Center, under Contract No. DAAD-16-03-C-0054, for which CDR (Sel.) Dylan Schmorrow serves as the Program Manager of the DARPA Augmented Cognition program and Mr. Henry Girolamo is the DARPA Agent. Any opinions, findings, conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Army.

References

Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., and Ververs (2005). "Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State." In Proceedings of the 11th International Conference on Human-Computer Interaction (Augmented Cognition International), Las Vegas, Jul 22-27, 2005

Norman DA. The problem with "automation": inappropriate feedback and interaction, not "over-automation". Philosophical Transaction of the Royal Society of London, 1990; B327: 585-593.

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

Raj, A. K., Kass, S. J., & Perry, J. F. (2000). Vibrotactile displays for improving spatial awareness.Proceedings of the International Ergonomics Association XIV Triennial Congress / 44th Annual Meeting of the Human Factors and Ergonomics Society (1), 181-184.

Sarter, N.B., Woods, D.D., and Billings, C.E. (1997). Automation Surprises. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd edition) (pp. 1926-1943). New York, NY: Wiley.

Schneiderman, H. and Kanade, T. (2000) "A statistical approach to 3d object detection applied to faces and cars". Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pages 746--751, Hilton Head Island, South Carolina, June 2000

US Army (2003), "The Objective Force Warrior: The Art of the Possible…a Vision," Available Online at http://www.ornl.gov/sci/nsd/pdf/OFW_composite_vision.pdf.

Ververs, Whitlow, Dorneich, Mathan, and Sampson (2005) AugCogifying the Army's Future Warfighter. In Proceedings of the 11th International Conference on Human-Computer Interaction (Augmented Cognition International), Las Vegas, Jul 22-27, 2005