Dorneich, M.C., Ververs, P.M., Mathan, S., and Whitlow, S.D, (2006). "Evaluation of a Tactile Navigation Cueing System Triggered by a Real-Time Assessment of Cognitive State," *Proceedings of the Human Factors and Ergonomics Society Conference 2006*, San Francisco, CA.

EVALUATION OF A TACTILE NAVIGATION CUEING SYSTEM AND REAL-TIME ASSESSMENT OF COGNITIVE STATE

Michael C. Dorneich, Patricia May Ververs, Stephen D. Whitlow, Santosh Mathan
Honeywell Laboratories
Minneapolis, MN

This paper details an evaluation of a tactile cueing system that was created to enhance the navigation of a complex route. The complexity of the task along with simultaneously challenging cognitive tasks also enabled the real-time assessment of cognitive state in various task load conditions. Honeywell has been working with the US Army's Future Force Warrior program to develop adaptive systems that will effectively manage the available cognitive resources used for information processing by the dismounted Soldier in highly dynamic, information rich environments. The appropriate allocation of cognitive resources is key to managing multiple tasks, focusing on the most important ones, and maintaining overall situation awareness. Non-visual navigation support would offload a typically visual task, such as viewing a paper map or computer-based map display, to a sensory channel that is underutilized, tactile sensation. Both benefits and costs to this type of automation support are explored in detail, where the evaluation supports the premise that strong automated support should only be used in high workload situations where the benefits outweigh the costs.

INTRODUCTION

This paper describes an evaluation of the benefits and costs of a Tactile Navigation Cueing System (TCNS). The TCNS is an adaptive system that mitigates non-optimal performance via adaptive automation. Cognitive state is assessed in real-time via a suite of neuro-physiological and physiological sensors to directly measure the person's current level of workload. When the assessment of workload is high, and additional tasks or information processing demands can not be met, this would be a candidate time to invoke automation. However, that automation may not be appropriate when the cognitive capacity is well matched to the current task demands. Thus the adaptive system should only be invoked when the person's ability to handle the task demands breaks down. In a joint human-automation system, where automation support is provided on an as-needed basis (as opposed to continuously), careful consideration must be given to the costs and benefits to determine the optimal time to invoke and disengage automation assistance.

Adaptive automation assistance triggered by real time classification of cognitive state offers many advantages over traditional approaches to automation. These systems offer the promise of leveraging the strengths of humans and automation - 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, such systems run the risk of introducing many of the same automation-induced problems observed in traditional human-automation interaction contexts. 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). However, it has been widely notated that poorly designed automation can have serious negative effects. For example, clumsy automation can relegate the operator to the status of a passive observer ñ serving to limit situational awareness (SA), and induce cognitive overload when a user may be forced to inherit control from an automated system. The current design explicitly considered the costs as well as the benefits of automation when deciding when and how to intervene in the participants workflow.

The dismounted Soldier of the future will be faced with greater decision making responsibilities that are enabled by the voluminous real-time information that will be available on the mobile battlefield. The Army will rely on small combat units with netted communications enhanced with information from distributed and fused sensors, tactical intelligent assets enabling increased SA and on-the-move planning (FFW, 2004). The FFW program seeks to push decision making requirements to the lowest levels and posits that with enhanced netted communications capabilities a squad can cover the battlefield in the same way that a platoon now does.

Honeywell is working with the U.S. Army's Future Force Warrior (FFW) program to build an adaptive system to support the appropriate allocation of cognitive resources that isis key to managing multiple tasks while focusing on the highest priority activity and maintaining SA. If one considers the navigation in unfamiliar and 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 TCNS prototype incorporated functionality to assist participants with the navigation task of avoiding threat zones and arriving at their target destination. In hostile areas the system would generate navigation plans based on knowledge of the mission's geographical objective and information about enemy locations gathered from the FFW communications networks. The system would provide users with tactile cues to guide them along safe and secure paths to their objective. Navigational assistance would be invoked when the system assessed that the participantís cognitive state profile indicated workload was high and the participant needed to navigate through an unfamiliar route. The authors recognize that this is a very specific type of automation and would only be considered in cases where context information such as geo-coordinates about the current location and final destination are available. Additional information such as terrain and building structures would need to be known before route navigation cueing would be deemed beneficial to the user.

Navigation aiding via tactile cues is a relevant research topic because the visual nature of the battlefield as well as the new display devices will quickly overwhelm the visual input channel; the visual domain is an overused means of presenting information to the individual who also must use vision to navigate the battle space and to attend immediate threats in the environment. Our working premise is that non-visual navigation support would offload a typically visual task, such as viewing a paper map or computer-based map display, to a sensory channel that is underutilized, tactile sensation. This application could be considered a feasibility prototype of eventual Future Force Warrior capability that would reduce attentional bottlenecks by offloading typically visual tasks to the underutilized tactile sensory channel.

System Description

The TCNS is the focus of the evaluation described in this paper. Also discussed is the cognitive state assessment performed in the high and low task load conditions.

Tactile navigational cueing system. The Tactile Navigational Cueing belt used eight vibrating tactors in conjunction with position information from a GPS system and bearing information from the Dead Reckoning Module (DRM) to guide the participant to known locations. Tactile cues were provided to the user by means of a tactor belt worn around the waist. Tactors fired to direct participants toward the bearing of their next waypoint The rate of firing the tactors increased from 1 to 2 to 8 Hz as the participant approached each waypoint. When a waypoint was reached the system provided navigation cues relative to the next waypoint until the participant reached the appropriate destination.

Under heavy information processing demands imposed by operational tasks such as responding to radio communications and maintaining awareness, participants would likely be unable to quickly and safely navigate a complex route; therefore if the cognitive overhead of navigating to an objective was temporarily alleviated, available resources could be used to complete the tasks at hand. The cognitive state was assessed during periods of high and low task load conditions. Ideally this cognitive state classification would be used to activate the tactile cueing device until the destination was reached. Thus the navigation task would go from being cognitively intense (involving reading a map, mental transformation from 2D to 3D space, etc) to one that is essentially a reactionary task to external stimuli that requires less attentional resources. The overall effect is to lower the task load and cognitive demands, allowing users to improve performance on the navigation task while not adversely affecting other tasks being done simultaneously. 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).

Cognitive state assessor. Cognitive state classification was based on spectral EEG data collected from Advanced Brain Monitoring's (ABM) EEG sensor headset. The sensor headset acquired six channels of EEG using a bipolar montage. Differential EEG are sampled from bipolar channels CzPOz, FzPOz, F3Cz, F3F4, FzC3, C3C4 at 256 samples per second with a bandpass from 0.5 Hz and 65 Hz (at 3 dB attenuation) obtained digitally with Sigma-Delta A/D converters. Quantification of the EEG in real-time is achieved using signal analysis techniques to identify and decontaminate eye blinks, and identify and reject data points contaminated with electromyography (EMG), amplifier saturation, and/or excursions due to movement artifacts.

Cognitive state classification was driven by kernel-based classifiers. Classifiers were trained for each individual using power spectrum estimates collected during close isomorphs of experimental tasks. Experimental tasks placed participants in conditions of high and low workload by manipulating task load. Over the course of the training run, the system developed a model of power spectrum profiles associated with various cognitive states of interest. During experimental runs, the classifier examined each power spectrum estimate and associated it with the most likely cognitive state. The output of the model was an assessment of a participant's cognitive load.

Automation Etiquette

Automation technologies triggered by real-time assessments of cognitive 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 neuro-physiologically triggered automation that can have a detrimental impact on performance.

Many neuro-physiological 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. In addition, adaptive assistance can alter the task demand that the controller is participant to. As a consequence, neuro-physiological measures may not effectively reflect the overall task demand imposed by the task environment. Unless the task context is assessed and considered using non-physiological sensors (e.g., accelerometers to indicate movement or headtrackers to indicate direct of focus or active scanning of the environment), a neuro-physiologically triggered adaptive system could potentially return control to the user under circumstances that may be beyond the capability of a user to handle.

Operationally, pulses from the Tactile Navigation Cueing System "tugged" the participants in the direction the participant was expected to go. The system was designed to be invoked when the cognitive state classification indicated cognitive workload was high and the participant needed to navigate through an unfamiliar route. However, turning the system off as soon as cognitive workload fell below some threshold would leave users disoriented in an unfamiliar area. Thus once the system is turned on, the navigation mitigation persists until users arrive at the safe destination**.** The system is designed to be invoked when *workload* is high, rather than any time a user needs to navigate through an unfamiliar route, since there is a potential loss of SA when the person is not forced to navigate on his or her own. This cost is realized if a Soldier were to find him or herself in the area at a later time, or their commanders assume they know their way around because they have been there before. The lack of knowledge of the area could have detrimental effects. While this loss of SA may be acceptable in high task load situations, it is an unnecessary cost when the Soldier is capable of navigating the area on his or her own.

Cost/Benefit Approach

As with any adaptive automation system, the TCNS stands to potentially inherit many of the problems commonly observed with highly automated human-in-the-loop systems. The system was designed with close consideration of several identified problems with human-automation interaction, as highlighted by Sarter, Woods, and Billings (1997).

Uneven distribution of workload. Many systems require the user to play the role of 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. The TCNS was designed to be triggered based on assessments of cognitive state and task context. As a result, users could receive automated assistance automatically in difficult task contexts that induce high workload -- users would not be distracted from the task at hand to configure the automation's intervention. Parameterization of the automated system could be supported by the assumed netted communications infrastructure that is a central component of the FFW program. For example, likely ambush locations could be continually assessed using information from human and electronic surveillance assets. Real time access to this information would allow the system to come up with route plans without explicit intervention from the user.

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.

Once effective mitigation strategies are triggered they effectively reduce the cognitive load on the user, thus an index of cognitive state looses its value as an indicator 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 can be turned off on the basis of context-related information. Navigation cues were terminated only after a user arrived at the destination, since stopping cueing once it has started is likely to result in even greater disorientation, since there is an inherent loss SA when being cued versus navigating with a map. The loss of SA was deemed an acceptable cost when cognitive overload threatened a complete breakdown in performance; however this cost must be accounted for when deciding how and when to "turn off" mitigation.

*New coordination demands***.** Poorly designed automation may fail to keep users informed about task status. 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.

Mitigation design should allow human intervention if the system is unable to handle a situation effectively. In the simplest example, users should have the ability to turn off the mitigation if it is not performing to expectations.

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. Sarter and her colleagues, argue that sophisticated automation 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 participants who used the system received extensive training in the use of cueing system under automation in the actual contexts of use. Participants progressed on to task scenarios only after they were able to successfully demonstrate use of automation in training scenarios.

METHOD

Participants. Eight male participants volunteered to participated in this evaluation. They ranged in age from 21 to 42 years of age with an average age of 29.5 years. None of the participants had military experience.

Experimental design. This was a two (cueing automation: on/off) by two (task load: low/high) within subjects design. The order of the task load conditions was held constant with high task load received first. In order to assess the true benefits and costs of the cueing system in both low and high task load conditions, the mitigation was forced on and off to provide a completely crossed design. The scenario was performed twice, once under automation and another time without automation. This order was counterbalanced. Each scenario contained periods of high and low task loads. The two versions of the scenario were similar but not identical to avoid any learning effects. The entire data collection averaged approximately 17 minutes per subject, with a range of 11to 25 minutes.

Operational scenario. The participant played the role of a platoon leader (PL), managing three squads while reporting to his or her company commander (CO). Participants navigated a complex route in an unfamiliar area to avoid video surveillance detection and virtual mine fields while executing a mission to navigate to an objective to set-up a fortified surveillance watch. Secondary tasks included scanning the environment for potential improvised explosive devices (IEDs), monitoring radio communications to organize another evolving mission, maintaining radio counts, and performing a periodic mathematical task.

Navigation task. The participants navigated along an unfamiliar and unmarked route. They were given a map with landmarks, and potential mine locations. They were to navigate to an objective without triggering a mine and without being detected by security cameras.

Maintain radio counts. A simulated CO relayed messages about entities encountered by his or her three platoon leaders over the radio, of which the participant was one. The messages contained reports of civilians, enemies, or friendlies spotted. The participant maintained a running total of civilians, enemies, and friendlies reported to him, while ignoring the counts reported to the other two platoon leaders. This task relied heavily on the participants' ability to keep the three counts in working memory until he was asked to report the counts. This task also required the participants to focus their aural attention to listen for their call-sign and ignore the messages directed to the other platoon leaders.

Mission monitoring. The participant organized the execution of a series of bounded overwatch maneuvers by three squads under his command, where one squad moves while the other two squads protect the moving squad. Participants kept track of the status of all three squads ó either "ready to move" or "ready for overwatch." Once all three squads reported that they were in position (two squads ready for overwatch and one squad ready to move), participants ordered the appropriate squad to move forward. This task required the participants to keep the track of the three squads, their locations, and readiness to advance in the mission in working memory until the final team was in position.

Visual search of IEDs. The field in which the participant was navigating contained multiple IEDs. The IEDs were round flat discs of various colors. Participants were instructed to radio in to report the sighting and approximate location. This task forced participants to visually scan their environment.

Interruption task. For this experiment, a simple math problem was periodically presented to the participants as an interruption task during the scenarios. This task was representative of any type of unanticipated interrupt that requires significant cognitive resources and an immediate response from the platoon leader. Participants were interrupted twice per minute in both high and low task load periods. This interruption task had the potential of disrupting any of the tasks that required continual rehearsal, such as the working memory tasks of mission monitoring and maintaining counts. Also, the head-down time with the PDA had the potential of disrupting the visual search for the IEDs in the environment.

*sa*SA s*Research questions*. Would the tactile navigation cueing be intuitive to learn and successfully guide the participants to avoid danger zones and reach the destination? Would there be a cost to the use of the tactile cueing such as a loss to situation awareness regarding information normally acquired en route? Would the sensor-driven classification of cognitive state detect a change in cognitive state between low task load and high task load conditions? Would cognitive state changes correlate with changes in performance?

Hypothesis. The use of tactile cueing would enhance performance on the navigation tasks in high task load conditions, while potentially lowering performance when mis-applied during low task load periods. In order to assess the true benefits and costs of the cueing system in both low and high task load conditions, the mitigation was forced on and off to provide a completely crossed design.

RESULTS

Navigation task. The mitigation directly targeted performance on the navigation task. Mitigation showed no statistically significant ($F_{1,5}$=.27, p=.63) effect on composite run time in the low task load condition: 98.3 seconds (unmitigated) vs. 107.7 seconds (mitigated). Under high task load, however, the Tactile Navigation Cueing system enabled a statistically significant ($F_{1,5}$=8.69, $p < 0.05$) reduction in composite run time, from 562.1 sec (unmitigated) to 389.2 sec (mitigated). Much of this composite run time score was driven by the number of incursions. A 30-second plenty was assessed if the participant intruded a mine zone area. All six participants saw their survivability improve as measured by the number of mines they encountered. Under unmitigated high task load conditions, all six participants triggered mines at least four, and up to, seven times. With mitigation, however, five of six participants avoided the mines altogether, and the one remaining participant encountered fewer unmitigated than when mitigated. Clearly the high task load participants were able to navigate to the objective more quickly and more safely.

Counts task. Participants in high task load condition performed at equivalent levels having a 29.9% accuracy when unmitigated and 35.1% accuracy when mitigated. There was no difference in these accuracies ($F_{1,5}$=0.42, p=.55). In the low task load blocks participants saw a statistically significant ($F_{1,5}$=7.71, $p< 0.05$) decrement in performance in the mitigated condition, 26.2% accuracy, while unmitigated accuracy was 43.3%. This is consistent with the hypothesis that the application of a mitigation can result in a cost to performance if not appropriately applied to the situation. For example, it is possible that the mitigation (tactile buzzing) proved to be a distraction to competing tasks when walking in a straight line in the low task load block.

Mission monitoring task. Participants performed equivalently on the mission monitoring task in high task load when mitigation was available. Performance in high task load was similar in the unmitigated (35.2%) and mitigated (49.3%) conditions ($F_{1,5}$=.92, p=.38). Likewise, in low task load blocks, the accuracy was similar with 83.3% in the unmitigated condition and 100% in mitigated ($F_{1,5}$=1.00, p=.36). The results trended in the direction of the mitigation positively influencing performance.

Interruption task. The math interruption task was used to assess attention and cognitive resources available at any given moment. There are three measures associated with the task; they were response time, solution time, and accuracy.

Due to data logging issues, only two of six participants' data were recorded for the low task load condition of the math task. All six participant's data were used in the analysis of the high task load condition for the math task.

Participants responded to the interruption alert much more quickly in the low task load condition, as expected. In the low task load condition, the mitigation actually slightly increased response time from 3.3 sec (unmitigated) to 5.2 sec (mitigated), although not significantly ($F_{1,1}$=1.00, p=.36). In the high task load condition, where we expected to see the benefits of mitigation, participants responded faster under mitigation, improving response time from 22.0 sec (unmitigated) to 9.1 sec (mitigated). Although the difference was not statistically significant ($F_{1,5}$=1.20, p=.32), the reduction was in the expected direction.

Once interrupted, participants' time to actually solve the math problem presented to them did not vary significantly across any of the experimental conditions. The data suggest that once the participants were interrupted, their entire attention was focused on solving the math problem. Participantsí accuracy in solving the math problems was considerably reduced in the high task load condition, as compared to in the low task load condition.

Under low task load, accuracy was similar in both mitigation conditions with 83.3% accuracy in the unmitigated condition and 100% in the mitigated case ($F_{1,5}$=6.84, p=.23). Under high task load, accuracy significantly ($F_{1,5}$=7.26, $p < 0.05$) improved from 47.0% (unmitigated) to 68.0% (mitigated). Again, the benefits of the mitigation were seen in high task load conditions.

Visual search for IED. Participant performance didn't vary significantly either under task load or mitigation conditions. In low task load, participantsí search accuracy was not significantly ($F_{1,5}$=0.29, p=.61) different: 41.7% (unmitigated) and 33.3% (mitigated). In high task load, participantsí search accuracy was unchanged from 50.0% in both the unmitigated and mitigated cases ($F_{1,5}$=0.00, p=1.0). One might reasonably expect that the mitigation would free up resources to scan the environment. However, with no mitigation the participant is forced to navigate and scan the environment looking for visual cues, where they would also detect IEDs. Unfortunately the data are inconclusive.

Cognitive state classification. We used a backward elimination, a heuristic search procedure through the space of possible feature subsets to identify a subset of features that would provide reliable classification. Feature selection was based on the training data collected prior to the individual scenarios. With an appropriate selection of channels we were able to classify cognitive state in the high and low workload states with an accuracy that exceeded 70% for all participants. We observed classification accuracy as high as 95% for one participant. See Mathan et al. (2005) for more information.

DISCUSSION

The goal of this evaluation was to investigate the viability of a navigational cueing device for benefiting performance is a non-laboratory navigational task. We also wanted to investigate whether the cognitive state classifier could reliably detect differences in workload induced by task load variations. The experimental results showed that the cost/benefit tradeoffs of cueing automation are pronounced. The Tactile Navigation Cueing System, by relieving participants of the cognitively challenging task of navigating through an unfamiliar area, resulted in the improvement of almost all tasks in the high task load condition. However, when the Tactile Navigation Cueing System was invoked during low task load periods, it was so distracting that almost all tasks suffered as a result. When the tactile cueing was used to augment the navigation task under high workload conditions, the participants had shorter run times and fewer navigation incursions into the mine zones. The addition of the mitigation did not interfere with performance on the radio count task or mission monitoring task and even improved accuracy on the interruption task. When the mitigation was triggered in the low task load conditions, performance was adversely affected leading to poorer recall in the maintaining counts working memory task. We conclude that the unnecessary cueing was a distraction to this task. The mitigation proved most effective when used during high task load periods. Costs were involved when mitigations were inappropriately invoked during low task load periods resulting in no added benefit and sometimes a performance decrement. The finding of greater than 70% accuracy of cognitive state classification based on EEG data alone was promising in an environment many have found intractable. Closing the loop with an accurate assessment of cognitive state, in order to appropriately trigger mitigation, is one way to improve human-system integration in operational settings.

ACKNOWLEDGMENTS

This paper was supported by a contract with DARPA and the U.S. Army Natick Soldier Center, DAAD16-03-C-0054, for which CDR Dylan Schmorrow served as the Program Manager of the DARPA Improving Warfighter Information Intake Under Stress/AugCog program and Mr. 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.

REFERENCES

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 the *IEEE Proc of the International Conference on Systems, Man and Cybernetics*, New York: IEEE.

Future Force Warrior. (n.d.). Retrieved September 24, 2004, from http://www.natick.army.mil/ffw/content.htm.

Mathan, S., Mazaeva, N., Whitlow, S., Adami, A., Erdogmus, D., Lan, T., & Pavel, M. (2005). Sensor-based cognitive state assessment in a mobile environment. In G. Salvendy (Ed.) *Proceedings of the 11th International Conference on Human-Computer Interaction* (1st Annual Augmented Cognition International conference). Mahwah, NJ: Lawrence Erlbaum.

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 44th Annual Meeting of the Human Factors and Ergonomics Society*, 181-184, Santa Monica, CA: HFES.

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