NASA/TM–20205003444

CAST SE211 Training for Attention Management Final Report

Angela R. Harrivel, James R. Comstock, Jr., Lawrence J. Prinzel, Chad L. Stephens, Kellie D. Kennedy, and Alan T. Pope
Langley Research Center, Hampton, Virginia

July 2021

NASA STI Program Report Series

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- Help desk contact information:

https://www.sti.nasa.gov/sti-contact-form/
and select the "General" help request type.

NASA/TM–20205003444

CAST SE211 Training for Attention Management Final Report

Angela R. Harrivel, James R. Comstock, Jr., Lawrence J. Prinzel, Chad L. Stephens,
Kellie D. Kennedy, and Alan T. Pope
Langley Research Center, Hampton, Virginia

National Aeronautics and Space
Administration

Langley Research Center
Hampton, Virginia 23681-2199

July 2021

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Contents

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

Executive Summary

NASA Langley led the implementation of all outputs of a research Safety Enhancement adopted by CAST in 2014 to address Training for Attention Management.[1] Attention-related human performance limiting states (AHPLS), which may reduce pilot airplane state awareness, may be indicated objectively by covert or physiological markers of limited performance. Existing NASA Crew State Monitoring (CSM) technologies were applied to sense and quantify these covert markers. Monitoring cognitive state changes on a moment-to-moment basis provides quantitative data regarding the onset of AHPLS. CSM uses psychophysiological signals and machine learning to predict cognitive state passively. Also, line-oriented scenarios were designed and implemented in simulated flight experiments to examine the ability of different flight scenarios to induce AHPLS for the purposes of ground-based training. Attention management is considered here as a trainable skill.

A series of Human-in-the-Loop studies, entitled "Scenarios for Human Attention Restoration using Psychophysiology" (SHARP), assessed the efficacy of using CSM as a means of detecting AHPLS while performing the simulated flight scenarios. This report describes recommendations for training to improve the self-detection of the onset of AHPLS, among other methods including: proper scanning, response to alerts, and detection in a co-pilot. Self-detection may be improved via a "SHARP" display of state predictions in near-real-time designed to augment flight instruction and improve debriefs.

These concepts have been subjected to expert review during a workshop and multiple external technology demonstrations. Recurrent themes in the associated feedback include: low-time and cultural issues, high-time issues, training requirements and resource limitations, the expected privacy and comfort considerations, and calls for more customization per pilot to increase the efficiency of training. Next steps are recommended, including simulation realism and mindfulness/mind wandering studies for training, and further sensor development. Various avenues for training may be taken up, from awareness campaigns, to scenario adoption, to full CSM SHARP Display or Instant Replay system implementation. However, all instantiations should be without jeopardy to the pilot trainee.

---

[1] https://www.skybrary.aero/index.php/SE211:_Airplane_State_Awareness_-_Training_for_Attention_Management_(R-D)

Acknowledgements and Disclaimers

This report is submitted as completion of the Technologies for Airplane State Awareness (TASA) work done under SE211. It captures the ideas and work of the full TASA-SE211 team. Team members at the NASA Langley Research Center included Nijo Abraham, James Jones/AMA, Dan Kiggins/NIA, Mary Carolyn Last/AMA, Nicholas Napoli/NIA, and Ken Toro. Team members at the NASA Glenn Research Center included Joanne Walton, Eric Baumann, David Fuller, Dan Gotti/USRA, Steve Hall, Kristen Hauser, Jeff Mackey/VP, and Rick Powis/VP.

We acknowledge the collaboration of Timothy Drake, Daren Gulbransen, Mark Rubin, Trent Smith, Mike Snow, Steve Taylor and James Wilkerson of Boeing, John Duncan and colleagues of American Airlines. We acknowledge the input of Miguel Martin of Pan Am International Flight Academy, James Wilborn of the FAA, and Sophie Bousquet, Benoit Papaix and Vincent Sibelle of Airbus. Additionally, we greatly appreciate all of the 2015 Scenario Workshop attendees, and multiple excellent interns over the years. This work was performed in part under NASA Space Act Agreements SAA1-1155 and SAA1-1112. We also acknowledge Trey Arthur, Steve Williams, Kyle Ellis, Vincent Houston, Dan Johnson/AMA and all our fellow TASA colleagues. Portions of this work were performed in collaboration with the NASA Langley Data Science Team. We gratefully acknowledge the contributions of Charles Liles, Robert Milletich, and Tina Heinich.

Given the inherent nature of research, the recommendations provided at the completion of this research SE are not presented as design specifications. Substantial risks and uncertainties remain. The implementation of these and other ideas and suggestions for future operational use should be undertaken by those properly experienced and directly knowledgeable in education and flight instruction, and under all applicable regulations. Statements, recommendations and references are not intended as an endorsement of any products or institutions. This information has been prepared by the SE 211 team for NASA ARMD using information and feedback provided by a wide variety of professional stakeholders and experts. Subject matter expert feedback is in summary and anonymous form, and complete only to the extent the contacted subject matter experts were willing to provide time, facility access, and information.

Training for Attention Management Safety Enhancement Research - Motivation and Background

Attention-related human performance limiting states (AHPLS) can cause pilots to lose airplane state awareness (ASA), and their detection is important to improving commercial aviation safety. The Commercial Aviation Safety Team reviewed and reported on (CAST, 2014a) international airplane accidents between 2001 and 2010. Eighteen of them were attributed to loss of control inflight (LOC-I). They found that 13 of those 18 accidents, accounting for more than half of the fatalities resulting from LOC-I, involved flight crew loss of ASA. Further, they found that distraction was involved in all 18 of the LOC-I events. Distraction was divided into two types: channelized and diverted attention. When diverted, crew were distracted from aviating by actions and thoughts associated with decision making (sometimes under high workload). When channelized, crew focused on one instrument or response to the exclusion of other important sources of information. Additionally, confirmation bias was described as making a decision based on faulty information or incorrect reasoning which favors one understanding of an event. As a result, research on AHPLS including channelized attention, diverted attention, startle, surprise, and confirmation bias (as selected and defined by CAST) has been recommended in a Safety Enhancement (SE) entitled "Training for Attention Management" (CAST, 2014b; Harrivel, et al., 2016).

Per the Detailed Implementation Plan (CAST, 2014b) released to charter the SE211 research, "The goal is to provide design and training recommendations to the Joint Implementation and Data Analysis Team of CAST (JIMDAT) for review and reference to reduce the occurrence and enable recovery from attention issues, and receive feedback from training providers that training is effective." One output focused on AHPLS detection methods, and a second output addressed training scenarios and AHPLS mitigations. Technical progress and research team recommendations are presented below for each of these outputs. Feedback from aviation safety community subject matter experts follows in a third section. As a member of CAST, NASA reports these findings for the consideration of the JIMDAT and CAST regarding future ground-based line-oriented training designed to reduce accidents by enabling the avoidance of, and recovery from, loss of control in flight due to attentional performance limitations.

The NASA SE211 team aimed to assess the prediction of pilot cognitive state in real time through the use of Crew State Monitoring technology, and to test and evaluate realistic Line-Oriented Simulation scenarios intended to induce attentional human performance limiting states (AHPLS) in pilots using human-in-the-loop (HITL) simulation studies. Training methods have been conceptualized for mitigating the detrimental effects of AHPLS. Key components of the presented state detection and induction techniques have been designed to augment flight instructors, have been tested in motion-based flight simulation, and have been subjected to expert review.

Attention management can be considered as a trainable skill, and there is real-time functional magnetic resonance brain imaging evidence of sustained attention ability improvement with closed-loop training (deBettencourt, et al., 2015). The end goal is pilots better able to manage their own attention while flying based on awareness and training received from instructors while on the ground. Indeed, "extensive research indicates that self-regulation - the ability to direct one's attention, thoughts, moods, and behavior in line with one's personal goals - is among the most critical skills in life." (Mrazek et al., 2018) However, the idea is not new, and game-based training of attention management has been studied. Workload-coping and attention-management skills have been developed through structured video game experience and improved check ride ratings (Hart and Battiste, 1992). Also, "efficient control and management of attention under high task load are argued to be skills that can improve with proper training and generalize to new situations." (Gopher, et al., 1994) Further, rote memorization, for example of standardized test event responses, vs. event recognition abilities may offer limited generalizability in the face of unexpected, off-nominal events (Casner, et al., 2013).

Regarding wearables in general, there has been support for adopting wearable devices for managing workload and pilot monitoring, and to enable single-pilot operations. This includes support for crew alerting, synthetic vision, checklist aids, communications, pilot fatigue and incapacitation monitoring. (Moehle and Clauss, 2015) Further, SITA OnAir's CEO Ian Dawkins and chief strategy and marketing officer François Rodriguez have also expressed support for wearables: "Wearables will first change the way the flight deck and the cabin communicate, says Rodriguez, who is currently the company's lead on wearables within the cabin. No longer will flight attendants need to pick up a wired telephone to speak with the cockpit, with wrist-to-flight-deck talking the way forward." [2]

There exists literature addressing challenges of realistic upset recovery training in simulators (Burki-Cohen, 2010), and recommendations for training skills and techniques for observing pilot behaviors (Research Integrations, Inc., 2011). A full treatment of startle, surprise and distraction provides a definition for startle and surprise, and recommends behavioral therapy mitigations such as conditioned neutral responses (FAA, 2015). Additionally, there has been much work done to improve flight path monitoring (Civil Aviation Authority, 2013; CAST 2013; Flight Safety Foundation, 2014). Distraction, preoccupation/attention-tunneling, and lapses in monitoring are addressed as human vulnerabilities, alongside high workload (as a dominant stressor), tiredness and fitness. However, none address the use of the real-time detection of cognitive states for self-awareness or the augmentation of the instructor with feedback of objective trainee mental state information.

Regulatory information that influenced the team approach includes "Qualification, Service, and Use of Crewmembers and Aircraft Dispatchers," 14 CFR Part 121 FAA-2008-0677; Amendment No. 121-366. This Federal Aviation Administration (FAA) Final Rule revised requirements to allow air carriers to modify pilot training programs, and requires pilot monitoring training to be incorporated into existing requirements for scenario-based flight training. Under CFR 14, Parts 121 and 135, the Advanced Qualification Program (AQP) is "a voluntary alternative to the traditional regulatory requirements for pilot training and checking." [3]

Also, FAA Advisory Circular (AC) No. 120-35D, "Flightcrew Member Line Operational Simulations: Line-Oriented Flight Training, Special Purpose Operational Training, Line Operational Evaluation," Issued March 13, 2015 by responsible office Aviation Flight Standards-200.[4] This AC provided guidelines for the design and implementation of Line Operational Simulations (LOS) for flight crew members, and called for an interdependent relationship between their human factors, Crew Resource Management (CRM), flight operations, and safety initiatives. It describes a means by which LOS scenarios are developed, scripted, tested, and evaluated. In the case of Line-Oriented Flight Training (LOFT) and Line Operational Evaluation (LOE), it allows for approval by the Administrator for use in an operator's training program. The methodology set forth also achieves the FAA mandate to ensure that each certificate holder provides the highest level of safety in the public interest, while meeting the agency's responsibility to reduce or eliminate the possibility or recurrence of accidents in air transportation.

> "The FAA has proposed developing official guidance for airline pilot training that includes a curriculum focused on establishing pilot-monitoring duties. "The FAA will develop guidance defining pilot monitoring duties and responsibilities that air carriers can use to develop pilot training and evaluation. The guidance will address the definition of pilot monitoring in the operational environment, and it will provide the basis for development of a curriculum and syllabus by carriers. The FAA plans to complete this action prior to January 31, 2017," H. Clayton Foushee, director of the FAA's office of audit and evaluation." [5]

Given this background and regulatory environment, the SE211 team approached output 1 as described next.

---

[2] https://runwaygirlnetwork.com/2015/04/20/how-crew-applications-will-make-wearables-worthwhile-in-aviation/

[3] https://www.faa.gov/training_testing/training/aqp/

[4] https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1027170

[5] http://www.aviationtoday.com/av/embedded-systems/86932.html?hq_e=el&hq_m=3200518&hq_l=2&hq_v=0ee6605e5d#.VpZ2hvlsOM9

I. Output 1: “Air carriers will provide research organizations with access to operational expertise that can help improve effectiveness and feasibility of detection methods.”

Technical Progress

Regarding state detection methods, the SE211 team considered that AHPLS which may reduce pilot airplane state awareness may be indicated objectively by covert or physiological markers of limited performance. Existing NASA Crew State Monitoring (CSM) technologies were applied to sense and quantify these covert markers. First, CSM will be further defined and background will be discussed. Second, the definitions of the cognitive states of interest as selected by CAST will be presented. Third, CSM system, development and technical details will be given. Fourth, a series of studies and associated results published thus far will be described. Finally, future applications will be described.

A mental or cognitive state is a frame of mind or set of cognitive processes going on in the mind. Such cognition is usually employed on purpose to accomplish a task in which the human operator is engaged. They are not emotions, not medical conditions, and not related to psychiatry. Cognitive states can come and go in the same way that aircraft energy states change. Monitoring these state changes on a moment-to-moment basis provides quantitative data regarding the onset of AHPLS. CSM uses psychophysiological signals and machine learning to predict cognitive state passively during operational activities. The application of CSM is intended to address natural, human vulnerabilities during off-nominal situations considering increasing traffic in the airspace and complexity in the cockpit. See Table 1 for further characterization of CSM. This approach encourages training around “how to think” about thinking and paying attention, with a focus on managing one’s own head. On any given day, a human pilot can be overcome by unexpected, off nominal events. However, humans have innate ability to detect social cues, e.g., when others are paying attention, confused, or startled. With training and quantitative techniques, this ability can be used to avoid or recover from AHPLS to the advantage of aviation safety.

Table 1. Characterization of CSM.

What CSM is not:	**What the SE211 team is not trying to do:**
A lie detector	Tell trainers they are not good at their job
A psyche probe	Tell pilots they are not good at their job
A mind-reading system	Create robotic or inhumanly-vigilant pilots
A clinical diagnosis tool	Burden the training footprint without credit
Psychoanalysis or personality profiling	Help trainees pay attention during the training itself
Emotion sensing	

A 1992 “analysis of incidents in the Aviation Safety Reporting System database reveals that civil transport flight crew members often relate their mistakes to experiencing certain states of awareness such as absorption and preoccupation. As automated systems become more capable and comprehensive, there is the danger that crew members will spend more time performing a passive monitoring function. Hazardous states of awareness occur most often under just such conditions.” (Pope et al., 1992) Currently, such “advances in aircraft automation have significantly contributed to safety and changed the way airline pilots perform their duties - from manually flying the aircraft to spending a majority of their time monitoring flight deck systems. Despite the importance of pilots’ monitoring skills, the FAA does not have effective processes to assess these skills both in training and during flight. Further, inspectors do not know how to assess a pilot’s ability to monitor the state of the aircraft, beyond observing call-outs

(FAA report AV-2016-013). These hazardous states of awareness, including AHPLS, can be detected with CSM, and still remain worthy of further research.

A major goal of the SE211 Project was to leverage multi-modal psychophysiological sensing for engagement and workload prediction. The relevant neuroergonomic scientific literature also addresses fatigue and emotion sensing (Fairclough & Gilleade, 2013; Thomas et al., 2015; Wilhelm & Grossman, 2010; Wilson et al., 2003). Prior work has not employed multi-modal sensing in a real-time system using operationally-relevant, realistic flight scenarios as undertaken by the team at NASA Langley (LaRC). An underlying premise of the approach of the present project is that no single physiological marker, behavior, or subjective report is sufficient when used alone to assess mental state. A single indicator may result in a false positive or negative assessment, whereas multiple indicators leverage the strength of machine learning techniques, enable convergent validity, and help account for individual differences.

Solovey, Afergan, Peck, Hincks, and Jacob (2015) describe a processing pipeline that includes a calibration phase. Based on this offline process, a model is built for recognizing quantitative levels of a cognitive state and adapting system behavior accordingly in a real-time closed-loop process. Also, Berka et al. (2007) state: "First, it is necessary to define a set of relatively pure tasks that consistently elicit the targeted cognitive states to provide training data for the model and to validate the methods for cognitive monitoring." However, neither of these examples employ tasks designed to induce diverse cognitive states that are qualitatively different from one another, in particular, attention-related human performance limiting states (AHPLS) of interest to CAST; rather, these examples employ tasks that induce different quantitative levels of one particular cognitive state, e.g., cognitive workload.

The NASA SE211 team further proposed and described a means of refining cognitive state classification models based upon subject matter expert judgment and a means of presenting AHPLS detection derived from such models to interested parties in a readily comprehensible fashion.

As introduced above, AHPLS as selected and defined by CAST include: channelized attention, diverted attention, startle / surprise, and confirmation bias. Channelized Attention is defined here to be focused on one instrument or one response to the exclusion of all other relevant inputs, comments, or alerts; "tuning out" any information that may have led to fully understand the problem faced. Channelized Attention is considered a type of distraction by CAST. It is to be focused on one instrument or one response to the exclusion of all other relevant inputs, comments, or alerts. It is to "tune out" any information that may have led to fully understanding the problem faced. Some use the word "Perseveration" to describe similar phenomena (Dehais et al., 2010). Also it may be dubbed "Coning of Attention" or "Fascination" wherein a "pilot who, under stress of attempting to perform a demanding and unfamiliar task, allows their attention to be confined to one aspect of the task." (Benson and Rollin Stott, 2006) Also, "task saturation" may be used, such as: "history has shown that loss of situation awareness (as a result of channelized attention, distraction, or task saturation) is involved in a majority of all Ops error class A mishaps." (Gawron, 2004)

Diverted Attention is considered a second type of distraction by CAST. It is defined here as to be distracted by actions or thought processes associated with a decision.

Startle is defined here as a bodily response to intense and abrupt auditory, visual, or tactile stimuli.

Surprise is defined here as a pre-emotional response caused by a violation of expectancy which, if sufficiently intense, can trigger a brief stress response (Burgoon & Jones, 1976; Cannon, 1932; Rammirez-Moreno & Sejnowski, 2012).

Confirmation Bias is defined here as making a decision based on faulty information or incorrect reasoning (sometimes when task-saturated). It is the cognitive phenomenon in which decision-making is biased early in the decision-making process in favor of one understanding of events or one course of action at the expense of others (CAST, 2014a).

Due to the confounding potential of subjectively high workload as a concomitant stressor and a useful means of inducing AHPLS, the SE211 team chose to include it as a mental state to be induced and detected. Low Workload is defined here as a subjective state when task demand requires minimal resources to complete. Nominal Workload occurs when demand requires secondary/effortful/additionally-motivated resources to complete. High Workload occurs when demand outweighs resources to complete. (Hart & Staveland, 1988)

The induction and detection of cognitive states as defined above were tested during a series of HITL experiments entitled "Scenarios for Human Attention Restoration using Psychophysiology" (SHARP) to assess the efficacy of using CSM as a means of detecting AHPLS. All participants gave informed consent to participate in studies with protocols approved by the Independent Review Board of the NASA Langley Research Center. The series began with CSM data acquisition capabilities being integrated for the first time with fixed-base flight simulation. (Figure 1). Also, "benchmark tasks" were used to induce the states of interest with laboratory-based tasks for classifier training purposes prior to flight simulation. Study achievements included the integration of simultaneous psychophysiological measures during HITL simulation, the delivery of preliminary state classification results, the progress made on real time state identification, and lessons learned before moving up to a costly motion simulator. With respect to a unimodal case using electroencephalography (EEG) signals, multi-modal classification using galvanic skin response (GSR) in addition to the EEG signals produced increased state discrimination accuracy (90% vs. 86%). Using EEG, GSR, and heart rate variability, multi-state (channelized attention, diverted attention, and workload) accuracy averaged 89% (Harrivel et al., 2016a).

Figure 1. Physiological signal measurement during a flight deck simulation in the Crew Systems and Aviation Operations Branch at NASA Langley Research Center.

In parallel with these efforts, in a separate study, participants performed an attentional task while their brain activity was monitored with functional near infrared spectroscopy (fNIRS). Higher state prediction accuracy using support vector machines was observed when noise in the fNIRS hemoglobin signals was filtered with an adaptive model compared to static regression (84% ± 6% versus 72% ± 15%) (Harrivel et al., 2016b). The intention is to mature Frequency Domain fNIRS technology for artifact reduction and operational field use, and incorporate it in future CSM systems. For a description of the fNIRS system under development as part of SE211, please see Appendix B (NASA/TM—2020-220348, Mackey et al, 2020).

Subsequently, a motion-based flight simulation study was undertaken called "SHARP 1" which was implemented in the Research Flight Deck (RFD) of the Cockpit Motion Facility (CMF) at the NASA Langley Research Center. The data collection and analysis system developed and implemented for the SHARP 1 Study consisted of off-the-shelf hardware and software capabilities commercially available at the time of the experiment. Optical fibers (developed in-house) and instrumentation for accomplishing fNIRS were built into the RFD for the left seat only. Other sensors were worn by both the left and right seat pilots during the study. These included electroencephalography (B-Alert by ABM, Carlsbad, CA), galvanic skin response, respiration and electrocardiogram sensors (Nexus Mark II by Mind Media, Herten, Netherlands). Each of the physiological sensors systems were integrated with the eyesDX, Inc. Multi-modal Analysis of Psychophysiological and Performance Signals (MAPPS™) data aggregation and synchronization software suite. MAPPS was further used to process raw data from the physiological sensors systems and, through the use of MatLab and Python-based data analysis capabilities, for real-time assessment of AHPLS. The human operators involved in the study were assessed for characterization of AHPLS during Benchmark Task and Flight Simulation Scenarios. The architecture of the initial data collection and analysis system relied upon the MAPPS software such that each physiological sensor system operated and stored data to a separate computer (Figure 2). This approach to data management resulted in a distributed data storage configuration across numerous computers. The NASA team identified the architecture of the initial data collection and analysis system as a barrier to translating the research system to commercial aviation training facilities for adoption and use. This limitation and development of the system for subsequent experiment and demonstrations is described below around the SHARP 2 study.

AHPLS were detected with CSM using benchmark tasks to induce Channelized Attention, Startle / Surprise, Diverted Attention and Confirmation Bias. AHPLS induction by scenario event is validated in two ways: by the video assessment of subject matter experts using overt behavioral markers and flight performance (to be discussed further below in section II B), and by the convergence of state classifier prediction outputs with the intended state induction via the scenario events of an approximately 90 minute long line-oriented flight training scenario profile (to be discussed further below in section II A). Supervised machine learning classification methods were employed to predict state with respect to the intended state induction. Using a select set of features and a “combined” classifier training method, multistate (channelized attention and startle) prediction accuracy averaged 0.64 +/- 0.14 across thirteen participants and was significantly higher than that for a “separate” training case (Harrivel et al., 2017).

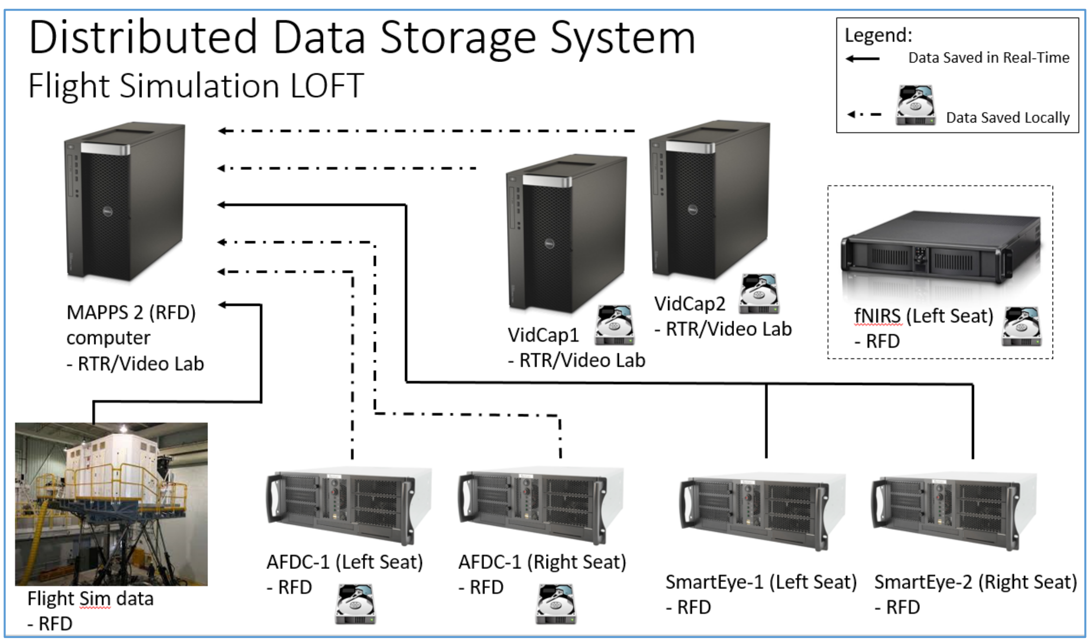


Figure 2. SHARP 1 LOFT data management plan

The team reported on the analysis of fNIRS benchmark task data from SHARP 1 to assess the usefulness of various fNIRS-derived features to discriminate cognitive attentional states using data collected during SHARP 1. Four different types of features were extracted: power spectral quantities, statistical summaries, ratios and differences quantifying anti-correlated brain network activation to represent task engagement, and correlational quantities capturing the co-activation of the networks. Relative feature importance was assessed using multiple ML models. The use of individual and combined subsets were compared to the use of all available features. The combination of feature subsets was found to produce superior state prediction accuracy. The power spectral and task engagement features were found to be most important to state prediction accuracy. This informs the fusion of data from multiple passive physiological sensors, down-selection of features and tuning of machine learning algorithms (Harrivel et al., 2018). SHARP 1 data additionally will be made available for a public data science competition to maximize the state prediction accuracy and determine the best-available machine learning methods. Results are pending competition completion and will be incorporated into future CSM systems.

The CSM Prototype Translational System (CPTS) deployed in external technology demonstrations was developed with Project Echo and as part of the SHARP 2 study as described further below. The sensors chosen for the development of the CPTS are primarily commercial off the shelf (COTS) sensors (Table 2). Some of these sensor packages were originally developed by the manufacturers for self-awareness of one's own mind and body state to help better their mental and physical health. COTS sensors were chosen so that the tool will be able to adapt in accordance to the psychophysiological sensor market. Additionally, the sensors listed here are much smaller, less obtrusive and easier to apply than the sensor suite used in SHARP 1.

Table 2. Implemented sensors for the CPTS.

Name	Manufacturer	Measurements
Muse	Interaxon (Toronto ON, Canada)	EEG, ECG, Head Acceleration
E4 Wristband	Empatica (Boston, MA)	Blood Volume Pulse (BVP), Heart Rate, Hand Acceleration, Skin Temperature, Galvanic Skin Response (GSR)
Spire	Spirehealth (https://spirehealth.com/)	Respiration Rate
Tobii Pro Glasses 2	Tobii (Sweden)	Eye Tracking

Implementation of the sensors listed in Table 2 requires various programming languages and communication protocols. This section provides the list of the software that will be required to build all the tools that are necessary for the seamless transfer and manipulation of the data. These software packages (Table 3) should be installed on the computer that acts as the mainframe for the tool. These selections depended on market conditions and product availability, and additional sensor modalities may be incorporated with further development.

Table 3. CPTS required software packages.

Software	Purpose
Windows 7	Compatibility with sensors and middleware platforms
Lab Streaming Layer	Interfaces with sensors to provide synchronized data to NeuroPype
NeuroPype (Intheon, San Diego, CA)	Acts as the middleware platform that runs the Python based machine learning code
Python 2.7	a) Provides the framework for developing machine learning codes that can be executed in the NeuroPype environment b) Interfaces and enables communication between Muse-io and LSL c) Platform for developing SHARP monitor visualization of state prediction outputs
Muse-io	Provides user interface to the Muse headband and protocol to receive headband signals
Empatica BLE Server	Provides data communication and control for the Empatica E4 wristband

An illustration of the high-level architecture of the system is shown in Figure 3, and discussions of the roles each part play in the system are presented in the following section. The overarching goal of the system is to process the input signals and provide feedback to the flight instructor (or the pilot) via the Scenarios for Human Attention Restoration using Psychophysiology (SHARP) monitor.

In Figure 3 the sensors on the left communicate via Bluetooth to the receiving computer. The receiving computer uses sensor specific software to unpack raw signals which are recorded and transmitted across a Lab Streaming Layer (LSL, 2012) to NeuroPype (Intheon, formerly Qusp). NeuroPype contains the scripts where the machine learning model will classify the state of the user. This state is then displayed to the output screen through the SHARP display monitor. Please see Appendix C for additional CSM Prototype Translational System technical details (LAR-18996-1, US-Patent 10,192,173).

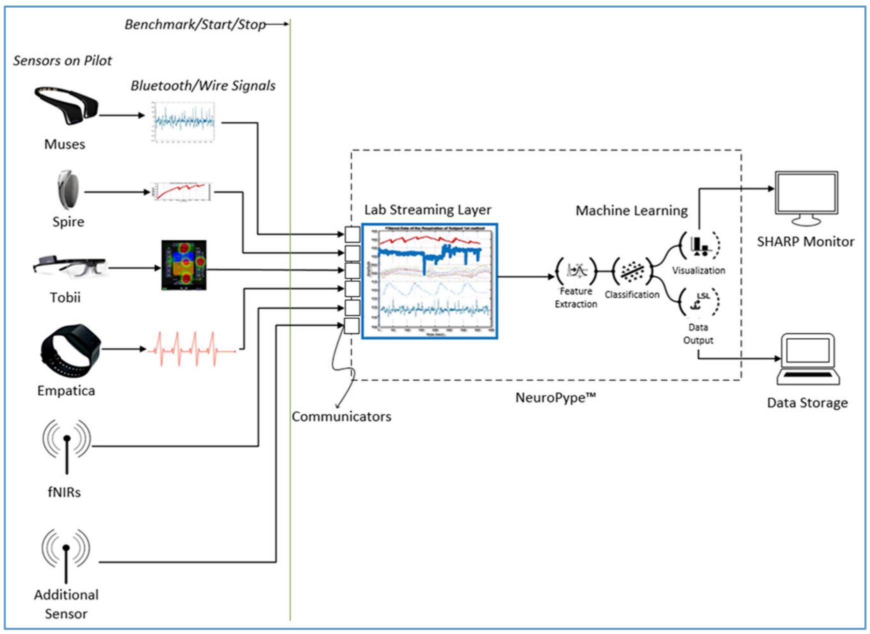


Figure 3: CSM Prototype Translational System high-level architecture

Iterative development of the CPTS which achieved the above-described system was accomplished with SHARP 2 and Project Echo. A major focus of Project Echo was to address the barriers and limitations of the SHARP 1 study data collection and analysis system developed and implemented using the eyesDX MAPPS software suite. The subsequent data collection and analysis system for the final flight simulation study was called "SHARP 2." The system was developed in-house and implemented using an alternative data aggregation and synchronization software suite based in open source platforms LSL and NeuroPype. Furthermore, the SHARP 2 study data aggregation and synchronization software capability greatly reduced the dedication of computers to each physiological sensor system thus simplifying the complexity of the overall system from many computers to a pair of computers: one for benchmark task presentation, and the second for benchmark task operation and data collection and analysis.

SHARP 2 was implemented in the Research Flight Deck (RFD) on a fixed base of the Cockpit Motion Facility (CMF) at the NASA Langley Research Center using the results of Project Echo. Additional differences between SHARP 1 and 2 include: in-house benchmark task automation, the use of two special purpose operational training (SPOT) scenarios in addition to the LOFT for additional state induction opportunities, and trials of less-obtrusive sensors. The sensors are listed in Table 2 above. Importantly, this included switching from the Smart Eye system embedded in the flight simulator to a portable headworn glasses (Tobii Pro) for eye tracking that would not require the modification of flight simulator cabs at line training facilities.

SHARP 2 study aims were to test the automated benchmark tasks, to explore ways to increase the ease of use and portability of sensors while sacrificing minimal loss of state prediction accuracy, to iteratively develop mobile recording systems and consider additional scenarios for future use during demonstrations at airline training simulator facilities. To these ends, participants wore both legacy and less-obtrusive sensors to enable comparative assessment. State prediction accuracy using a small 4-channel electroencephalography (EEG) device (among multiple additional physiological sensors) was found to be comparable to that found with a more cumbersome EEG system for channelized attention at 60%, and to be only slightly lower for startle at 67%. These results warrant continued development of real-time classification of AHPLS using less-obtrusive sensors (Harrivel et al., 2018a). Additional SHARP 2 results are presented around the wake encounter and throttle anomaly events below in section II A. At the completion of SHARP 2 and Project Echo, the CPTS was ready for a series of external technology demonstrations (as part of a "Road Show") described in section III.

The prototype system developed by the NASA SE211 team to translate research into useable concepts and application of that capability into the commercial aviation training context are not the final solutions to address Training for Attention Management challenges. The efforts made by the SE211 team, including NASA personnel and industry/academic partners, are the first steps towards addressing these problems. The SE211 team collaborated with international researchers to write about and present on further reaching capabilities which propose to carry the prototype system into alternative and future applications (Stephens et al., 2018). The CPTS fits into the category of biocybernetic adaptation because it involves adapting technical systems (SHARP Display Instructor tool) to cognitive states (AHPLS). Biocybernetic adaptation involves a "loop upon a loop," which may be visualized as a superimposed loop that senses a physiological signal and inserts its influence into the operator's task at some point (Figure 4) (Stephens et al., 2018). The purpose of the collaborative report and the international conference session (HCII 2018) in which it was presented are to detail the structure of biocybernetic systems regarding the mental states of interest for adaptive systems, their processing pipeline, as well as the adaptation strategies to pave the way towards machine awareness of human state for self-regulation and improved operational performance.

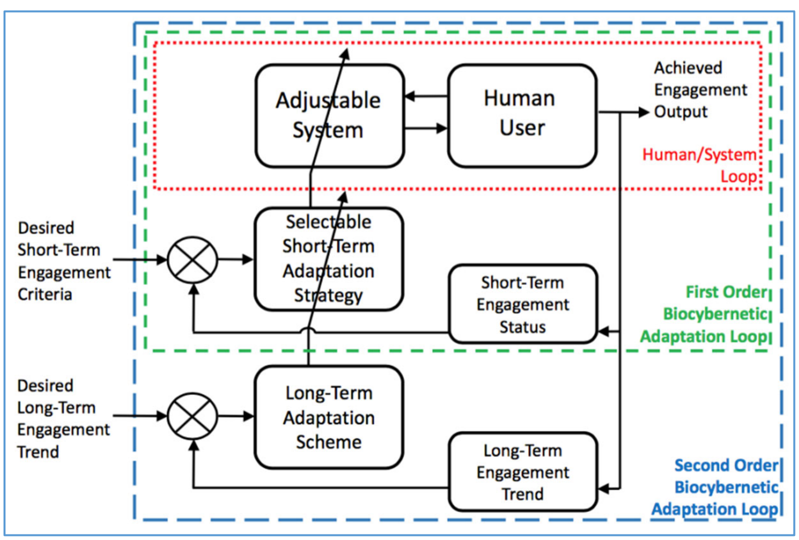


Figure 4. Depiction of multiple levels involved in an adaptive system involving biocybernetic control capabilities.

An important consideration in designing such systems and their use in training contexts is how the end user considers the constructs of workload and attention and their interaction. Traditionally, most of the research has focused on mental workload-based biocybernetical adaptation. The usability of the mental workload construct remains limited. This construct, although theoretically and practically interesting, remains ill-defined (Mandrick et al., 2017) and provides a non-specific and generic index rather like a thermometer. Moreover, mental workload should not be viewed as the resultant of external demand on an individual passively adapting to it, but rather as an active process that depends on the human operator's level of engagement. Thus, human cognitive performance has to be considered as the byproduct of the level of task demand by the level of task engagement. Interestingly enough the concept of engagement is related to a triad of attentional states that are: attentional disengagement, attentional over-engagement and attentional in-engagement. Thus, this concept accounts for neurophysiological and behavioral phenomena, and it can be characterized with portable measurement tools using for instance an EEG engagement index (Pope et al, 1995). The CPTS capability developed and tested by the NASA SE211 team built upon this work and expanded upon it by incorporating multivariate assessment to determine what temporal and magnitude changes in physiological signals reflect trainee state changes that warrant mitigation.

The SE211-team-developed CPTS is biocybernetic adaptation on a broader time scale, proposed to be applied in an aircrew training context by influencing the instructor-trainee interaction. Here the adaptation involves complementing the usual observations of airline training personnel/instructor pilots by informing them, in the real-time training setting, of the occurrence of attention-related human performance limiting states experienced by their trainees. The loop is closed when their state information is conveyed to the trainee as part of each session debrief (Harrivel et al, 2017). Another solution proposed and described is the dynamical reallocation of task between automation and human operators to the triggering of cognitive counter-measures (Roy et al, 2016; Dehais, et al 2015). Such a capability would have implications for highly automated or autonomous systems which require human operators to provide input to or minimally interact with. Both of the potential biocybernetic adaptations described are examples of designs which fit into a catalogue of adaptive solutions to mitigate the decline in attentional performance and improve human performance.

Findings and Recommendations for Training

The team's initial source of feedback regarding these SE211 concepts was from the professional line pilots who were recruited to participate in our HITL studies, and from pilots undergoing training which the SE211 team was privileged to observe. Participant responses (listed in this section) cover the taboo of psychological topics, the effectiveness of current training practices, simulator realism, self-awareness and workload management. The LOFT brought underlying issues that may diminish flying performance and personal workload saturation thresholds into focus, vs. technical flying skill. Training pilots to recognize and address the potentially insidious root causes of mistakes is recommended vs. individual procedural, protocol, and execution mistake correction. Elevating this awareness and creating an ongoing curriculum similar to that of CRM[6] in the late 1980's is recommended.

- This experiment has the potential "to open the can of worms that holds the psychological and biomedical side of flying airplanes which pilots do not want to talk about let alone address in meaningful ways. … In conversations with Line Check Airmen, training department heads, chief pilots, and line pilot colleagues about flight operations issues, these subjects never come up as potential causal factors in performance. This is likely because they touch on difficult issues of breaking traditional training protocols and matrixes, or because they are simply too hard to deal with from a personal perspective."
- Pilots absorb information as well as their stress level allows. Being aware of your own perception and therefore your own SA limits is important.
- Sometimes you have to channelize for a period of time without interruption or distraction to problem solve. In these times you narrow down to what you need to survive.
- Pilots must be thinking at 16 miles per minute. They have scans, flows, procedures and check lists to stay organized. With training, they can handle more but still safely.
- The PF and PM helped each other be aware of their locus of attention by stating things like "I'm inside for a little bit."
- Situational Awareness[7] vs. knowledge was discussed.
- Alerts seemed sometimes to be problematic: "EICAS is in the way." "We acknowledge, we ignore."
- The systems/tasks shed by the automation may be unexpected.
- To get appropriate/representative physiological responses. State-dependent learning and sim realism issues. Although it seems the simulator can overload, fatigue and disorient.
- Sometimes experience is the best teacher, e.g., if a checklist is "in the way."
- Prioritizing what to do now vs what to do later, what to delegate, when to hold, pull yourself out and reset your assumptions was discussed in the context of workload management. Being non-sequential is for those who are more advanced in experience, although not following SOPs can lead to problems with shared first officer SA, especially if low-time.
- The trainer noted that growth occurs when you are stressed. However, strict scripts (FAA, union restrictions) don't allow customization to achieve that for all pilots.
- The issue of "selling" type ratings vs. quality training arose numerous times.

---

[6] definition/components: communications, situational awareness, problem solving, decision making, and teamwork (Helmreich et al., 1999)

[7] definition: "the perception of the elements in the environment within a volume of and space, the comprehension of their meaning, and projection of their status in the near future" (Endsley, 1995)

II. Output 2: “Air carrier training organizations, in conjunction with research organizations, will use these methods to develop scenarios and training-based mitigations.”

A. Training Scenario Development

Technical Progress

The SE211 team endeavored to examine the ability and efficacy of different flight scenarios to induce AHPLS. This test assessed whether flight scenarios can be created to reliably induce human performance limitations, for example for use in training. NASA’s approach has been to design and implement realistic, high-workload training scenarios and training-based mitigations which are practically feasible for complementary use with psychophysiology and behavioral monitoring capabilities. This is to provide “effective training feedback to flight crews” and to enable NASA, air carriers and training organizations to work together on recommendations on new definitions and/or revisions to current training syllabi as required to accommodate new LOFT scenario design methods, use of CSM, and recommended training-based mitigations.

The CAST directive requested NASA to design these training-based mitigations and conduct research on how psychophysiology may be employed, in a practical and useful manner considering the training environment of commercial airline pilots. Relevant training areas include proper instrument scan for glass flight deck instruments, detection of hazardous states of awareness in themselves and in other crew members, and how to optimize attention awareness. This optimization would improve responsiveness to, and the proactive avoidance of, loss of attitude and energy aircraft states. Further, remaining attentively engaged supports better compliance with flight deck alerts with proper standard operating procedures when they do occur. The goal is to provide more objective methods for:

- Inducing AHPL during LOFTs with purpose to expose pilots to the potential opportunities that may arise
- Having more quantitative measures to better detect the impending onset of hazardous states of awareness
- Using the LOFT scenarios and CSM technologies together to provide the foundation for new and innovative training methods that enable pilots to learn how to recognize AHPL states and potential operational conditions they may encounter, and effective techniques for attention management

The research safety enhancement approach taken includes establishing basic concepts and theories, developing and validating new concepts in collaboration with university researchers, proving innovative techniques through analysis, simulation, and laboratory testing, and, ultimately, demonstrating the most promising concepts in operational environment tests. Measurement technologies to assess Hazardous States of Awareness include behavioral (i.e., task performance, control inputs, errors), subjective (self-reported), and physiological techniques. A major focus is on direct measurement of brain electrical activity (measured with EEG) and hemodynamic activity (measured with functional Near Infrared Spectroscopy (fNIRS)). Additional CSM physiological measures also figure prominently in the research and applications. These measures were collected to enable the application of machine-learning-based analysis techniques to the physiological and performance data. Resulting training-based mitigations are described in section II B.

The “SHARP 1” study, entitled “Scenarios for Human Attention Restoration using Psychophysiology (SHARP)”, was a HITL study conducted using custom-developed state induction training scenarios. Simulated flight scenario event sets were designed in collaboration with subject matter experts and line-operational commercial airline pilots. A research laboratory-based system for data acquisition and real time state prediction was integrated within the Cockpit Motion Facility at LaRC as described in section I. The study resulted in the successful development of methods and scenario details to inform the commercial aviation community of the techniques employed by the NASA SE211 team for creating realistic flight simulation scenarios that reliably induce pilot states such as channelized attention, confirmation bias, and startle/surprise (the AHPLS).

AC 120-35D - Flightcrew Member Line Operational Simulations [8] presents guidelines for the design and implementation of LOS which includes LOFT, SPOT and LOE. An objective of CAST SE211 was to develop, script, test, and evaluate LOFT and SPOT scenarios. Therefore, the scenarios used in this study were intended to support (a) the development of methods, guidelines, and best practices for creating training scenarios that induce the attention-related human performance limitations as observed in incidences and accidents associated with the loss of airplane state awareness, (b) the assessment of whether these methods can be used to support training in the airline operational environment, and (c) the provision of a means to assess the concurrence of CSM classifier attentional state prediction.

The data were analyzed to assess the ability and efficacy of different flight scenarios to trigger AHPLS as these states might reduce their aircraft state awareness. Additionally, the scenarios were evaluated for potential use in Line-Operational Simulation training programs, including LOFT, SPOT, and LOE.

The 90 minute duration LOFT scenario (Figure 5) was a gate-to-gate scenario with the following event sets: (a) four instrument anomalies, (b) a wake vortex encounter, (c) hydraulics and anti-skid failures, (d) airborne traffic conflicts, (e) a dispatch request to return to airport, (f) a runway change, (g) a missed approach, (h) a runway incursion, and (i) an air traffic control (ATC) taxi clearance error (via live, scripted ATC). Instrument anomalies were inserted throughout to check for detection. For further details, see (Stephens et al., 2017).

AHPLS induced by flight scenario events during SHARP 1 included:

- Channelized Attention
  - LOFT Hydraulic System Pressure alert
  - LOFT Flap Asymmetry
  - SPOT Fuel leak
  - SPOT Display failure
- Startle / Surprise
  - LOFT Wake turbulence
  - LOFT Runway incursion
  - SPOT Fuel imbalance + turbulence
  - SPOT Altimeter failure/throttle anomaly
- Diverted attention
  - Unexpected air traffic control calls
  - Noticed messages, warnings and alerts

[8] https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1027170

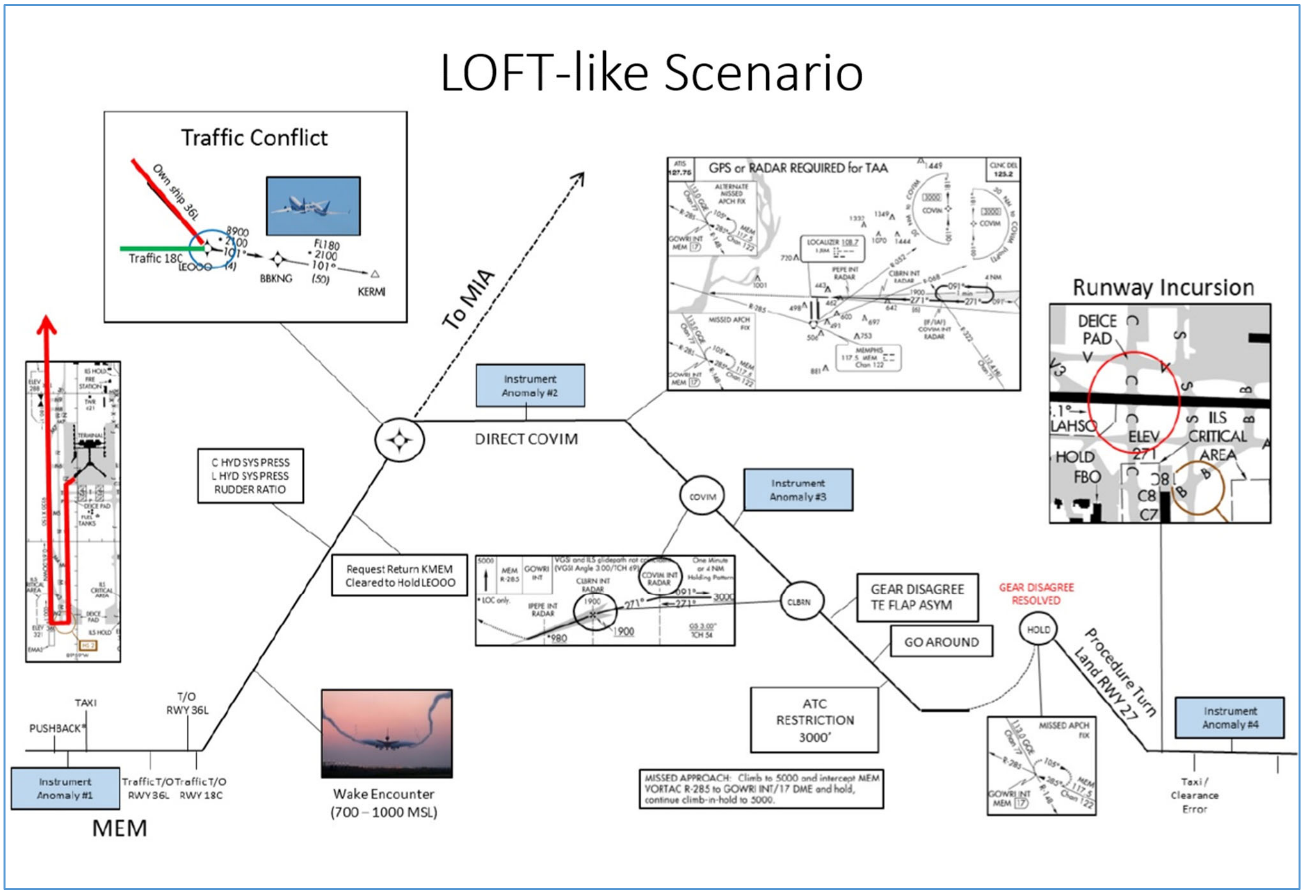


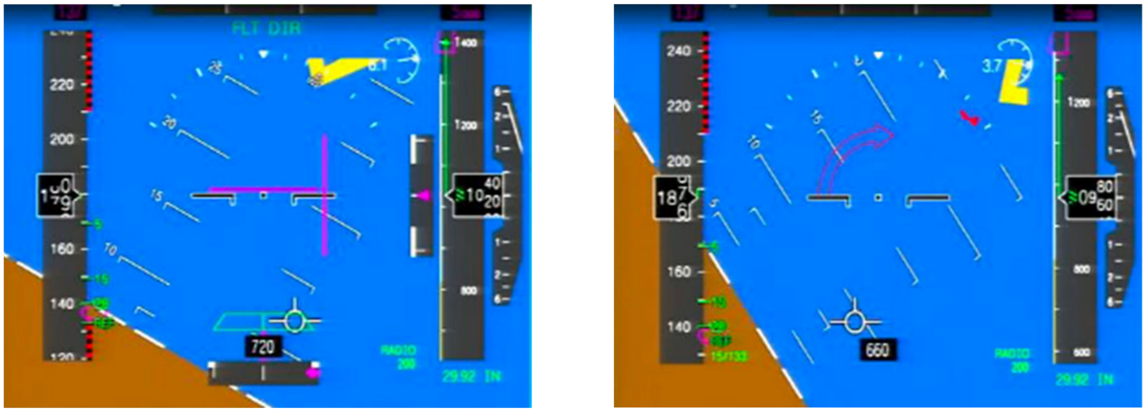
Figure 5. An overview of the LOFT scenario.

The wake encounter occurred after takeoff in the LOFT scenarios at 700 ft above ground level (Figure 6). This was the first "Startle / Surprise" event encountered by the pilot participants. The response may be a reflexive startle, or emotional based on a violation of expectations, or both. There are popular examples of startle (Parker et al., 2008) and surprise in popular television episodes and films which depict recognizable behavioral indicators of such states. Parameters of the event were determined using reported behavioral response pilot measures measured during a wake encounter (Ahmad et al., 2014).

Channelized-attention-inducing event number one began with "R HYD SYS" and "ANTISKID" messages, with proximate traffic indicated (Figure 7). The "TE FLAP ASYM" and a go-around (Figure 8) with ATC-amended level-off represent channelized attention event number two (Figure 9).

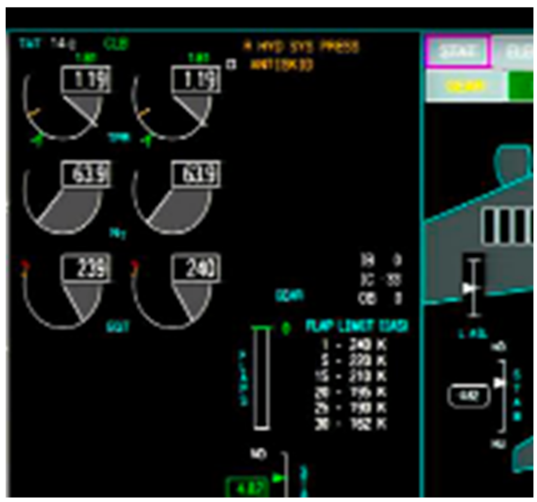
Four instrument anomalies (Figure 10) were also presented during the LOFT scenario. This tested their use as a means of evaluating anomaly detection via instrument scan.

The runway incursion event (Figure 11) was designed to be a "pilot deviation" (cross hold line on active runway of other traffic) Category B event requiring the flight crews to make corrective/evasive action to avoid a collision, but was not expected to result in a collision unless the flight crew exhibited poor attention management (FAA 2015a; Jones & Prinzel, 2011).

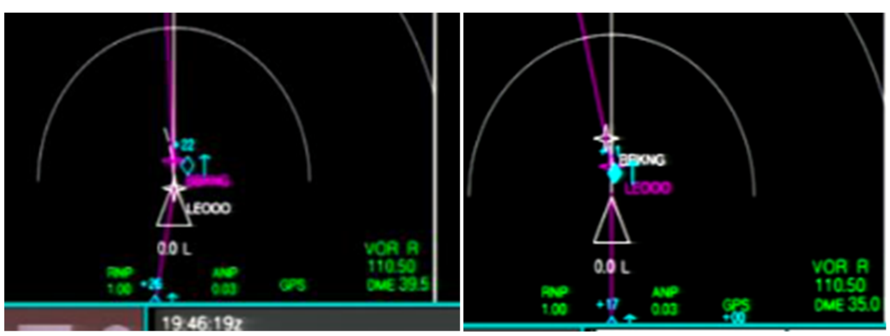


Normal Recovery Unusual Attitude Recovery

Figure 6. Displays during recovery from a wake encounter used as startle-surprise-inducing event number one, as simulated in the Cockpit Motion Facility at LaRC.



R HYD SYS and ANTISKID Events



‘Other’ Traffic Indication Proximate Traffic Indication

Figure 7. Displays during channelized-attention-inducing event number one (upper) simulated in the Cockpit Motion Facility at LaRC. This was followed by display (lower) of a potential traffic conflict.

Figure 8. Displays during the go-around during approach to Runway 36L simulated in the Cockpit Motion Facility at LaRC.

TE FLAP ASYM message

Flaps indicator

Figure 9. Displays during channelized-attention-inducing event number two simulated in the Cockpit Motion Facility at LaRC.

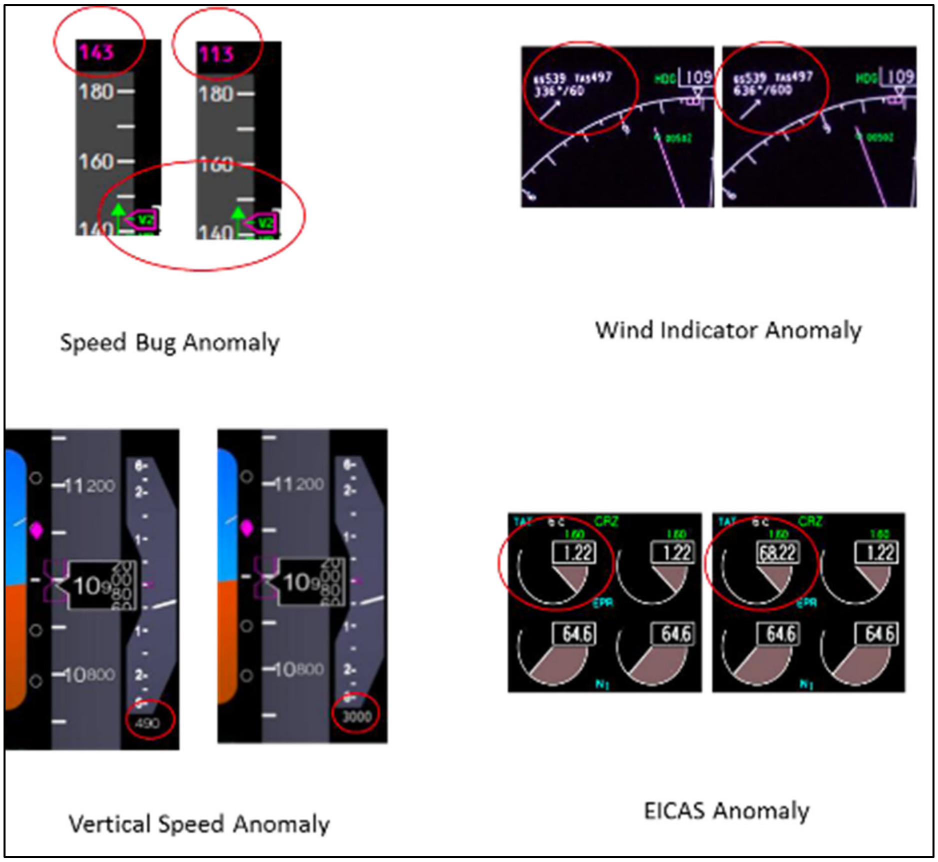


Figure 10. Notional Examples of Instrument Anomalies.

Figure 11. Displays during a runway incursion event used as startle-surprise-inducing event number two, as simulated in the Cockpit Motion Facility at LaRC.

Video recordings of the scenarios were reviewed by flight instructors (this is discussed further in section II B), who remarked on the state induced and recorded the time of the induction and duration. These responses indicated that state induction was achieved by the developed scenarios as intended for each of the states. The LOFT scenario was found to be highly effective for producing startle/surprise responses for the wake encounter and runway incursion events (when the event occurred as implemented without technical issues) and for channelized attention event set number one (hydraulic system and antiskid failures). The second channelized attention event set was marginally effective owing largely to the highly variable nature of scenario segments which, to maintain realism, allowed degrees of freedom for pilot responses. As a consequence, the TE FLAP ASYM and behavioral indicators were not always manifest in the LOFT scenario. A number of potential opportunities to also evaluate CSM data for diverted attention were also identified, including a variety of messages as shown, plus unexpected ATC calls. ATC was provided live by experienced controllers for SHARP 1.

Overall, the flight crews exhibited acceptable threat and error management and NOTECHS (Non-Technical Skills) and Line/LOS behavioral markers were found to be good to acceptable across pilots. Pilot Technical Standards were found to meet FAA published standards and were evaluated against multiple phases of operations during the LOFT scenario. Finally, the ATC taxi clearance error event set demonstrated that approximately half of the flight crews accepted the erroneous taxi-in clearance without cross-checking and verification.

Findings and Recommendations for Training

Prior recommendations around flight path monitoring include appropriate use of briefings and deviations from previously-briefed plans, workload management and task prioritization, and specific interventions for specific threats to effective flight path monitoring (Flight Safety Foundation, 2014). The SE211 work is not intended to replace those excellent recommendations. Rather, we build upon them by presenting methods addressing how, via informed training practice, one might achieve self-management of attention, and thus the ability to maintain awareness and attention priorities and to recall and enact appropriate interventions despite the occurrence of AHPLS. The recommendations in this SE211 report point to methods that should be considered for progressive training approaches. Methods are presented to address identified gaps in current flight path monitoring, such as a lack of feedback regarding attentional lapses and dramatic variances in monitoring skills (Flight Safety Foundation, 2014), failures to check and fully process mode annunciations (Sarter, et al., 2007), and a lack of explicit emphasis on monitoring skills identified as lacking by CAST (CAST 2014b).

A scenario may be designed to avoid "simulator mindset" in which the flight crew may (a) not have the same responses as in a real-world event and (b) exhibit high vigilance due to an expectation of the occurrence of off-nominal events which are otherwise-rare. In these cases, crew may experience different physiological and behavioral responses. Indeed, pilots have been found to have more difficulty managing upset when it is presented unexpectedly, and training should include elements of surprise (Landman et al., 2017).

For an assessment of differences in response times to the same wake encounter event given simulation with and without a motion base, and for further details regarding the events used in the SHARP 2 study, please see NASA/TM–2020-220576, "Flight Simulation Scenarios for Commercial Pilot Training and Crew State Monitoring" by Comstock et al. These results support the use of motion-based simulation for the implementation of these training scenarios for events which require the assessment of response time to punctuated events. The value of costly motion-based simulator time is an important consideration. There are many aspects of training for a particular aircraft type or procedures where simulator motion would not make a difference. However, for scenarios in which detecting the onset of the event is critical, simulator motion makes a difference. These include the practice of unusual attitude or upset recovery scenarios. From a sensory standpoint, human motion perception is good at detecting change but not absolute position or subthreshold changes. As evidence of this, earlier research on eye-tracking and simulator motion (Comstock, 1984) showed that both correct and even reverse-direction pitch motion resulted in faster response times in simulator testing of responses to pitch disturbances (Comstock, 1984).

B. Training-based Mitigations

Technical Progress

The application of psychophysiological methods to the assessment of human operators in aviation contexts and other environments has a long history in what is now the Crew Systems and Aviation Operations Branch at NASA Langley Research Center (Pope & Bowles, 1982). The research approach including biocybernetic technologies to augment the human-machine interaction (Pope, Bogart, & Bartolome, 1995) set the course for numerous research projects and has influenced the current work in Training for Attention Management. The TfAM Team at NASA Langley undertook a research-based approach to address sub-optimal attentional states identified and defined by CAST (CAST, 2014a) which built on the perspective the team had developed previously by expanding the effort into novel concepts and prototypes. The approach taken by the team to develop an example training-based mitigation is described. The previous work leveraged to accomplish this effort is also specified.

Previous work focused on crew state monitoring of mental states identified by pilots in themselves during incidents they experienced (Comstock, Harris, & Pope, 1988; Pope & Bogart, 1992). The perspective of identifying hazardous states shifted to developing a system capable of characterizing an optimal attentional state and providing feedback to the operator to enable optimal engagement and workload (Pope et al. 1995). More than a decade of work at NASA Langley and Old Dominion University focused on developing and testing this capability which established the concept of adaptive automation (Scerbo, 2007; Stephens, Scerbo, & Pope, 2012). The application of adaptive automation was intended for real-time operations such as implementation of a system in the flight deck of an aircraft. Although the initial focus was on operational purposes, our team and other researchers have identified the possibility of such a system being used outside of the cockpit as a ground-based training capability (Hettinger et al., 2003; Pope, Stephens, & Gilleade, 2014).

Conceptualizing a biocybernetic system as a training technology to promote optimized attention was demonstrated in research conducted at NASA Langley (Neilson, Stephens, & Pope, 2012) and expanded upon in publications by our team (Pope & Stephens, 2011, 2012, 2015). The primary mitigation approach developed under the TfAM SE translates attention augmentation methods described in previous team publications to the commercial aviation training context with a focus on training attention management skills to lessen the impact of AHPLS on ASA. The capability developed by our team involves a real-time crew state monitoring technology and psychophysiological sensors for real-time attentional and workload state prediction displayed to the flight instructor/check airman. This prototype version of the attention and workload display based on reduced data from the CSM sensors is shown in Figure 12.

"Project IR" is a program for instructors to view an "instant replay" (IR) of SE211 participant HITL performance, perform post-experimental assessment of state inductions and provide feedback via a "SHARP display" tool (Figure 13). The results can be used to inform machine learning method development and capture instructor intuition as ground-truth to support objective measures of attentional state. SHARP Display and Instant Replay tool system investments and intended payoffs are presented in table 4.

An objective of this work was to create an end-to-end system designed to detect the occurrence of cognitive states such as the AHPLS simultaneously in real time using psychophysiological data recorded via multiple sensing modalities. Data from the individual physiological modalities is fused to take advantage of any synergistic information they provide. An underlying premise of this approach is that no single physiological marker, behavior, or subjective report is sufficient when used alone to assess mental state. A single indicator may result in a false positive or negative assessment, whereas multiple indicators leverage the strength of machine learning techniques, enable the use of convergent validity and mutual information, and help account for individual differences.

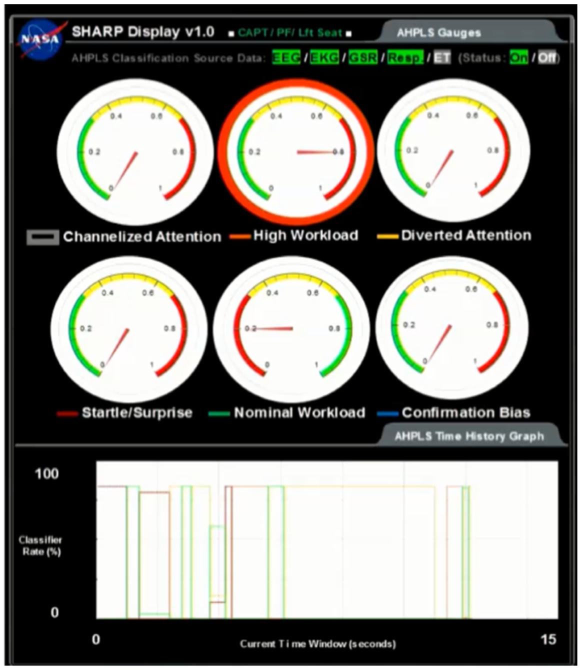


Figure 12. A screen capture of the operational display for the System for Human Attention Restoration using Psychophysiology (SHARP) which depicts the moment-to-moment attentional state of the human operator whose physiological measures are being recorded, processed, and analyzed.

Figure 13. The "Instant Replay" display incorporates a mosaic of displays including the SHARP Display and video feeds of cameras and graphics of the flight simulation the flight crew experiences during a LOFT-like scenario.

Table 4: Investments and payoffs associated with a SHARP AHPLS detection system.

Investments in training operations	Training payoffs
Trainee application of watch, headband, eyeglasses and stick/clip-on sensors	Production of documentable objective measures vs. unreliable self-reporting and inconsistent subjective assessment
Trainee time invested in performing benchmark task training	Increased accuracy of prediction due to personalized baseline
Trainer presentation of flight scenarios and AHPL events	Induction of states not currently trained
Trainer observation of state prediction visualization displayed via tablet	Increased efficiency and improved quality of debrief due to objective measures
Trainer and Trainee review state induction and recovery (or lack of recovery) from the state	Improved Trainee attention management

Further objectives are to introduce: (a) the novel use of tasks specifically designed to induce diverse cognitive states that are qualitatively different from one another as a basis for developing cognitive state classification models, (b) a means of refining classification models based upon subject matter expert judgment, and (c) a way to present cognitive state detection derived from such models to training personnel or the human operators themselves in a readily comprehensible fashion. This could be in ground-based simulated flight, or in actual flight for training or operational purposes (LAR-18996-1, US-Patent 10,192,173).

The TfAM team reviewed the video recordings of the LOFT-like flight simulation sessions from the SHARP 1 Study and documented occurrences of the AHPLS experienced by the participant pilots in the experiment. The assessment was performed by the research team as a group with consideration for behavioral and performance indicators of attention and workload decrement. This consideration also included each pilot in the two-person crew separately and together. The labels were recorded on a standardized record form and digitized for use in machine learning-based analyses for AHPLS classification.

SMEs at Boeing Flight Services reviewed under Space Act Agreement the video recordings of the LOFT-like flight simulation sessions from the SHARP 1 Study and documented occurrences of each of the AHPLS of interest as having been experienced by the participant pilots in the experiment. The assessment was performed by two SMEs at Boeing while considering behavioral and performance indicators of attention and workload decrement. This also included each pilot in the two-person crew separately and together. The labels were recorded on a standardized record form and digitized for use in machine learning-based analyses for AHPLS classification.

A public and international Kaggle competition[9] was held using the SHARP 1 dataset. This challenged data scientists to build models to detect ineffective attentional state events from aircrew's physiological data. Under contract with Booz Allen Hamilton, the winning models were assessed and further improved. The resulting best real-time-compatible methods used autoencoding to deal with biological noise, and two 6-layer dense neural networks. In this way non-nominal attentional states (AHPLS) are identified at rates greater than 0.8, enabling the identification of times when the collection of additional information, attention aids or cautionary measures are warranted. These methods are described further by Terwilliger et al. (2020).

In an effort named "SHARP 3" these concepts were disseminated to the training community via a "Road Show" of external technology demonstrations throughout fiscal 2018. The goal was to gather and assess instructor feedback on training feasibility of the system and scenarios and regarding candidate mitigation methods to inform further development. The feedback is presented and described in section III below.

---

[9] https://www.kaggle.com/c/reducing-commercial-aviation-fatalities/overview

Findings and Recommendations for Training

Training Mitigation Conceptualization and Implementation activities were performed to varying levels for various ideas as presented in this report. However, all ideas were presented - regardless of conceptualization and implementation maturity - to instructor SMEs for their feedback throughout the Research Safety Enhancement activity. Recommendations for training mitigations are presented here based on the research team's experience and the feedback received.

Mitigation categories are presented in the detailed implementation plan (DIP) (CAST, 2014b) for the SE211 work. Briefly, these are: a) improved instrument scanning, b) detection in others, c) detection in the self, and d) standard operating procedures following and response to alerts. Recommendations are discussed in turn in this section below.

Further, a NASA/CAST Attention Management Workshop attended by various prominent aviation safety and training stakeholders and subject matter experts was hosted in June 2015 by NASA Langley, the Crew Systems and Aviation Operations Branch, and the SE211 team. The 23 non-NASA attendees were international and represented industry, academia, unions, training and regulatory agencies. This workshop also informed the development of scenarios presented in section II A. The anonymous, unfiltered discussion comments will be included as applicable to each mitigation category below.

Overall, there are many levels of mitigation that may be adopted by training organizations at their choosing within existing regulatory flexibility (please see Background discussed on page 5 above). At a minimum, an awareness campaign (e.g., see the FAA's awareness campaign on fatigue) around attention being a trainable skill could be implemented. Posters and instructional videos (similar to those used to raise awareness of spatial disorientation effects and optical illusions) could be used. The description of a "mental illusion" as an analogy to an "optical illusion" may be helpful. Additionally, training organizations may wish to adopt scenario events as-presented or design their own using the ideas presented herein (section II A). Finally, training organizations may wish to adopt a full SHARP system (sections I and II B). This would require additional research and development which builds upon the novel technical work performed and disseminated by the SE211 team (section III). Further discussion of recommendations and next steps are provided at the conclusion of this report (section IV).

a) Improved instrument-scanning behaviors in both nominal and off-nominal conditions, with emphasis on scan patterns for glass cockpits

Eye-tracking is a mature sensing modality capable of providing evidence of visual attention. However, there exists no standard scan for glass cockpits. If resources do not allow the development of a full SHARP system, eye tracking gaze location data might be collected to inform trainers regarding whether an item on the flight deck important to situational awareness was addressed visually by a trainee. This alone however does not inform regarding cognitive attention. However, it may be a valuable start. Eye tracking solutions are currently available from multiple vendors. Also, relevant research for example addresses real-time in-cab feedback. Fitzharris et al. found that real-time driver monitoring feedback, including auditory and haptic warnings, is effective in reducing fatigue events during commercial trucking (Fitzharris et al., 2017).

Kuo, et al. demonstrated the utility of continuous ocular metrics for assessing workload and fatigue in operational environments. Gaze predicted max and mean workload, and pupil diameter was related to operator fatigue. No active manipulation of operator state through task or environmental changes were implemented. They state: "The negative relationship between gaze spread and workload is consistent with the phenomena of visual tunneling, whereby visual scanning behavior decreases as tasks become cognitively demanding." (Kuo et al., 2017)

While the mechanism for the effect is not fully explained, Aidman et al. found that real-time feedback resulted in improved alertness and driving performance (Aidman et al., 2015).

Regardless, if eye-tracking is headworn, a practical but important recommendation is to be sure to include the ability to accommodate reading glasses and the need to dynamically doff and don them.

In addition to those few examples of relevant research ongoing, these topics are analogous to those in the automotive field, which also addresses issues associated with cameras and privacy. The existence of

Cadillac's Supercruise and other similar systems are offered as evidence that these issues can be overcome.[10]

Feedback gathered during the June 2015 workshop on scan-related mitigations include the following unfiltered, anonymous discussion comments. (Acronyms are provided in Appendix A.)

- Eye tracking = crew state monitoring (CSM), three years to recommend this to CAST, looking for ways to train this topic
- Eye-trackers do pupil dilation for workload, Bady in 1982 discovered that pupil dilation = indicator of workload
- Body temp up, GSR up, EKG and EEG signs
- Don't have instruments to scan for that, when mode is changed they don't know what's going on/ not focusing on mode change
- Go from cognitive thinking to associative thinking, not same connection with an FMA, FMA is an awareness item, there's nothing that draws me to the FMA, it's doing what it's supposed to do, if you want to get guys to be aware of FMA, it must be a part of the crosscheck
- CAT Approach, people missing out on critical things, example of approach we might take, get into sim with instructor, often in sim ask people to enunciate to make them pay attention, COULD take an occulometer that is unobtrusive and the instructor can get a display and look at where they're looking in real time and can report on this in real time back to the pilot/ Instructor can observe scan patterns and identify pilots who need more training
- Some instructor's make pilots verbally announce FMA change, had incident in Columbia where crew with stepdown altitudes (ASI) /20 yrs ago with piece of cardboard? / He was good at saying you aren't looking at the PFD
- What if we could give a tool where every instructor could do the same thing?
- Showing someone in a debrief after flight training would be a good reinforcement
- Just because someone's looking at something it doesn't mean they're actually seeing it! So let's say we take the eye-tracking approach, how do we do the other part? Not only can we make sure they're looking at the right things, how do we know if they're understanding what they're looking at?
- How do you understand what they're thinking? Really we want more.
- sometimes we do STOP training and they go into approach mode and ……. 10,000 hrs and make same mistake
- What's proof of understanding? (Humans are very adaptable)
- There are other tests that might be applicable to it
- You kind of have to bite this elephant one piece at a time
- What we're doing now to get FMA awareness is getting guys to yell out FMA changes, guys verbalizing changes, worried about reaction to seeing it
- Recommended to learn monitoring skills in the classroom and through practice, do what you're already doing, just better, NASA's like, that's not enough, use rxn time accuracy and error rates, many ex. Of where you could monitor in the sim to measure this, eye-tracking is a suggestion. How can we use that tool? Where were they looking? Was it effective?
- Could combine both the tool of eye-tracking and seeing if they act on it
- What is the optimal scan pattern? Then challenge of knowing whether they're understanding it.
- Also with scan pattern, pay attention to scan rate, scan too rapidly, not understanding what you're seeing

[10] https://media.gm.com/media/us/en/cadillac/news.detail.html/content/Pages/news/us/en/2017/apr/0410-supercruise.html

b) Recognition of channelized or diverted attention in one flight crew member by the other flight crew member, and appropriate methods of intervention and correction

For detection in others, frequency of communication and missed callouts may be useful. When and how to interrupt a fellow crew member is also relevant. Feedback gathered during the June 2015 workshop related to intra and interpersonal recognition mitigations include the following unfiltered, anonymous discussion comments. (Acronyms are provided in Appendix A.)

- When designing scenarios, consider:
    - getting the 'crew' into a state or states, not just each pilot flying or monitoring
    - A recommended goal 'crew' state is inducing the PF into a channelized state (trying to recover from or avoid a LOC-I event) while the PM is distracted (and can not help the PF because the PM may not recognize the event or is at least behind in awareness regarding the events).
- This may be accomplished as follows:
    - divert the PM with a complicated task, such as an inability to enter RVSM airspace, a fuel balance or transfer issue, a medical emergency, necessitating attention to activities such as fuel or reroute / replanning, ATC, dispatch contact.
    - 1 - fully engage the mental resources of the PF with an off-nominal flying task, to induce loss of ASA as evidenced by low energy state or unusual attitude
    - for realism connect with dispatch issues (altimeter failure, auto-pressurization failure, APU MEL'd out)
    - 2 - engage the PF's mental resources with, for example, a pitot-static issue leading to unreliable airspeed, AOA freeze-up, or a subtle sensor failure leading to a sub-threshold roll.

c) Self-diagnosis methods for flight crew members to recognize and recover from channelized attention, confirmation bias, startle/surprise, and diverted attention

SHARP system development is most related to this mitigation type, and has been described in sections I and II B above. Here, the SE211 team proposes that exposure to each attentional state via a state-inducing scenario event during ground-based training simulation, importantly followed by instructor-led debrief informed by quantitative data regarding the trainee’s cognitive state at times crucial to airplane state awareness, could improve a pilot’s ability to manage their own attention during operational flight.

Discussions at the Langley 2015 workshop related to crew self-diagnosis mitigations are summarized here. The novel SHARP system and video replay concepts (LAR-18996-1, US-Patent 10,192,173) were illustrated to the workshop attendees to prompt feedback and consideration.

Implementation of the SHARP Display would take advantage of the detection methods of SE211 Output 1 to provide objective feedback to the instructor (by presenting the state prediction outputs) to further improve the self-recognition of performance-limiting attentional states by the pilot. Practice at self-recognition would be provided via the instructor in-training or post-training debrief. Essentially, this represents training in self-diagnosis for pilots themselves to recognize and recover from non-optimal states such as channelized attention, confirmation bias, startle/surprise, and diverted attention. The analogy of feeling/learning one’s own hypoxia symptoms has been a useful discussion and dissemination tool for the research team. This is differentiated from self-detection where state prediction outputs are presented directly to the pilot crew themselves, such as autogenic or closed-loop feedback training (Cowings and Toscano, 2000; Stephens et al., 2018). Additionally, research supports the use of mindfulness training to improve attentional control under high workload (Meland et al., 2015).

Subject matter experts supported the effectiveness and usability of a video replay tool to:

- allow trainers to observe the experimental scenarios with all of the pertinent information about the flight situation
- gather ground truth for refining state classification models
- learn from trainer observations of crew behavioral cues

d) Reinforcement of proper airline procedures regarding recognition of and response to flightdeck alerts.

Some ideas addressing SOPs and alert response have been provided by Barshi et al. For example, the idea of "pause points" could support training pilots to manage their own attention by their setting aside time to give "full attention" to critical checklists and "Do" lists. Also, not constantly monitoring unchanging displays, and performing tasks in serial (vs. in parallel or multitasking) are relevant to training for attention management (Barshi et al., 2016).

A relevant FAA advisory circular, AC 120-71B - Standard Operating Procedures and Pilot Monitoring Duties for Flight Deck Crewmembers, addresses safe operations as founded on comprehensive standard operating procedures (SOPs) and provides guidance on integrating pilot monitoring duties into SOPs.[11]

Feedback gathered at the Langley 2015 workshop related to procedural mitigations include the following unfiltered, anonymous discussion comments.

- How might we do something similar with training where pilots were not compliant with procedure? Why were pilots getting clear warning systems and not responding correctly or at all? Because high workload and high stress, either reverted to incorrect procedure or didn't do anything. Is there anything we can do more of to prevent this from happening?
- So with cognitive saturation, have 3 red lights, know that all are bad but one's worse, how do you prioritize your failures? It's something that needs practice.
- There's a level of denial in there, "This has not happened to me in my whole career, there's something wrong with the system."
- First thought is often, "That can't be," then, "I don't care."
- Noted a difficult behavior to solve via training in which Pilots with high time and experience perform operations without following SOPs (i.e., skipping steps) which causes breakdown in CRM.
- Pilot callouts as a method to maintain SA

[11] https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1030486

III. Feedback from SMEs

Feedback gathered by the SE211 team throughout the multi-year effort, in addition to that provided above in this report, includes the following unfiltered, anonymous discussion comments relevant to training for attention management in general, and detection methods and scenario development. Notes in square [brackets] were added by the authors for clarity and understanding. This feedback was received through the 2015 workshop (introduced in section II B), pilot study participant feedback (both ad-hoc and via a LOFT questionnaire), and during a "road show" style series of external Technology Demonstrations held throughout government fiscal year 2018.

The Technology Demonstration locations were selected according to potential to diversify the adaption of developed state induction scenarios by airline sub-populations, demonstration opportunity, and accessibility of instructor SMEs. An equipment kit was packaged and transported to each location ahead of the visit to deploy the SHARP system. In some cases, developmental data were successfully acquired, even during simulated flight.

<u>Feedback and Discussion related generally to training for attention management</u>

Attention management may best be trained as an overarching skill or technique, not only tied into specific flight scenarios. (Acronyms are provided in Appendix A.)

- Including examples from accidents and incidents into the pre-briefing as it relates to specific maneuvers
- Training regarding how attention played into real accidents and incidents might be a treatment/mitigation strategy itself (or a part of one)
- The use of real events to motivate and show the value of training was noted. Meta-awareness and skills vs. event-driven training were discussed in the context of the Air France accident.
- Would it be effective to include information about human performance and limitations in the pre-brief, to introduce the concepts? HPL and CRM are similar in that they focus on cognitive skills and perhaps a similar approach could be used.
- A trainer indicated it would be helpful to have a CBT to introduce these concepts, perhaps because, as it was noted, CRM is evaluated in training, not necessarily taught.
- Whether training for attention management fell under airmanship vs. problem-solving skills vs. professionalism vs. CRM vs. event-driven training was discussed.
- The discussion and experiences demonstrated the import of attention management training (including detection of, inducement of, and mitigation of AHPLS) that is integrated into existing training requirements (e.g., a discussion focused on how required "extended envelope training" and "manual handling" and specific "pilot monitoring" training that must be part of all U.S. training by 2018 could be utilized to allow for attention management training.
- Need for quantitative and more diagnostically sensitive methods for detecting and inducing those states (i.e., for the instructor to have confidence that the flight crews witness similar responses as that expected on the line)
- Crew state monitoring (CSM) tools and techniques and methodologies that may provide a more objective means in opposition to current practice that relies on instructor expertise ("intuition") and highly subjective and qualitative correlates (e.g., pilot exclamations of surprise) or, most often, is assumed and verified during the post-session facilitated debrief (at which time the opportunity, if not induced, is no longer available).
- The training of general skills vs. knowledge of flows and procedures was apparent. "Make yourself organized, then it's easier to remember." [the word "organized" was recurring] "Stop studying one hour before bedtime." "Ask the captain to repeat information as needed because they will assume you know what they briefed." "Fly mentally in your head." These are examples of overarching skills being trained.

Feedback and Discussion related to Output 1: Detection Methods

Workshop attendees on Crew State Monitoring

- I'm convinced, that within the aviation community, we will see changes in how training is being conducted thanks to all your collective effort.
- One factor that greatly concerns me with distraction and channelization is that of time dilation. …it is all too easy to become so immersed in a task that significant amounts of time pass without the pilot realizing that it may have been several minutes since they last 'stepped back' and rebuilt their mental picture of the situation.
- I believe that your suggestion of training pilots on what their internal state feels like in such a channelized or distracted condition will add a critical component to the pilots' and trainers' toolkits.
- [Given that the workshop leads and experts would be available to travel to and work "on-site" with an airline in support of producing the practical research and objectives.] This may well be an excellent opportunity for [our airline] to leverage this expertise in supporting our training content development team in these targeted training enhancements."
- Current training is marginally acceptable, but "attention management" and "effective pilot monitoring" require enhanced and more innovative training approaches.
- For Line-Oriented Flight Training (LOFT) and Special Purpose Operational Training (SPOT) scenarios, the most critical need was enhancement of line operational realism.
- Current Line Operational Simulation (LOS) scenarios are often designed on the basis of an aviation accident or incident data, without specific focus on induction of AHPL states, not effectively addressing the issues of channelized attention, diverted attention, startle/surprise, and confirmation bias.
- Crew State Monitoring technologies could be innovative but require research on efficacy, easier implementation, and greater acceptance.
- Eye tracking technologies were highly rated and may be enhanced with psychophysiological measures.
- Current training is marginally acceptable, but that "attention management" and "effective pilot monitoring" requires enhanced and more innovative training approaches
- Crew State Monitoring technologies were considered potential innovative approaches to training for attention management, but the technologies require maturation, more research on efficacy, easier implementation, and greater acceptance
- Eye tracking technologies were highly rated as potential CSM technology, and may be enhanced with other psychophysiological measures, as adjuncts to behavioral markers currently in use in line-oriented training
- Need someone to tell you you're channelized at first so you can recognize it, it's difficult to cue yourself because you are already saturated
- Self detection is important in single pilot missions
- Detection of an alpha wave state could show inattention and help detect channelized attention
- Market to trainers not pilots
- Crew alert monitor, if no one touches anything for a while the plane alerts the pilot, protects from becoming lost in a discussion or activity
- Behavioral assessment is most common way trainers currently check during training, these are all very subjective especially intuition, how to you qualify that?
- These measures have a lot of noise, have to use multiple measures and context (environment, expected response to tailor to specific situation)
- Other objective measurements include self-debriefing, the pilot can say when they felt saturated or channelized and the trainer can confirm
- Feedback depends on whether it is LOFT or LOE
- Self-assessment is very important

- Both active and passive destruction of CRM (active would be a pilot focused on telling stories rather than flying)
- Ethiopian 409 off of Africa, mystery, look at video to look at where CRM destruction happened
- Summary data more valuable for CRM, maybe real time would be too distracting?
- Addressing non-optimal behavior is best done through verbal debrief, mostly after the sim
- Trainers check situational awareness with behavioral assessment most commonly currently
- Objective measurement of SA should be real time display, followed by summary
- Interrupting sim session would be unhelpful so post-talk would be better, don't want to stop what is happening, but rather debrief later (but not lecture)
    - get away from systems training and move toward psychological training
    - SA best addressed through verbal debrief
    - What do you envision would be immediate feedback given by metrics? [onset of AHPL state]
- Don't know of any regulatory barriers
- Short burst of alpha before a successful put, also alpha when your eyes are closed
- Tradeoff between group and participant specific, specific is less cost effective but usually gives better results
- Measuring a mixture of alpha, beta, and theta, just EEG, for others things, more than just EEG would be monitored
- How to determine when it is good and when is it bad to be absorbed in what you are doing, what are the signs that you are doing what you should or shouldn't be doing
- Teachers and parents report behavior of ADHD kids, video games technology was used, neuro-feedback in the game was not much different than without the game
- Can you actually measure startle/surprise, confirmation bias, channelized attention?
    - show directionality, instead of raw numbers that don't mean anything to trainers
    - usually measure human performance based on how airplane acted - it's great that we are shifting towards looking at what's going on in the human in order to mitigate these problems
    - Eye tracking = crew state monitoring (CSM), three years to recommend this to CAST, looking for ways to train this topic
    - Eye-trackers do pupil dilation for workload, Bady in 1982 discovered that pupil dilation = indicator of workload
    - Body temperature up, GSR up, EKG and EEG signs
    - appreciate making it easy for the rest of the world to understand (outside of the research world)
    - it will work as long as it is valid and reliable
    - figure out root cause of problems by looking at physio measures
- When trainers just "don't have a good gut feeling" about a student, it's good to have the physio data to look back at and validate this intuition
- How do we train pilots to have cognitive and physiological control?
- Need trainers to be able to recognize inattention, a simulator may not be the best place
- Videos can be used to train, pilots would watch what was happening to other people, most helpful if there is an instructor walking you through it (this can be a problem because people will think "oh this will never happen to me" and so they don't worry about it, you need to put people into a situation where it happens to them to successfully train against it)
- -progress through stages of continually increasing things you need to monitor, helps individual train themselves to be able to pay attention to more and more at a time (simulator would be ATC traffic)
- Show both instructors and trainees an inattentional blindness video to show they are susceptible (to look but not see)

Study participants on wearing psychophysiological sensor gear

- Didn't notice it much after a while
- Considering all of the things it was measuring, it was pretty comfortable, even toward the end of the day
- Virtually invisible. After a few minutes, it's hard to even know it's there
- Awkward and uncomfortable, but manageable
- A few comments concerning the fNIRS being uncomfortable
- Gear didn't necessarily hinder flight performance, but still remained "a little distracting" for some
- Would have been better to have finger sensors on stick hand
- Restricted my motion on several occasions as the wires would seem to be binding against the harness or the chair or something else in the cockpit

<u>Feedback and Discussion related to Output 2: Scenarios and Mitigations</u>

Workshop attendees on scenarios

- Give pilots unusual circumstances in safe environment to force them to think
- Suggest teaching stress training with true unexpected events in a no jeopardy event and teach them that it is expected to perform poorly.
- LOFT and SPOT scenarios most critical need was enhancement of event sets that have greater line operational realism (e.g., should emphasize the concurrent task demands, distractions, interruptions, and delays often encountered by flight crews)
- Current LOS scenarios are often designed on the basis of an aviation accident or incident data without specific focus to create the necessary conditions to induce AHPL states.
- Scenario and training often do not effectively address the issues of channelized attention, diverted attention, startle/surprise, and confirmation bias (AHPL focus of SE 211)

Study participants responses and expert opinions on the LOFT scenario

A take-home final questionnaire (available as an appendix in NASA/TM–2020-220576 by Comstock et al., 2020) was given to all SHARP 1 study pilot participants with instructions to complete the questionnaire package and return via U.S. Mail. There was a 70% return rate achieved. Each question asked is provided below with a corresponding data graph, where appropriate. Some questions asked pilots to provide a comment or further explanation or details, and those responses are provided (if a response was given). Pilot responses to the SHARP1 final questionnaire provided a wealth of data in terms of current LOFT scenario implementation at airline training centers and substantial information for work on Output 2 of SE211. Thirty-four questions were asked that revealed significant and valuable information on how to enhance LOFT scenario and implementations, and on potential avenues to explore for further scenario development specific to the construction of training for attention management scenarios which address AHPLS.

The SHARP data results for the LOFT scenario evince that the scenario was rated to be "excellent" / "very good" (82%) with 68% of pilots responding that NASA LOFT scenario was of higher quality than airline LOFT scenarios they had experienced. "Quality" was undefined. One pilot rated the scenario as "fair" but commented that the reason was due to issues with the simulator (i.e., non-standard flight displays and controls). The NASA LOFT scenario was also judged to be "excellent" to "very good" in comparison of realism to actual commercial flight operations and hazards encountered on the line. The LOFT scenario was found to be highly effective to producing startle/surprise responses and channelized attention (please refer to section II A for more details).

Additional ad-hoc feedback includes the following:

- Artificialities built into a training scenario (to achieve a certain outcome) versus flying like you train has the ability to produce negative training or promote gaming the system knowing subconsciously that 'it's only a training exercise.'
- Except for the very few snap decisions with associated maneuvers we have to execute (V1 cut, Rejected Take-off, Rejected landing, wind shear), we train to deal with things in a controlled, methodological, and procedural manner. From that standpoint knowing how the body reacts during episodes of being startled and distracted came into focus.
- Scenarios starting in high altitude operations have a lot of potential because that is where breakdowns in performance originate, and FAA Advanced Qualification Program (AQP) could allow for creating scenarios in that phase of flight.

Responses on low-time and cultural issues

- Multiple professionals indicated that young pilots with low time and limited experience lead to majority of errors internationally.
- An approach in which young/new pilots call out mode annunciation changes to maintain common situation awareness with Capt. However, it can be difficult to balance communication against cultural respect [for crews].
- The contrast between a Part 121 and Part 129 training flight crew and the noted differences needed in approach to them (e.g., non-U.S. crews may not be as communicative and have more difficulty in understand the Western-built aircraft and use of English as a communication medium can impede training. Part 129 pilots tend to need additional training in pitch/power relationships and that typical pilot has only 200 hours in contrast to U.S. that require 1500 hours).
- A challenge with foreign students is difficulties connecting with students, getting them to talk and getting feedback. [Presumably some objective feedback might be helpful.] The ability to customize training per student was again brought up.
- Due to the large general aviation component of US aviation, international training relies much more on theoretical knowledge.
- There was a major difference across trainer mannerisms. Trainer nationality is purposely matched to the trainees as possible with local trainers for improved communication.

Responses on training requirements and resource limitations

- Currently, moment-to-moment actions by a pilot involve OODA loop decision cycles of observe, orient, decide, and act [for aviate, navigate, communicate, and system monitoring]
- Standard Recurrent Training follows a specific and approved formula: 1st day CBT, Sim Sessions with known maneuvers, and Practical Test Standards Testing.
- A Trainer indicated that the major constraint was that Training is driven by Practical Test Standards Testing that Pilots see on the final day of the Recurrent Training. [Could infer from this statement that more flexibility and not so much teaching to the test would permit more effective training.]
- Fitting all of the Practical Test Standard requirements into the time allowed limits what and how it can be covered.
- Strictly following regulations perpetuates mediocrity (does not enable innovation in Training).
- Cost defines what is possible for an Airline to accomplish through Training.
- Time is biggest limiting factor, with all of the FAA requirements it is difficult to fit anything else in to the Training experience.
- Airline recurrent training models have become saturated with a matrix of 'must accomplish' items to satisfy FAA training and recency directives. As such, sim instructors do very little teaching and spend most of their time evaluating performance; with or without jeopardy. The pre-sim briefing is typically very thorough however the debrief ends up being an afterthought because the full period has been exhausted with events.

- Line Check Airmen follow a very scripted set of events. There is a strong pressure to pass pilots. If a Pilot is not demonstrating proficiency then it is possible for the Trainer to request that the Airline extend the training time by 1 day or more. AQP was noted as a possible means to allow training customization per pilot.

Responses on high-time issues
- Flying with a crewmember who has downbid from captain back to a first officer position: “When he/she was a captain he/she developed his/her own pace, processes, and priorities. Consequently, relinquishing them or communicating them tactfully in a stressful environment is extremely difficult.”
- Whether any pilot or airline as a whole recognizes and acknowledges it there are issues of quickness, retention, analysis, memory, multi-tasking, eyesight, hearing, etc. for pilots approaching age 65 that even years of experience cannot overcome.
- Periods of self-doubt and extended decision making times … diminish the supply of attention available.
- Performance level comparisons between what that pilot remembers he/she was able to accomplish at age 50 and their most recent sub-par performance… leads to a ‘competency confidence deficit.’ Doubt and second guessing can become critical time eaters when it comes to making decisions.

Responses on instructor assessment
- Training a Pilot on a specific event and then testing them on a similar event
- Repetition and consistency
- Handling it in mixed-up or distracting situations
- The performance of actions at the proper time as opposed to simply following procedures without fully understanding them.
- A trainer stated that you can “see the bell ring in their faces.”
- The whole performance is important and relying on intuition is big
- Procedural vs. experiential training
- Managing tasks vs. airmanship and manual flight skills
- Non-jeopardy sessions are important to permit learning to occur and assessment of that learning. It had been previously mentioned that trainees will talk more in non-jeopardy debriefs.

Feedback and Discussion resulting from Technology Demonstrations

A technical interchange meeting with a commercial aviation training organization included a presentation of results from recent studies and development efforts as well as demonstrations of the SHARP system for attention management skill training. A feedback session resulted in insightful assessment of the prototype system developed by the NASA SE211 team. Suggestions included those for best use of the training-based mitigation technology and limitations of the system given current regulatory constraints in commercial aviation training. Several Flight Instructors indicated that the system would be useful for debrief sessions following LOFT session to promote recognition of AHPLS in the self and the other pilot participating in the simulation session. One Flight Instructor noted that the tool could enable pilots to clarify errors in their perception of the mental experience the flight simulation scenario is intended to induce. Another insight was that the training tool could facilitate pilots in training to buy into the simulated induction of AHPLS for more effective training targeted to those limiting states. Another suggestion for use of the SHARP capability was for assessment of trainee state prior to the simulator session. This application would permit the Flight Instructor to identify distraction in the trainee which might limit effectiveness of training session.

Some limitations noted by the Flight Instructors included the need for the user (i.e., the Flight Instructor) of the information provided by the system to divide their attention among the AHPLS information provided and the typical tasks of behavioral assessment and performance assessment of the trainee. A suggestion to overcome this limitation included the potential for the AHPLS information to be combined with the timeline of the scenario events. Additional feedback on the prototype system include suggestions for improving the display components through simplification of the indicators. Similar feedback included translating the moment-to-moment information into trending or predictive state estimation to provide the Flight Instructors with a look-ahead of the trainee's attentional state in addition to the historical information of the previous state of the trainee.

The NASA SE211 team also provided an exhibit at the 21st World Aviation Training Summit in Orlando, Florida (April 17-19, 2018). This provided substantial exposure to a wide audience of training/simulation industry professionals, and major domestic but also many international and non-major airlines. In general, instructors liked the human fallibility demonstration, evidence-based training and real time aspects to support teaching moments, especially when the trainee disagrees regarding attentional focus. One instructor said he observed a real-world-operations incident similar to one the team presented in the booth via video as an example from the SHARP flight simulation study (where the pilots called "positive rate" while in fact descending). Notably, many comments included how the NASA LaRC SE211 team's work "is the future." Distraction, workload and startle are important issues presented on by many at WATS.

Further NASA SE211 team engagement in-person with industry subject matter experts at major airlines, including presentation of results from recent studies and development efforts, demonstrations of the SHARP system for attention management skill training, and technical interchange. Comments from the SMEs included how the NASA LaRC SE211 team's work addresses future planning efforts undertaken by major airlines to address pilot shortage and change in pilot demographics/experience level. In some cases, feedback provided was overwhelmingly positive and discussion included the potential for incorporating the SHARP technology into specific training for risk management. Unfiltered, anonymous discussion addressed the following.

- Reduce the training footprint based on skill and readiness
- Reduce the failure rates based on readiness to test
- Identify pilot readiness to retire based on performance rather than age?
- Crew Resource Management
- Identify skills needed per student

- Understand the visual information but not the physical feeling
- They need more box time to get the right feel
- Lots of flying experience but not experience with all the new displays and information management
- They need more computer time to understand the appropriate scan pattern and information management systems
- No pilot I know would do this (call "positive rate" while in fact descending) but they did.
- Two aviation training problems:
- Instructors saying a student isn't progressing, isn't retaining feedback.
- This tool can help us with that issue. They send them to cognitive training to help, but what kind of cognitive training are they giving them? Because 10 out of 10 they come back and they're fine. Because it has nothing to do with their needs.
- Just being able to know where they're looking is helpful
- Really good pilots but are older technology pilots, need to learn the new aircraft without having to go through the learning footprint.
- A major jump in technology and they are very concerned about "I don't know if I'll do well in this new type of technology, I don't even know if I would like it" and there's no way to help them ease that transition. Knowing where [s/]he's fixating and looking, it would help anyone in that situation, maybe even accelerate [her/]him. It's good to keep this experience. We're trying to find new, innovative ways to help [her/]him make that jump in technology.
- Simulator is a fairly benign environment. A lot of things are missing. We need to elevate the LOFT scenarios in our actual training.
- For example ATC
    - Runway crossings, incursions, missed calls, similar call signs
    - All the Philly transatlantic start with 7, Another one starts with 9
- CAE [Inc.] has a cognitive test battery they gave to all instructors, then give it to potential instructors to see where the fit is.
- Low pay, retired, bored,
- Our current process is a really great pilot, put them in the sim, good job, but what about all the extra information they're using, and providing feedback, and understanding what is happening/doing wrong.
- Use it for current cadre as an instructor and select similar
- Use it for training the trainer
- Determine if they're able to transition among tasks because it's even harder as an instructor. Capt, controller, and simulator operator… it's a daunting task.
- So much information in the aircraft. So where do they look? How do we help them with that?
- Training module that directs them through an information management approach technique
- Install devices on a pilot and establish metrics, but inappropriately focused attention, I can see how you measure that and that a pilot is having a problem but if a pilot consistently has this issue, how do we correct that?
    - You're focused on that inappropriately, I'm not sure we can break that.
    - Take the pilot who is loaded in heads down, but he's cognitively distracted and flies in the prohibited area on accident.
    - How do I say, you're looking at the wrong thing?
- Hiring/training for pilots but also hiring/training for check airmen
- The potential for expanding the capabilities of the system into boredom in the flightdeck was discussed.

IV. Concluding Discussion, Limitations, and Next steps

Concluding Discussion

The feedback was presented above in unfiltered form to allow the reader to draw their own conclusions. In general, recurrent themes in SME feedback included low-time and cultural issues, high-time issues, training requirements and resource limitations, the expected privacy and comfort considerations, and calls for more customization per pilot to increase the efficiency of training.

Myths commonly associated with applying state detection via crew state monitoring technologies to training include impressions that time, cost, privacy and comfort issues cannot be overcome. However, the team has received multiple instances of feedback where opportunities for improving the efficiency of training were identified. In particular, these are around calls for improving less standardized and rote training with training more customized to a particular pilot trainee's needs. That customization could be based in part on quantitative assessment and improvement of attentional control skills, which could be an over-arching component of proficiency. Privacy issues could be overcome with security and data protocols, and use of the information solely for safety purposes. These issues are already being overcome, for example, in the automotive industry. Finally, objective feedback the team has received indicates only minor issues with wearing sensors as reported in the "Feedback and Discussion related to Output 1: Detection Methods sub-section of section III. The greatest concerns are around the use of wires, which can be overcome with cable management and new wireless sensors as sensor technology evolves (see the CPTS description in section I).

Various avenues for training may be taken up, from awareness campaigns, to scenario adoption, to full SHARP Display or Instant Replay system implementation. At a minimum, an awareness campaign (e.g., see the FAA's awareness campaign on fatigue) around attention being a trainable skill could be implemented. Posters and instructional videos (similar to those used to raise awareness of spatial disorientation effects and optical illusions) could be used. The description of a "mental illusion" as an analogy to an "optical illusion" may be helpful. Additionally, training organizations may wish to adopt scenario events as-presented or design their own using the ideas presented herein (see section II A). Finally, training organizations may wish to adopt a full SHARP system (see sections I and II B). This would require additional research and development which builds upon the novel technical work performed and disseminated by the SE211 team (see section III).

Recommended training might be:

- Skill- or conditioning -based with an in-sim biofeedback tool similar to hypoxia symptom training
- Computer-based training akin to cognitive-behavioral therapy and self-recognition for "anger management"
- A video for awareness with follow-up instructor check
- Made a part of CRM, TEM, ab-initio training, or recurrent LOFT / SPOT

Limitations

NASA has obtained feedback in particular regarding a debrief (and possibly scenario state induction validation) tool that provides quantitative information regarding whether the trainee is experiencing AHPLS. However, additional resources should be applied to further explore additional mitigation methods as described in the "Findings and Recommendations for Training" sub-section of section II B.

Important considerations and concerns include that non-jeopardy training is more likely to be accepted, that the use of sensors needs to not require a technician, and that harmonization, integration and certification issues will be non-trivial.

Additionally, while reviewing video recordings of training sessions, it can be difficult to discern what is and is not effective performance. Typically, video replay requires that the pilot trainee recalls their step-by-step mental process for the trainer to point out any errors or "teachable moments." Also, video recording review is limited because it is not possible to review the entire session due to time limitations - only a limited number of events can be reviewed as a lesson to learn from. The selection of events to view would depend upon the instructor, and on the system for facilitation of efficient event selection.

A further limitation of this work is that Confirmation Bias has not been fully explored. Confirmation Bias is not a "cognitive state" as we've defined here. Rather, it is a decision-making process where one does not seek out information which challenges a current understanding. Thus, it is manifest in decision making behavior and a complex chain of thought patterns, and not detectable in the same manner as the other AHPLS. This is not to say it is unimportant. Indeed, the Flight Safety Foundation provides an article describing it and suggests: "Be aware of confirmation bias and actively search for cues challenging the current understanding of the situation." [12]

Given the inherent nature of research, these recommendations are not presented as design specifications. Substantial risks and uncertainties remain. The implementation of these and other ideas and suggestions for future operational use should be undertaken by those properly experienced and directly knowledgeable in education and flight instruction, and under all applicable regulations. Statements, recommendations and references are not intended as an endorsement of any products or institutions.

Next steps

Recommended next steps include the translation of research-simulated scenario events, the improvement of simulation realism, the consideration of mindfulness or mind wandering studies for training, and further CSM sensor development.

The exploration of which aspects of mindfulness training may be appropriate for airline training syllabi, and mind wandering research studies to address questions of meta-awareness, are recommended. Some existing research in the area of task-unrelated thoughts is relevant (Casner et al., 2013; Casner and Schooler, 2014; Meland et al., 2015).

Considering how to teach pilots to remain calm in the face of stress may be another way to mitigate performance decrement. Scenario-based state induction and state detection as discussed in this report could be a useful method of accomplishing stress inoculation training. The Flight Safety Foundation also provides a Briefing Note on this topic, stating that one of the best ways to reduce stress is to learn to recognize the symptoms,[13] which is relevant to the mitigation of human-performance-limiting states.

Finally, Training Scenario Development is needed to translate LOFT scenarios such as those implemented in SHARP 1 from the research environment to external training facility simulators. Partnerships with simulation developers will be required to ensure that sufficient capabilities exist to completely and realistically simulate the events required to induce the cognitive states of interest. The events should be selected and the scenarios designed to minimize any increase in the existing training time footprint. However, very importantly, all instantiations should be without jeopardy to the pilot trainee to facilitate experience of AHPLS so recovery from them may be learned.

[12] https://www.skybrary.aero/index.php/Fuel_Leak_and_Confirmation_Bias_(OGHFA_SE)

[13] http://www.skybrary.aero/index.php/Stress_and_Stress_Management_(OGHFA_BN)

V. References

Ahmad, N., VanValkenburg, R., Bowles, R., Duparcmeur, F.L., Gloudemans, T., van Lochem, S., and Ras, E. (2014). Evaluation of fast-time wake vortex models using wake encounter flight test data. Atmospheric and Space Environments Conference. Atlanta, GA. June 16-20, 2014.

Aidman E, Chadunow C, Johnson K, Reece J. (2015) Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. Accident Analysis Prevention. 81:8-13. doi: 10.1016/j.aap.2015.03.041

Barshi, I., Mauro, R., Degani, A., Loukopoulou, L. Designing Flightdeck Procedures. NASA/TM-2016-219421, October 2016. http://www.sti.nasa.gov or http://ntrs.nasa.gov/]

Benson, A. J. and Rollin Stott, J. R. (2006). Spatial Disorientation in Flight. In Ernsting'sAviation Medicine, 4E., 433 -458.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., Craven, P. L. (2007) EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. Aviation, Space and Environmental Medicine, Vol. 78, No. 5, Section II, May 2007.

Burgoon, J. K. & Jones, S. B. (1976). Toward a Theory of Personal Space Expectations and Their Violations. Human Communication Research, 2, 131-146.

Burki-Cohen, Judith, 2010, Technical Challenges of Upset Recovery Training: Simulating the Element of Surprise, AIAA Modeling and Simulation Technologies Conference Toronto, Ontario, Canada https://arc.aiaa.org/doi/10.2514/6.2010-8008

Cannon, Walter (1932). Wisdom of the Body. United States: W.W. Norton & Company.

Casner, S. M., Geven R. W., Williams K. T. (2013). The Effectiveness of Airline Pilot Training for Abnormal Events. Human Factors, 55, 3, 477-485.

Casner, S. M., and Schooler, J. W. (2014). Thoughts in flight: Automation use and pilots' task-related and task-unrelated thought. Human Factors 56(2), 433-442.

Civil Aviation Authority Paper, Loss of Control Action Group, 2013, Monitoring Matters, Guidance on the Development of Pilot Monitoring Skills

Commercial Aviation Safety Team, Flight Deck Automation Working Group, Performance-based operations Aviation Rulemaking Committee, Operational Use of Flight Path Management Systems, Final Report of the / September, 2013

Commercial Aviation Safety Team. (2014a). Airplane State Awareness Joint Safety Analysis Team Interim Report. Retrieved from http://www.skybrary.aero/index.php/ Commercial_Aviation_Safety_Team_%28CAST%29_Reports.

Commercial Aviation Safety Team. (2014b). CAST Safety Enhancement SE 211 detailed implementation plan (DIP). Retrieved from http://www.skybrary.aero/bookshelf/books/2541.pdf.

Comstock, J. R., Harris, R. I., & Pope, A. T. (1988). Physiological Assessment of Task Underload (No. N89-19846). Hampton, VA: National Aeronautics and Space Administration.

Comstock, J. R., Jr. Oculometric indices of simulator and aircraft motion. NASA Contractor Report 3801, June 1984 (NASA Grant NGT 47-003-800). (Available at https://ntrs.nasa.gov)

Comstock, J. R., Jr., Prinzel, L. J., Harrivel, A. R., Stephens, C. L., and Kennedy, K. D. (2020). Flight Simulation Scenarios for Commercial Pilot Training and Crew State Monitoring. NASA/TM–2020-220576, Langley Research Center, Hampton, VA.

Cowings, P.S., and Toscano, W.B. (2000). Autogenic-feedback training exercise is superior to promethazine for control of motion sickness symptoms. J Clin Pharmacol. 2000 Oct;40. (10):1154-65.

deBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB. (2015). Closed-loop training of attention with real-time brain imaging. Nature Neuroscience, 18, 3, 470-475. doi:10.1038/nn.3940

Dehais F., Tessier C., Christophe L., Reuzeau F. (2010) The Perseveration Syndrome in the Pilot's Activity: Guidelines and Cognitive Countermeasures. In: Palanque P., Vanderdonckt J., Winckler

M. (eds) Human Error, Safety and Systems Development. Lecture Notes in Computer Science, vol 5962. Springer, Berlin, Heidelberg

Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., Gateau, T.: Automation surprise in aviation: real-time solutions. In: Proceedings of the 33rd Annual ACM conference on Human Factors in Computing Systems, 2525–2534 (2015).

Endsley, M. (1995) Towards a theory of situation awareness in dynamics systems. Human Factors, 37, 32-64.

FAA (2015a), FAA Runway Safety Report 2013-2014. Department of Transportation, Federal Aviation Administration, Office of Runway Safety, Washington DC. May, 2015.

FAA, Defining Startle, Surprise, and Distraction: A State-of-the-Art Technical Review to Support the Development of FAA Technical and Advisory Guidance, and of Line-Oriented Simulation Scenarios for Training, University of Central Florida, FAA CRA 13-G-007, 2015b.

Fairclough, S., and Gilleade, K. (2013) Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. International Journal of Autonomous and Adaptive Communications Systems. 6:1, 63–79.

Federal Aviation Administration Report Number: AV-2016-013, Date Issued: January 7, 2016, Office of Inspector General Audit Report, ENHANCED FAA OVERSIGHT COULD REDUCE HAZARDS ASSOCIATED WITH INCREASED USE OF FLIGHT DECK AUTOMATION

Fitzharris, M., Liu, S., Stephens, A. & Lenné, M. (2017). The relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations, Traffic Injury Prevention, 18:sup1, S71-S78, DOI: 10.1080/15389588.2017.1306855

Flight Safety Foundation, A Practical Guide for Improving Flight Path Monitoring, Final Report of the Active Pilot Monitoring Working Group, November, 2014

Gawron, V. (2004) Psychological factors. In Spatial Disorientation in Aviation. p. 145-195.

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. Journal of Experimental Psychology: General, 129, 308-339.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. Human Factors, 36, 387-405. http://hfs.sagepub.com/content/36/3/387.abstract

Harrivel, A., Heinich, C., Milletich, R., Comstock, J., Stephens, C., Last, M. C., Napoli, N., Abraham, N., Toro, K., Kennedy, K., Pope, A. (2018a). Comparative EEG Sensor Analysis for Attentional State Prediction. Abstract accepted to *The Aerospace Medical Association's 2018 Annual Scientific Meeting*. Dallas, TX.

Harrivel, A. R., Liles, C., Stephens, C., Ellis, K., Prinzel, L., & Pope, A. (2016a). Psychophysiological Sensing and State Classification for Attention Management in Commercial Aviation. Paper presented at the American Institute of Aeronautics and Astronautics, SciTech, San Diego, California.

Harrivel, A. R., Stephens, C. L., Milletich, R. L., Heinich, C. M., Last, M. C., Napoli, N. J., Pope, A. T. (2017). Prediction of Cognitive States during Flight Simulation using Multimodal Psychophysiological Sensing. Paper presented at the American Institute of Aeronautics and Astronautics, SciTech, Grapevine, Texas.

Harrivel, A., Neuroergonomics (2018b) talk citation

Harrivel, A., Weissman, D., Noll, D., Huppert, T., and Peltier, S. (2016b). Dynamic filtering improves attentional state prediction with fNIRS. Biomedical Optics Express. 7(3), 979-1002.

Hart, S. G. & Battiste, V. (1992). Field test of a video game trainer. Proceedings of the human factors society 36th annual meeting-1992, Vol 36, Issue 17, 1291-1295. https://doi.org/10.1518/107118192786749450

Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati(Eds.), Human mental workload (pp. 139-183). Amsterdam: North Holland.

Helmreich, R. L., Merritt, A. C, Wilhelm, J. A. (1999). The Evolution of Crew Resource Management Training in Commercial Aviation. International Journal of Aviation Psychology 9, 1, 19–32.

Hettinger, L., Branco, P., Encarnacao, L. M., & Bonato, P. (2003). Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. Theoretical Issues in Ergonomics Science, 4(1-2), 220-237.

Jones, D., & Prinzel, L. (2011). Simulator evaluation of runway incursion prevention technology for general aviation operations. NASA-TP-2011-217045. Hampton, VA: NASA.

Kuo, J., Lenné, M., Myers, R., Collard-Scruby, A., Jaeger, C., and Birmingham, C. (2017). Real-Time Assessment of Operator State in Air Traffic Controllers Using Ocular Metrics, Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 61(1), 257 - 261. doi: 10.1177/1541931213601547

Lab Streaming Layer, originally written at the Swartz Center for Computational Neuroscience, UCSD. This work was funded by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022 (2012) as well as through NINDS grant 3R01NS047293-06S1. https://github.com/sccn/labstreaminglayer

Landman, A., Groen, E., van Paassen, M., Bronkhorst, A, & Mulder, M. (2017) The Influence of Surprise on Upset Recovery Performance in Airline Pilots, The International Journal of Aerospace Psychology, 27:1-2, 2-14, DOI: 10.1080/10508414.2017.1365610

LAR-18996-1, Method and System for Detecting and Displaying Human Cognitive State using Multimodal Psychophysiological Sensing, US-Patent 10,192,173.

Mackey, J. R., Powis, R. T., Walton, J. C., Hauser, K. M., Hall, C. S., Gotti, D. J., and Harrivel, A. R. (2020). Frequency Domain Functional Near-Infrared Spectrometer (fNIRS) for Crew State Monitoring. NASA/TM–2020-220348, Langley Research Center, Hampton, VA.

Mandrick, K., Chua, Z., Causse, M., Perrey, S., Dehais, F.: Why a comprehensive understanding of mental workload through the measurement of neurovascular coupling is a key issue for neuroergonomics? Front. Hum. Neurosci. 10 (2016).

Meland, A., Ishimatsu, K., Pensgaard, A., Wagstaff, A., Fonne, V., Garde A. & Harris, A. (2015) Impact of Mindfulness Training on Physiological Measures of Stress and Objective Measures of Attention Control in a Military Helicopter Unit, The International Journal of Aviation Psychology, 25:3-4, 191-208, DOI: 10.1080/10508414.2015.1162639

Moehle, R. and Clauss, J., "Wearable Technologies as a Path to Single-Pilot Part 121 Operations," SAE Int. J. Aerosp. 8(1):2015, doi:10.4271/2015-01-2440.

Mrazek Expanding minds: Growth mindsets of self-regulation and the influences on effort and perseverance, Journal of Experimental Social Psychology 79 (2018) 164–180]

Neilson, B. N., Stephens, C.L. & Pope, A.T. (2012). Psychophysiological and Performance Effects of Commercial Brain-Computer Interface Technology. Applied Psychophysiology and Biofeedback, 37, 4, 309.

Parker, T. (Director), & Stone, M., Brian Graden, B., Hader, B. (Writers). (2008, October 29) Pandemic 2: The Startling [Television series episode]. In T. Parker & M. Stone (Executive producers), South Park. New York, NY: Comedy Central.

Pope, A. T., & Bogart, E. H. (1992). Identification of Hazardous Awareness States in Monitoring Environments. SAE PAPER 921136; SAE, International Conference on Environmental Systems, Seattle, WA, United States, July 13-16, 1992.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic System Validates Index of Operator Engagement in Automated Task. Biological Psychology, 40, 187-195.

Pope, A. T., & Bowles, R. L. (1982). A program for assessing pilot mental state in flight simulators. Paper presented at the American Institute of Aeronautics and Astronautics.

Pope, A.T. & Stephens, C.L. (2011). 'Movemental': Integrating Movement and the Mental Game. Presentation at the 29th Annual ACM CHI conference on Human Factors in Computing Systems, May 9-12. Vancouver, BC, Canada.

Pope, A.T., & Stephens, C.L. (2012) Interpersonal Biocybernetics: Connecting through Social Psychophysiology. Presentation at the 29th Annual ACM International Conference on Multimodal Interaction, October 22–26. Santa Monica, California, USA.

Pope, A. T., & Stephens, C. L. (2015) The Interplay of Biocybernetic Adaptation and Biofeedback Training. World Usability Day @ HEC: User Engagement and Usability, November 12, 2015. Montreal, Canada.

Pope, A.T., Stephens, C.L., & Gilleade, K. M. Biocybernetic Adaptation as Biofeedback Training Method, Chapter 5 in Advances in Physiological Computing edited by Fairclough and Gilleade, Springer 2014.

Rammirez-Moreno, D. & Sejnowski T. (2012) A computational model for the modulation of the prepulse inhibition of the acoustic startle reflex. Biological Cybernetics, 106, 3, 169-176.

Research Integrations, Inc. for the FAA, Flight Crew Training for NextGen Automation Final Report, September, 2011

Roy, R.N., Frey, J.: Neurophysiological markers for passive brain–computer interfaces. In: Clerc, M., Bougrain, L., Lotte, F. (eds.) Brain-Computer Interfaces 1: Foundations and Methods. Wiley, Hoboken (2016).

Sarter, N.B.; Mumaw, R.J.; Wickens, C.D. (2007). "Pilots' Monitoring Strategies and Performance on Automated Flight Decks: An empirical study combining behavioral and eye-tracking data." Human Factors: The Journal of the Human Factors and Ergonomics Society Volume 49 (June 1, 2007): 347–357.

Scerbo, M. W. (2007). Adaptive automation. In R. Parasuraman & M. Rizzo (Eds.), Neuroergonomics: The brain at work (pp. 239-252). Oxford: Oxford University Press.

Solovey, E. T., Afergan, D., Peck, E. M., Hincks, S. W., and Jacob, R. J. K. (2015) Designing implicit interfaces for physiological computing: Guidelines and lessons learned using fNIRS. ACM Transactions on Computer-Human Interaction, vol. 21, no. 6 (2015) article no. 35.

Stephens, C.L., DeHais, F., Roy, R., Harrivel, A.R., Last, M.C., Kennedy, K.D., & Pope, A.T. (2018). Biocybernetic Adaptation Strategies: Machine Awareness of Human Engagement for Improved Operational Performance. Lecture Notes in Computer Science series, Volume 10915.

Stephens, C., Pope, A., Kennedy, K., Harrivel, A., Roy, R., Dehais, F., (2018). Biocybernetic Adaptation Strategies: Machine awareness of human state for self-regulation and improved operational performance, 32nd International BCS Human Computer Interaction Conference, Belfast, Ireland, July 2018.

Stephens, C.L., Pope, A.T., & Scerbo, M.W. (2012). Adaptive Automation for Mitigation of Hazardous States of Awareness. In G. Matthews (Eds.) The Handbook of Operator Fatigue. Burlington, VT; Farnham, Surrey; Ashgate.

Stephens, C., Prinzel, L., Harrivel, A., Comstock, J., Abraham, N., Pope, A., Wilkerson, J., Kiggins. D. (2017). Crew State Monitoring and Line-Oriented Flight Training for Attention Management. International Symposium on Aviation Psychology. Dayton, Ohio.

Terwilliger, P., Sarle, J., Walker, S., Espenschied, W., Stephens, C., Harrivel, A. (2020). A ResNet Autoencoder Approach for Time Series Classification of Cognitive State, MODSIM World 2020, Norfolk, VA, 5-7 May 2020, http://www.modsimworld.org/conference-papers/2020

Thomas, L., Gast, C., Grube, R., Craig, K. Fatigue detection in commercial flight operations: Results using physiological measures. Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics. 2015: Las Vegas, Nevada.

Wilhelm, F., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. Biological Psychology, 84, 552–569.

Wilson GF & Russell CA (2003) "Real-time assessment of mental workload using psychophysiological measures and artificial neural networks," Hum. Factors, vol. 45, no. 4, pp. 635-643, 2003.

VI. Appendices

A. Acronyms

- Advisory Circular (AC)
- Air Line Pilots Association (ALPA)
- Attentional human performance limiting states (AHPLS)
- Augmented Flight Deck Countermeasures(AFDC)
- Angle of Attack (AOA)
- Auxiliary Power Unit (APU)
- Advanced Qualification Program (AQP)
- Airplane State Awareness (ASA)
- Air Traffic Control (ATC)
- Blood Volume Pulse (BVP)
- Computer Based Training (CBT)
- Cockpit Motion Facility (CMF)
- Crew Resource Management (CRM)
- Crew State Monitoring (CSM)
- Department of Transportation (DOT)
- Electroencephalography (EEG)
- Electrocardiogram (EKG or ECG)
- Electronic Flight Book (EFB)
- Federal Aviation Administration (FAA)
- First Officer (FO)
- Flight Management Computer (FMC)
- Functional Near Infrared Spectroscopy (fNIRS)
- Federal Aviation Administration (FAA)
- Galvanic Skin Response (GSR)
- Heart Rate Variability (HRV)
- Human-in-the-Loop (HITL)
- Human Performance and Limitations (HPL)
- Joint Implementation and Data Analysis Team of CAST (JIMDAT)
- Line Operational Evaluation (LOE)
- Line Operational Simulations (LOS)
- Line-Oriented Flight Training (LOFT)
- Loss of Control In-Flight (LOC-I)
- Mode Control Panel (MCP)
- Minimum Equipment List (MEL)
- Navigation Display (ND)
- Non-Technical Skills (NOTECHS)
- Observe, Orient, Decide, and Act (OODA)
- Office of Inspector General (OIG)
- Out the Window (OTW)
- Primary Flight Display (PFD)
- Pilot Flying (PF)
- Pilot Monitoring (PM)
- Reduced Vertical Separation Minimum (RVSM)
- Safety Enhancement (SE)
- Scenarios for Human Attention Restoration using Psychophysiology (SHARP)
- Situational Awareness (SA)
- Subject Matter Expert (SME)
- Standard Operating Procedure (SOP)
- Special Purpose Operational Training (SPOT)

B. Frequency Domain fNIRS system description

This work was performed at NASA Glenn Research Center which complemented SE211-funded TASA work to produce a Frequency Domain fNIRS unit that meets integration requirements for HITL flight simulation testing. Acceptance testing must be performed, and IRB approval must be obtained prior to use in HITL testing for TRL 3 system demonstration. (NASA/TM—2020-220348, Mackey et al, 2020)

Functional Near Infrared Spectroscopy (fNIRS) Device

- Developed and fabricated at GRC
- One of a suite of instruments for the CSM
- Measures changes in oxygen levels in the brain non-invasively
- Based on the concept that hemoglobin absorbs near infrared (NIR) wavelengths of light
  - Measures scattered light rather than transmission of light
- Commercial instruments are too heavy for intended use in flight simulators and other test scenarios
- fNIRS Device uses Frequency Domain (FD) to reduce signal artifacts generated by movement and provide better signal to noise margins
- Uses FD to reduce movement artifacts - US patent 9 848 812 B1 issued December 27, 2017
- Imagent manufactured by ISS
  - Weighs more than 80 pounds
  - 2 Electronics boxes are each 10" x 17" x 18"
  - Not very portable in field applications
  - a FD device developed with >20 year-old technology
- fNIRS Device Advantages
  - Single enclosure: 7" x 19" x 22"
  - Weighs ~15 pounds
  - Portable for use in field applications
  - a FD device using modern optoelectronic technology

Headgear

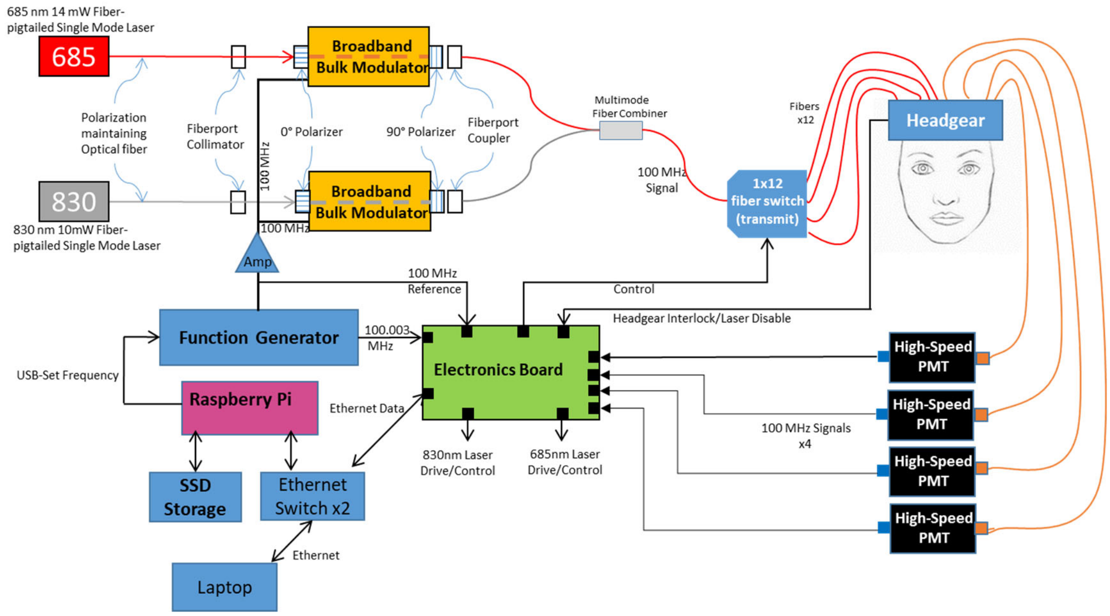
- 4 sets of transmit/receive optodes
- Headgear has multiple ports for adjusting optodes placement
- Complete redesign of previous GRC-developed headgear
- Block between Transmitters & Receiver to avoid surface scatter
- Interlock disables laser if headgear removed
- 3D printed design allows for quick, low cost modification
- Safe output of optodes for skin exposure verified by GRC Radiation Safety Officer

Hardware

- Custom designed for circuit boards and optical hardware
- Contains additional enclosures fabricated via 3D printing and painted with conductive coating to EMI protect circuit boards
- Electronics board designed to bypass amplification and filter stages as needed
- Function generator packaged inside enclosure for easy change of modulation and cross correlation frequencies

Software

- Hardware setup and control: modulation freq. selection, Timestamp, data start/stop, error checking
- Networking to facility control room
- Acquisition and storage of received waveforms
- Calculation, display and storage of AC, DC and Phase information
- Additional functionality added: Calculation, display and storage of Hb/HbO information
- Software can be updated remotely


685 nm 14 mW Fiber-pigtailed Single Mode Laser
685
830
830 nm 10mW Fiber-pigtailed Single Mode Laser
Polarization maintaining Optical fiber
Fiberport Collimator
0° Polarizer
90° Polarizer
Fiberport Coupler
Broadband Bulk Modulator
Broadband Bulk Modulator
100 MHz
100 MHz
Amp
Multimode Fiber Combiner
100 MHz Signal
1x12 fiber switch (transmit)
Fibers x12
Headgear
100 MHz Reference
Control
Headgear Interlock/Laser Disable
Function Generator
100.003 MHz
Electronics Board
USB-Set Frequency
Raspberry Pi
Ethernet Data
830nm Laser Drive/Control
685nm Laser Drive/Control
100 MHz Signals x4
High-Speed PMT
High-Speed PMT
High-Speed PMT
High-Speed PMT
SSD Storage
Ethernet Switch x2
Ethernet
Laptop

C. CSM Prototype Translational System (CPTS) Technical Requirements

CPTS Technical Requirements			Rack-mounted				Otherwise			
CPTS requirements	**Power Type**	**Power [A]**	**Weight lbs**	**H [in]**	**W [in]**	**L [in]**	**Weight lbs**	**H [in]**	**W [in]**	**L [in]**
Spire	internal rechargeable battery		<0.1 lb				0.6 lb			
Muse	internal rechargeable battery		0.2 lb	2	1.25	0.75	0.8 lb			
Empatica	internal rechargeable battery		<0.1 lb	8.25	6	2	0.4 lb			
Nexus (redundant, more obtrusive, legacy system)	removable rechargeable battery		2.8 lbs (includes EKG leads, GSR, resp belt)	5.5	4.75	1.75	7.0 lbs			
Tobii pro	removable rechargeable battery		0.6 lbs (glasses + amplifier box)	(size of normal eyeglasses)			4.8 lbs (includes battery charger)			
CSM 1 Laptop running Lab Streaming Layer (LSL) for local data acquisition from sensors *	120VAC 1 phase 60 Hz	1.8 A	5.2	14	9	0.75				
Miscellaneous cables and connections			2.2							

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01/07/2021	TECHNICAL MEMORANDUM	

4. TITLE AND SUBTITLE	
CAST SE211 Training for Attention Management Final Report	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Angela R. Harrivel, James R. Comstock, Jr., Lawrence J. Prinzel, Chad L. Stephens, Kellie D. Kennedy, Alan T. Pope	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER 340428.02.40.07.02

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
NASA Langley Research Center Hampton, VA 23681-2199	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)** NASA
National Aeronautics and Space Administration Washington, DC 20546-001	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** NASA/TM-20205003444

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT

NASA Langley led the implementation of all outputs of a research Safety Enhancement adopted by CAST in 2014 to address Training for Attention Management.1 Attention-related human performance limiting states (AHPLS), which may reduce pilot airplane state awareness, may be indicated objectively by covert or physiological markers of limited performance. Existing NASA Crew State Monitoring (CSM) technologies were applied to sense and quantify these covert markers. Monitoring cognitive state changes on a moment-to-moment basis provides quantitative data regarding the onset of AHPLS. CSM uses psychophysiological signals and machine learning to predict cognitive state passively...

15. SUBJECT TERMS

Aviation Safety, Training, Attention, Crew State Monitoring

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			HQ - STI-infodesk@mail.nasa.gov
U	U	U	UU	51	**19b. TELEPHONE NUMBER** *(Include area code)* 757-864-9658