Dorneich, M., Whitlow, S., Ververs, P.M., & Mathan, S. (2007). "Iterative Experimental Design to Mature Cognitive State Classification Techniques from the Laboratory to Field Settings", *Foundations of Augmented Cognition, 4nd Edition* (D.D. Schmorrow, D.M. Nicholson, J.M. Drexler, & L.M. Reeves, Eds.). Arlington, VA: Strategic Analysis. pp. 23-32.

Iterative Experimental Design to Mature Cognitive State Classification Techniques from the Laboratory to Field Settings

Michael C. Dorneich, Stephen D. Whitlow, Santosh Mathan, Patricia May Ververs

Honeywell Laboratories
3660 Technology Drive, Minneapolis, MN 55105
Michael.dorneich@honeywell.com

Abstract

This paper summarizes the research conducted on two parallel and complementary thrusts: cognitive state assessment (CSA) and mitigations development for augmented Cognition Systems. Cognitive workload classification research has been largely been limited to laboratory contexts with well defined tasks. However, as the technology matures, it is important to evaluate real-time, EEG- and ECG-based cognitive workload classification techniques when they are fully subject to the noise, motion, weather, and wide range of physical and cognitive tasks inherent in dismounted operational environments. The hardware and software infrastructure must reliably collect clean sensor data, mobile processing is needed to log and process the data, and the experimental design must reliably put participants in the cognitive states of interest. As mitigations mature, automation etiquette plays an important role in ameliorating the negative effects of strongly adaptive systems. In addition to the practical and system configuration challenges faced when moving from the laboratory to field studies, there are issues of experimental control and the characterization of cognitive state in less constrained task environments. A key component of an evaluation is to compare the classification results with "ground truth," i.e. the actual workload experienced by the participant. While in the laboratory it is possible to develop simple tasks where workload is manipulated precisely and consistently, in an unconstrained field environment it becomes substantially harder to manipulate workload precisely and to interpret and assess a user's true cognitive state without compromising operational realism. Advances in experimental design, data collection protocols, signal processing, classification, mitigation strategies, and experimental ground truth methodology successfully enabled classification of cognitive workload level in an unconstrained, fully-mobile, free-play operation with Soldiers executing missions in a challenging urban terrain environment.

1 INTRODUCTION

1.1 Adaptive Automation

Adaptive systems must make timely decisions on how best to use varying levels of (adaptive) automation to provide support in a joint human-automation system. Adaptive automation, where the automation adapts during execution to the current task environment, can either provide adaptive aiding, which makes a certain component of a task simpler, or can provide adaptive task allocation, which shifts an entire task from a larger multitask context to automation (Parasuraman, Mouloua, & Hilburn, 1999). In order for an adaptive system to decide when to intervene, it must have some model of the context of operations, be it a functional model of system performance, or possibly a model of the operator's functional state. Many adaptive systems derive their inferences about the cognitive state of the operator from mental models, performance on the task, or from external factors related directly to the task environment (Wickens & Hollands, 2000). For example, work in the 1980s in the field of "Associate Systems" developed human-adaptive, information and automation management technology, which were designed to share a common understanding of the mission, the current state of the world, the platform, and the operator him- or herself. Associate systems then used that shared knowledge to plan and suggest courses of action and to adapt information displays and the behaviour of automation to better serve the inferred operator intent and needs (Miller & Dorneich, 2006). The primary means used in associate systems programs to infer operator intent was logical deduction based on knowledge of the mission plan and the functional capabilities of the platform (Geddes, 1985).

While measurement of performance on the task can respond to unpredictable changes in the cognitive state of the operator, the method is only as good as the ability to measure performance (Scerbo, Freeman, Mikulka, Parasuraman et al, 2001). The use of behavioural responses to track cognitive function requires regular or periodic performance assessments to keep an updated assessment on performance capabilities. Many systems provide few opportunities to track overt responses to monitor operator performance (Parasuraman, 2003). Additionally, diagnosing cognitive state degradation via performance degradation occurs "after the fact" and thus an opportunity to proactively adapt the system to maintain performance is limited. Modelling techniques are only as good as the underlying model, and susceptible to model brittleness. Brittleness occurs whenever solutions are subject to system modelling assumptions that are necessarily an incomplete representation of the world (i.e. the system model does not account for all possible scenarios) and therefore subject to erroneous solutions when an important, but un-modelled, feature of the problem space impacts the choice of optimal solution (Smith, McCoy, & Layton, 1997). The more complex the task environment, the greater likelihood that the model will not have been designed to anticipate all aspects of human operator performance.

The aim of Augmented Cognition systems is to use physiological and neurophysiological sensors to detect the cognitive state of the operator - states where human cognitive resources may be inadequate to cope with relevant task demands. The goal is to enhance human performance when task-related demands surpass the human's current cognitive capacity, which fluctuates subject to fatigue, stress, overload, or boredom. Neurophysiologically and physiologically triggered adaptive automation offers many advantages over the more traditional approaches to automation by basing estimates of operator state on directly sensed data. These systems offer the promise of leveraging the strengths of humans and machines by augmenting human performance with automation specifically when assessed human cognitive capacity falls short of the demands imposed by task environments. With more refined estimates of the operator's cognitive state, measured in real-time, adaptive automation also offers the opportunity to provide aid even before the operator knows he or she is getting into trouble. This approach does not require instrumentation of systems to record behavioural actions required for task model-based systems.

The potential applications of augmented cognition cover a wide range of human-computer joint cognitive systems. One such application would be a closed loop adaptive system to help optimize performance of a stationary operator. For instance, such systems may include operators who interact with information displays, such as a unmanned air vehicle (UAV) ground control station operator (Snow, Barker, O'Neill, Offer, & Edwards, 2006), or an operator of a weapon control system such as the Tactical Tomahawk (Tremoulet, Barton Craven, Gifford, Morizio, et al., 2006). Other uses include using AugCog technologies for studying skill acquisition during telerobotic arm training (Krebs, et al., 1998). Cognitive state assessments during training could also be used to diagnose student difficulties in real-time and provide appropriate context-specific assistance to learners (Mathan & Dorneich, 2005). There are many mobile contexts in which knowledge of instantaneous cognitive state can be used to drive adaptive systems. Yet moving these technologies to mobile contexts remains a challenge if they are to be used in these types of operational environment. Examples include pilots, dismounted soldiers, ground vehicle operators, (Macuda, et al., 2006; Schnell, et al., 2006, Snow et al., 2006, Dorneich, et al., 2005). In all cases, if a truly adaptive system that manages information flow is implemented, the ability to operate in a dynamic operational situation along with a high degree of fidelity in the cognitive state assessment and temporal resolution is needed.

1.2 Dismounted Soldier Domain

The U.S. Department of Defense (DoD) has embarked on a process of change called Transformation to create a highly responsive, networked, joint force capable of making swift decisions at all levels and maintaining overwhelming superiority in any battle space (Parmentola, 2004). In response, the U.S. Army is shaping its Future Force to be smaller, lighter, faster, and smarter than its predecessor. One of the core capabilities of the Transformation is the availability of netted communications enabling information sharing and real-time collaboration enhancing the kind of situational understanding that drives decisive actions. The Future Force Warrior will have unparalleled connectivity to build situation awareness right down to the individual soldier. Mission success will be dependent on the individual Warfighter's ability to sort through the vast array of continuous information flow afforded by a full range of netted communications.

The increase in information flow will not come without a cost, however. Effective use of these information sources is constrained by the limitations of the human cognitive system. The real-time, dynamic exchange of information in a netted communications environment can be expected to increase the likelihood of information overload such that the postulated information superiority becomes rather a profound liability. The potential data overload coupled with the efficiency of information flow required in executing Army doctrine, places on over-reliance of critical information throughput on a single point of contact, the individual Soldier. One way to ensure that Soldiers are supported appropriately is to develop adaptive information management systems to support superior situation awareness on the battlefield. The efficacy of such a system is contingent on reliable and timely cognitive assessment.

Conducting military manoeuvres in operational environments, such as urban terrain, often does not allow an individual to remain stationary and can demand simultaneous cognitive and physical activity. Consequently, difficulties related to processing of physiological signals in real-world settings include factors associated with both participant motion and the operational environment itself. In harsh environments, Soldiers are subject to extremes of motion, multiple physical and mental stressors, and a wide range of cognitive activities (long periods of vigilance punctured by extreme periods of activity). Thus any approach to real-time assessment of cognitive state has to be robust to motion and noise artefacts.

1.3 Challenges Inherent in Mobile Cognitive State Classification

The effectiveness of neurophysiologically triggered adaptive systems hinges on reliable and effective signal processing, accurate cognitive state classification, and well-designed mitigation strategies. While this presents a difficult technical challenge in any context, these concerns are particularly pronounced in a system designed for mobile contexts. Assessment of an operator's state can be notably more difficult if the operator is permitted to move freely to perform cognitive tasks in conjunction with physical tasks. What is already a difficult problem – gathering clean and robust signals upon which to classify cognitive state – is further complicated by signal artifacts induced by motion.

Given the potential of augmented cognition systems in mobile contexts, methods have been developed to classify cognitive state in ambulatory contexts. This paper will describe the challenges inherent in mobile cognitive state classification, including the ability to collect robust and clean signals, the ability to create a mobile computing and data processing infrastructure, the ability to reliably classify cognitive state, and the ability to experimentally assess the accuracy and specificity of the algorithms in a mobile operational setting. The work described in this paper was developed for the dismounted Soldier – potentially one of the harshest, most mobile application domains for cognitive state estimation. Any cognitive state classification solution in this domain must be portable, efficient, and robust to extremes of conditions and motion. A robust solution that meets the challenges of this domain would result in techniques that are applicable to almost any other domain where motion is a key component of the operator's work environment.

1.3.1 Hardware

Sensors will have to be ruggedized and miniaturized as Augmented Cognition systems transition to the field. In addition, mobile processing requirements will have to be minimized.

1.3.2 Software: Real-Time Signal Processing Challenges

Utilization of research methods involving EEG in mobile operational environments necessitates the use of real-time algorithms for signal detection and removal of artifacts. Although real-time signal processing and classification of the EEG has been implemented previously (Gevins & Smith, 2003; Berka, Levendowski, Cvetinovic, Petrovic, et al., 2004), it has not been realized in a truly mobile, ambulatory environment.

Inferring cognitive state from non-invasive neurophysiological sensors is a challenging task even in pristine laboratory environments. High-amplitude artifacts ranging from eye blinks to muscle artifacts and electrical line noise can easily mask the lower amplitude electrical signals associated with cognitive functions. These concerns are particularly pronounced in the context of ongoing efforts to realize neurophysiologically driven adaptive automation for the dismounted ambulatory soldier. In addition to the typical sources of signal contamination, mobile applications must consider the effects of artifacts induced by shock, cable movement, and gross muscle movement. Specifically, artifacts related to participant motion include high-frequency muscle activity, verbal communication, and ocular artifacts consisting of eye movements and blinks; whereas artifacts related to the operational environment include instrumental artifacts such as electrical noise that created interference with the EEG signal (c.f. Kramer, 1991).

1.3.3 Experimental Assessment: Scenario Design Challenges

In addition to the practical and system configuration challenges faced when moving from the laboratory to field studies, there are issues of experiment control and the characterization of cognitive state in less constrained environments. It is essential to select tasks that are both operationally relevant and afford reasonable adaptations that improved performance. In the laboratory, it is possible to develop simple tasks where workload is manipulated precisely and consistently. Additionally, a user's performance can be collected and evaluated accurately. This makes it relatively easy to establish ground truth about a user's likely workload. However, when developing operationally relevant tasks in a field environment, it becomes substantially harder to manipulate workload precisely and to interpret and assess a user's performance without compromising operational realism.

2 METHOD

2.1 Augmented Cognition System

The Augmented Cognition closed-loop integrated prototype (CLIP) is depicted in Figure 1. The architecture is made up of three principal components: the cognitive state assessor, the augmentation manager, and the display and automation systems. These systems interact with the environment (either a virtual environment in early evaluations, or the real operational environment). The Cognitive State Assessor (CSA) combines all the psychophysical signals, and after processing them to remove noise artifacts, classifies cognitive state. The Augmentation Manager receives the cognitive state assessment, and in conjunction with any other context information available to it, makes decisions are made on how to adapt the work environment (via interface or automation management) to optimize joint human-automation performance. The Augmentation Manager invokes mitigations, either by changing the way information is displayed on the Human Machine Interface (e.g. changing the modality of incoming radio messages to text), or by invoking automation to take over parts or all of certain tasks (e.g., invoking a tactile navigation cueing system).

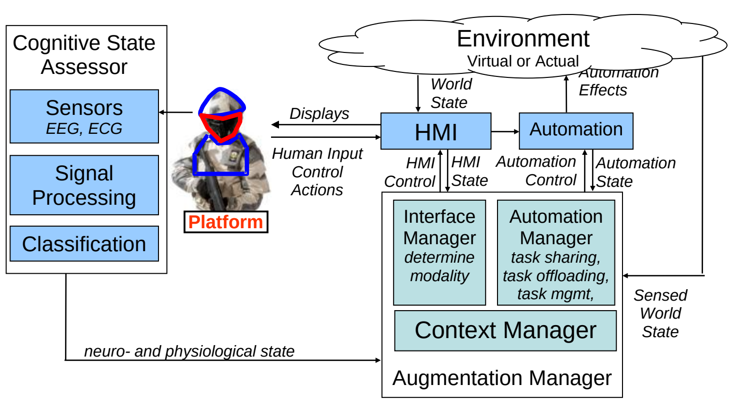

Figure 1. CLIP architecture.

2.2 Research plan

The research done to address many of the challenges described above focused on two parallel and complementary thrusts: cognitive state assessment (CSA) and mitigations (see Figure 2). Each iteration builds on the lessons learned from the previous iteration, with the goal of moving both CSA and mitigation technologies from the lab to the field.

Iteration 1 of the program was dedicated to proving that it was possible to reliably determine cognitive state in real time based on brain imaging (e.g., fNIR), external brain monitoring (e.g., EEG), body sensing (e.g., ECG-based arousal), and eye measures (e.g., pupillary reflexes), and to close the loop by driving adaptive automation with assessments of operator cognitive state. Phases 2-4 then continuously added realism by moving evaluations from a stationary virtual environment to a (semi-)mobile VE, outside to the field, and finally to the real operational environment. Throughout the phases, cognitive state classification has been matured from a laboratory, stationary setup with 32 EEG leads to a wireless mobile system with six leads. Multiple mitigations were explored throughout the phases, covering the gamut from task offloading, task sharing, and task and information scheduling to modality management.

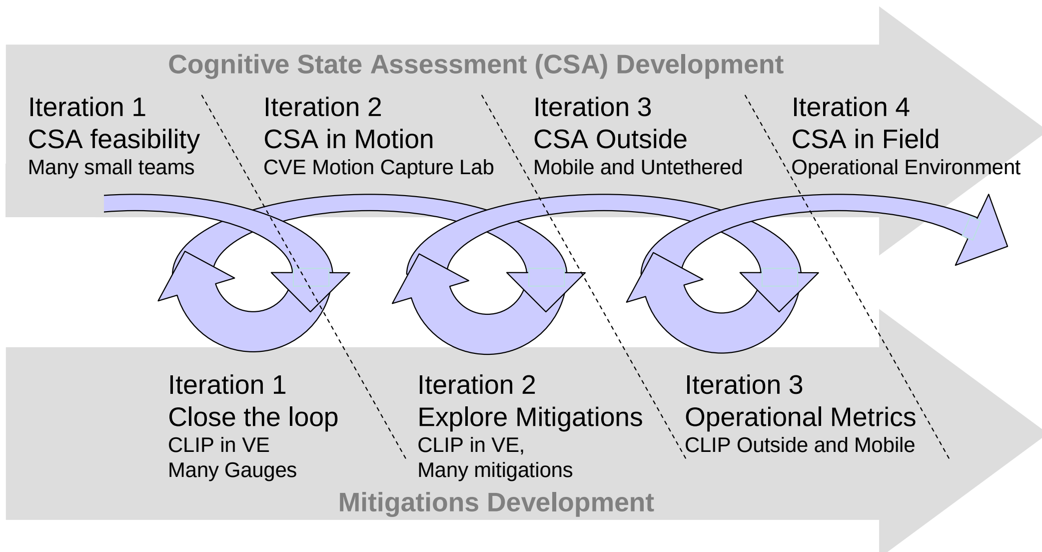

Figure 2. Spiral development over two principle research thrusts.

3 RESULTS

Iteration 1 CSA was the culmination of several years of research by many researchers (St. John, Kobus, & Morrison, 2003). Honeywell became involved in the Augmented Cognition Program in the Iteration 1 Mitigations Development phase by building a closed look system. .Space limitations preclude a discussion of each experiment over the four iterations. However, Table 1 briefly describes each experiment, the conditions under which it was conducted, and some results. The final column provides a reference where the reader can find more detailed descriptions of each experiment.

Table 1. Overview of experiments that developed CSA and Mitigations from the lab to the field.

Exp.	CSA	Mit.	Env.	Results	Reference
Iteration 1 CSA	20 Individual Cognitive gauges	None	Desktop PC task	11 of 20 gauges tested" successfully identified changes in cognitive activity during the task"	St. John, Kobus, & Morrison, 2003
Iteration 1 Mitigation	Engagement Index, Arousal Gauge, Stress Gauge, XLI Gauge, P300 Gauge	Communications Scheduler (CoS)	Virtual Environment (VE)	Under high workload conditions, the communications scheduler produced better situation awareness	Dorneich, Whitlow, Ververs, Carciofini, & Creaser, 2004
Iteration 2 CSA	Engagement Index, Arousal Gauge, XLI Gauge	CoS	Motion Capture VE	Gauges in combination successfully drove mitigations; able to classify despite mobility.	Whitlow, Ververs, Buchwald, & Pratt, 2005
Iteration 2 Mitigation	Engagement Index, Arousal Gauge, Stress Gauge, XLI Gauge, P300 Gauge	CoS; Tactile Navigation Cueing; Medevac Asst; Target Detection	VE	All Mitigations improved performance (e.g. 100% improvement in message comprehension; 125% in situation awareness;	Dorneich, Whitlow, Mathan, Carciofini, & Ververs, 2005
Iteration 3 CSA	Engagement Index, Arousal Gauge	None	Outside (grassy field)	Successfully collected usable data out in the field; identified the need for real-time artifact detection and removal	Dorneich, Whitlow, Mathan, Ververs, Pavel, & Erdogmus, 2005

Iteration 3 Mitigation	pattern classification approaches driven by estimates of spectral power from EEG	CoS; Tactile Navigation Cueing	Outside (grassy, wooded field)	significant improvement due to mitigations on tasks with operational relevance, in a mobile setting	Dorneich, Ververs, Mathan, Whitlow, Carciofini, & Reusser, 2006
Iteration 4 CSA	Support vector machine based on spectral power from EEG; fused EEG and ECG classification	Open-loop feedback of cognitive state to commander	Aberdeen Proving Grounds MOUT site	From 83% to 95% base classification accuracy for fused EEG and ECG	Mathan, Whitlow, Dorneich, & Ververs, 2007

3.1 Hardware

In Iterations 1, 2, and 3 the EEG collection system was a 32-channel BioSemi system where the sensors were cabled to a stand alone box. However, for Iterations 3 and 4 we migrated to a six-channel wireless EEG cap manufactured by Advanced Brain Monitoring (ABM) that was integrated into Honeywell's information architecture. The move to wireless helped considerably in reducing noise artifacts induced by events like cable-sway, as well as facilitating a mobile ensemble that could be integrated into a single backpack on a Soldier.

ECG signals were collected on various devices. In Iterations 1 and 2, ECG was collected on both a custom mobile Arousal Meter hardware, and with Cardiax PC ECG device running CardioSoft software. Iteration 3 used only the mobile Arousal Meter hardware. In Iteration 4, ECG was collected from the Hidalgo's Vital Signs Detection System (VSDS), a Bluetooth enabled system integrated into the Army's Warfighter Physiological Status Monitoring system. Since the classification algorithms are agnostic to the source of the signals, we were able to choose the hardware that produced the best signal quality in the environment in which it was being tested.

The Mitigations Development Iteration 1 evaluation was conducted on seven high-powered PCs wired together to run the five cognitive gauges, the architecture that synchronized all the signals, the ECG and EEG collection systems, the virtual environments, and the Communications Scheduler (mitigation). Each new evaluation's requirements drove a consolidation and reduction in the processing power and number of computers, until the final Iteration 4 evaluation was run on one mobile computer where all sensors communicated wirelessly.

3.2 Software

Artifact detection and reduction, necessary to create a "clean" signal that can be classified, is driven by a consideration of the characteristics of the noise artifacts themselves. How the noise artifacts are handled depends on where the noise lies in the frequency band in relation to the signal (see Figure 3).

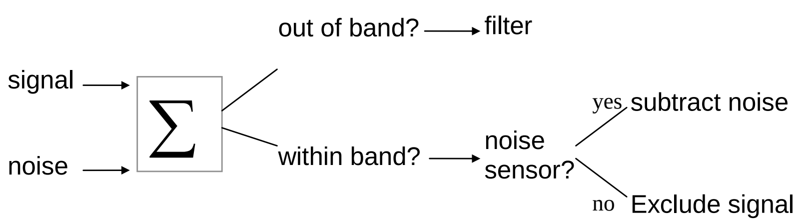

Figure 3. Artifact detection and reduction depends on the characteristics of the noise in relation to the signal.

Noise signals that lie out of the band of the signal of interest can be removed with filtering. Out-of-band artifacts, such as DC drift and 60Hz line noise, typically have well-known characteristics, and can be filtered out easily. Noise artifacts that lie within the same band as the signal require more sophisticated artifact detection and reduction. If the noise can be measured (e.g. eye blinks measured via a dedicated ocular sensor), the sensor data can be subtracted to de-contaminate the signal. However, when the noise cannot be directly measured, adaptive filtering can be applied to estimate the noise. When adaptive filtering is not feasible, the noise should be detected and the resultant data rejected so as not to compromise downstream classification. Signal processing approaches are specific to the signal type, and dependent on factors such as signal-to-noise ration and how specific artifacts affect the signal of interest.

The classification approach evolved over the iterations. Initial iterations focused on individual cognitive gauges, each driven from a subset of available physiological and neurophysiological signals. While the gauges as a group were moderately successful in driving the mitigations of the evaluations in Iterations 1 and 2, they did not show the kind of reliability and specificity needed to make moment-to-moment decisions based on cognitive state. Iterations 3 focused on pattern recognition algorithms that improved the accuracy and reliability of cognitive state assessment, but still required lengthy training sessions. (Mathan, Mazaeva, Whitlow, Adami, et al., 2005). Iteration 4 focused on discriminate function analyses approach, and was employed in the fully operational, mobile task evaluation. The approach used a support vector machine to discriminate between periods of low and high workload (Mathan, Dorneich, Whitlow, & Ververs 2007). In addition, the Iteration 4 CSA approach explored fusing EEG and ECG sensor data. Sensor fusion utilizes multiple sources of sensor data to create the fusion at the sensor level before the discriminate features are calculated. Composite classifiers employ multiple classifiers, where the output of each classifier is used to create a final determination (Dorneich, Mathan, Ververs & Whitlow, 2007). For more detail, see the companion paper in these proceedings (Mathan, Whitlow, Dorneich, & Ververs, 2007).

The mitigations development produced many lessons learned. Moment-to-moment, real-time, cognitive state assessment enables systems adaptation on the same time scale. Hence there is the potential for "strong" automation to mitigate undesirable situations where the cognitive state of the operator is misaligned with the current task demands. Like work in Associate systems, Augmented Cognition systems have principally targeted “higher criticality” applications, where consequences of error can have larger and more immediate consequences to the operator making decisions and taking actions (Miller & Dorneich, 2006). While the management of automation intervention promises 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 nature of Augmented Cognition adaptive systems mitigation is that they are "turned on and off" as needed, depending on the operators' cognitive state and context. 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 is only exacerbated in situations where the system is making independent decisions on when to invoke automation. Specific strategies to minimize the negative impact of automation were developed as part of the mitigations design. Automation etiquette therefore is an important aspect of mitigations design (Mathan & Dorneich, 2005).

3.2.1 Experimental Assessment

In order to calculate the accuracy of the classification approach, classifier results are compared to "ground truth." Ground truth is defined as the actual workload experienced by the participant at any given moment. The output of the classifier at any moment is then compared to the ground truth to determine the accuracy of the classifier. Experiments in the early iterations benefited from a task environment that afforded control over the task environment. From simple, well understood laboratory tasks (e.g. n-back tasks, cue and respond tasks, memory tasks) the task environment migrated to more operational tasks that were less controllable. By iteration 4, where participants were conducting normal military tasks in a free-play environment, experimenters no longer had control over task loads, or the ability to introduce workload probes without disrupting the free flow of events.

The principal issue in scenario/task design is to create detectable and sustained (5-10 minutes) of high or low workload multiple times within any single data collection session. Ground Truth was difficult to obtain directly from task characteristics, as discussed above. Thus indirect methods must be used. There are two classes of indirect methods: observation and subject self-reporting. Often human experts can observe the experiment and determine ground truth based on their knowledge of the task demands and the demonstrated behaviour of the subject. Subjects can do a post-scenario cognitive walkthrough, often with time-stamped videos, and report their self-assessment of their level of cognitive workload. These methods can be used individually or together to produce ground truth.

There are two important considerations when using either or both techniques. First, both the ground truth data and the classification data must be on a common (and accurate) time-stamp system, to allow for a moment-to-moment determination of classification accuracy. Second, subjects doing self assessments should have "high workload" and "low workload" defined in an understandable manner. Operationally, low workload can be defined as times when the participant would have been able to take on additional cognitive tasks. High workload can be defined as times when it was not possible for the participant to take on any additional tasks and in fact was not able to handle the current task load to the best of his or her ability.

4 DISCUSSION

Over the course of the development of augmented cognition adaptive systems for mobile settings, lessons have been learned in five key areas. Mobility requires thorough advanced signal processing algorithms which are essential for the use of the measurement of cognitive metrics, particularly in the harsh operational environment to remove or identify noise artifacts. There is not a one size fits all approach to cognitive state classification. Individualized measurements are necessary for each individual. In addition, due to the non-stationary of physiological data over time regular baselines will need to be captured to obtain high level of classification accuracy. There is a need to further ruggedize the physiological and neurophysiological sensors and sensor system to enable deployment of this capability. The assessment of classification effectiveness will always require the evaluation to capture context of the mission and task as well as user feedback as a basis of ground truth information. In addition to a complete understanding of the target environment, thorough interviews with participants and multiple raters of ground truth classification will help minimize any error in cognitive state classification due to poor insight into the cognitive loading requirements of the task environment. Finally, mitigation must be developed with a strong sense of automation etiquette since these systems run the risk of introducing many of the problems observed in traditional human-automation interaction contexts, but in a more exacerbated fashion due the strong, moment-to-moment automation interventions made possible by a reliable assessment of cognitive state.

5 ACKNOWLEDGMENTS

This paper was supported by contract with DARPA and the U.S. Army Natick Soldier Center, DAAD16-03-C-0054, for which Cmdr Dylan Schmorrow (Phase 1-3), and Dr. Amy Kruse (Phase 4) served as Program Managers 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.

Over the iterations of this program, there were many contributors we would like to acknowledge, including: Danni Bayn, Chris Berka, Mark Buller, Dr. Kelly Burke, Jim Buxton, Cynthia Blackwell, Ami Bolton, Barbara Brockett, Ben Buchwald, Jim Carciofini, Roger Carff, Janet Creaser, Mike Darga, Gene Davis, Bob DeMers, Fred Dupont, Don DuRousseau, LT Joseph Cohn, Col. Beau Freund, Kyle Gabler, Rob Gordon, Tony Ham, Dr. Tamara Hayes, Jeremy Higgens, Jessica Hodgins, Dr. Adam Hoover, Matt Johnson, Chris King, Dr. Scott Kerick, Maj. Latzka, Marty Loughry, Daniel J. Levendowski, Dennis Magnifico, Dr. Caroline Mahoney, Adam Malhoit, Steve Mullen, Dr. Eric Muth, Ken Parham, Dr. Lucas Parra, Dr. Randy Pausch, Dr. Misha Pavel, Justin Pisani, Jason Pratt, Dr. Denny Proffitt, Dr. Anil Raj, Trent Reusser, Reta Reynolds, Rose Mae Richardson, Dr. Bill Rogers, Jeff Rye, Dr. Paul Sajda, Dr. James Sampson, Jesse Schell, Dr. Roy Stripling, Allison Styer, Dr. Akasha Tang, Paul Tennant, and Bill Tharion.

6 REFERENCES

Berka, C., Levendowski, D., Cvetinovic, M.M., Petrovic, M.M., Davis, G., Lumicao, M.N., Zivkovic, V.T., Popovic, M.V., & Olmstead, R. (2004). Real-time analysis of EEG indices of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human Computer Interaction, 17*(2),151–170.

Dorneich, M., Whitlow, S., Ververs, P.M., Carciofini, J., and Creaser, J. (2004) Closing the Loop of an Adaptive System with Cognitive State. *Proceedings of the Human Factors and Ergonomics Society Conference 2004*, New Orleans, September.

Dorneich, M.C., Whitlow, S.D., Mathan, S., Carciofini, J., and Ververs, P.M. (2005). The Communications Scheduler: A Task Scheduling Mitigation For A Closed Loop Adaptive System. *11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Dorneich, M. C., Whitlow, S.D., Mathan, Ververs, P.M., Pavel, M., and Erdogmus, D. (2005). DARPA Improving Warfighter Information Intake under Stress - Augmented Cognition: Phase III Final Report. *Technical report for DARPA Augmented Cognition* Phase 3 under contract DAAD16-03-C-0054, Honeywell Laboratories, December 31, 2005.

Dorneich, M., Whitlow, S., Ververs, P.M., & Mathan, S. (2007). "Iterative Experimental Design to Mature Cognitive State Classification Techniques from the Laboratory to Field Settings", *Foundations of Augmented Cognition, 4nd Edition* (D.D. Schmorrow, D.M. Nicholson, J.M. Drexler, & L.M. Reeves, Eds.). Arlington, VA: Strategic Analysis. pp. 23-32.

Dorneich, M.C., Ververs, P.M., Mathan, S., Whitlow, S.D., Carciofini, J., and Reusser, T. (2006). "Neuro-Physiologically-Driven Adaptive Automation to Improve Decision Making Under Stress," *Proceedings of the Human Factors and Ergonomics Society Conference 2006*, San Francisco, CA.

Dorneich, M. C., Mathan, Ververs, P.M., & Whitlow, S.D. (in preparation). Cognitive State Estimation in Mobile Environments. In D. Schmorrow and K. Stanney (Eds)., *Augmented Cognition: A Practitioner's Guide*.

Geddes, N.D. (1985) Intent inferencing using scripts and plans. In *Proceedings of the First Annual Aerospace Application of Artificial Intelligence Conference*, Dayton, OH. September.

Gevins, A., & Smith M. (2003). Neurophysiological measure of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 113-132.

Kramer, A. (1991). Physiological metrics of mental workload: A review of recent progress. In D. Damos (Ed.), *Multiple Task Performance* (pp. 279-328). London: Taylor and Francis.

Krebs, H., Brashers-Krug, T., Rauch, S., Savage, C.R., Hogan, N., Rubin, R.H., Fischman, A.J. & Alpert, N.M. (1998). Robot-aided functional imaging: application to a motor learning study, *Human Brain Mapping*, *6*, 59-72.

Macuda, T., Poolman, P., Schnell, T., Keller, M., Craig, G., Swail, C., Jennings, S., Erdos, R., Allison, R., Lenert, A., & Carignana, S. (2006), Neural activity during a simulated rotary-wing flight used expert test pilots. In D. Schmorrow, K. Stanney, & L. Reeves (Eds.). *Foundations of Augmented Cognition* 2nd Ed. Arlington, VA: Strategic Analysis.

Mathan S, Mazaeva N, Whitlow S, Adami A, Erdogmus D, Lan T, & Pavel M.(2005). Sensor-Based Cognitive State Assessment in a Mobile Environment. *11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Mathan, S. & Dorneich, M. (2005). Augmented Tutoring: Improving military training through model tracing and real-time neurophysiological sensing. *11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Mathan, S., Whitlow, S.D., Dorneich, M.C., & Ververs, P.M. (2007). "Neurophysiological Estimation of Interruptibility: Demonstrating Feasibility in a Field Context", Proceedings of the

Miller, C.A. & Dorneich, M.C. (2006), From Associate Systems to Augmented Cognition: 25 Years of User Adaptation in High Criticality Systems. *Proceedings of the 2nd Augmented Cognition International*, San Francisco, CA.

Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 5-20.

Parmentola, J. A. (2004). Army transformation: Paradigm-shifting capabilities through biotechnology, *The Bridge, 34(3)*, The National Academy of Engineering of the National Academies. Retrieved January 28, 2005 from http://www.nae.edu/NAE/ bridgecom.nsf.

Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R., DiNocero, F., Prinzell, L.J. III (2001). The Efficacy of Physiological Measures for Implementing Adaptive Technology. *NASA Technical Report NASA/TP-2001-211018*.

Schnell, T., Macuda, T., Poolman, P., Craig, G., Erdos, R., Carignan, S., Allison, R., Lenert, A., Jennings, S., Swail, C., Ellis, K., & Gubbels, A.W. (2006). Toward the "cognitive cockpit:" Flight test platforms and methods for monitoring pilot mental state. In D. Schmorrow, K Stanney, L. Reeves (Eds.). *Foundations of Augmented Cognition* 2nd Ed. Arlington, VA: Strategic Analysis.

Smith, P.J., McCoy, E.C., Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: the effects of user performance. *IEEE Transaction on Systems, Man, and Cybernetics--Part A*, Vol. 27, No. 3, May 1997.

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.* Arlington, VA: Strategic Analysis.

St. John, M., Kobus, J.G., & Morrison, J.M. (2003). DARPA Augmented Cognition Technical Integration Experiment (SSC-SD TR 1905). Available at: <http://www-tadmus.spawar.navy.mil/tr1905.pdf>.

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. Arlington, VA: Strategic Analysis.

Whitlow, S., Ververs, P.M., Buchwald, B., & Pratt, J. (2005). Scheduling Communications with an Adaptive System Driven by Real-Time Assessment of Cognitive State. *11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Dorneich, M., Whitlow, S., Ververs, P.M., & Mathan, S. (2007). "Iterative Experimental Design to Mature Cognitive State Classification Techniques from the Laboratory to Field Settings", *Foundations of Augmented Cognition, 4nd Edition* (D.D. Schmorrow, D.M. Nicholson, J.M. Drexler, & L.M. Reeves, Eds.). Arlington, VA: Strategic Analysis. pp. 23-32.