The Crew Workload Manager: An Open-loop Adaptive System Design for Next Generation Flight Decks

Michael C. Dorneich, Bretislav Passinger, Christopher Hamblin,
Claudia Keinrath, Jiři Vašek, Stephen D. Whitlow
Honeywell Laboratories

Martijn Beekhuyzen
Delft University of Technology

This paper presents an open loop adaptive system intended to address workload imbalances in future, high-workload flight decks. Air traffic in Europe is expected to more than double by 2020. New technologies being proposed will significantly add to pilot roles and responsibilities, and has the potential to add further periods of high workload to pilot operations. The CAMMI (Cognitive Adaptive Man Machine Interface) program addresses human factors priorities in the aviation domain by developing concepts that balance operator workload, support added future operator roles and responsibilities and resulting new task and information requirements, while allowing operators to focus on the most safety critical tasks. The Crew Workload Manager (CWLM) is a research prototype that objectively measures, compares, and displays the workload between pilots, and can recommend task sharing or automate lower order tasks as necessary. It is expected that the CWLM will minimize the time pilots spend in unbalanced workload conditions, and thereby reduce errors and pilot fatigue, and improve crew resource management. An evaluation plan is outlined that utilizes the novel Shared Aviation Task Battery.

INTRODUCTION

The current Air Traffic Management (ATM) system is already operating at its limits. New innovations in the structure of ATM will be needed to handle the expected increase in traffic. However, such innovations will necessarily impact the distribution of tasks and responsibilities between the air and the ground, as well as the roles and responsibilities of the automation and human operators. These two bottlenecks lead to high workload situations, even under today's conditions. New concepts like precision 4D path following, self separation, and closer aircraft spacing will be needed to increase capacity and efficiency. These types of ATM innovations will necessarily affect aircrews operations and will impact their workload.

To meet the challenges of future ATM environments, programs like SESAR (SESAR, 2006) and NextGen (NextGen, 2007) have been launched to accommodate the trend of the air traffic growth and prepare the ATM (Air Traffic Management) for the demand of 2020 and beyond. These programs aim to develop new technological capabilities, more automated visualization and decision aids, changes in procedures, and changes in pilot roles and responsibilities. Such changes must be accompanied with a re-thinking of the design of on-board systems, where human performance capabilities and limitations should be clearly taken into account and drive the design. Human and machine agents must work together as a unit and gain a common understanding of the situation in which they are operating in order to achieve the ambitious operational, performance, and safety goals. Given the expected changes, pilots will be faced with managing increased levels of automation, different communication methods, and more decision making responsibilities. ATM operations today are characterized by high cognitive demand; future systems are only expected to become more demanding on pilots. The increased information integration requirements and automation that make up these future systems mean that pilots could be susceptible to dangerous deficiencies of situation and automation awareness. It is well established that integrated human-automation systemsperform complex tasks in dynamic environments better than manual or fully automated systems. But it is also understood that there can be a significant penalty in human attention and situational awareness when interacting with highly automated systems. Examples include the increased difficulty in trying to understand what automation is doing, locating relevant information, even programming the automation to perform its functions. This can degrade the efficiency and intended safety of the automated functions. Some prominent human-automation interaction problems are uneven distribution of workload, inappropriately aligned trust in automation, breakdown in mode and automation awareness, delays in finding, interpreting and integrating information, and human input errors. In multi-pilot aviation, Crew Resource Management (CRM) was developed to improve air safety by focusing on the cognitive and interpersonal skills needed to make optimal use of resources - automation, procedures, and pilots.

The Cognitive Adaptive Man Machine Interface (CAMMI) project is supported by the "Human-Centric Design (HCD) of Embedded Systems sub-program (SP-8)" of the ARTEMIS annual work program (ARTEMIS AWP, 2008). The program is expected to establish novel methodologies for the design and development of human-in-the-loop adaptive control systems, with a special emphasis on the cognitive state of the human operator in safety critical domains. The CAMMI program intends to develop adaptive pilot interface technologies that utilize both the context and real-time measurements of the pilots' cognitive state to trigger automation assistance when needed most. The goal is to keep pilot workload balanced and allow the pilot to focus on the most safety critical tasks, resulting in increased safety, higher capacity, and increased efficiency of operations. This approach will mitigate automation-related shortcomings, leverage human strengths, and augment human performance specifically when assessed human cognitive capacity falls short of the demands imposed by task environments.

Adaptive Systems

The CAMMI system is designed as a human-computer interaction system classed as an adaptive automation system. Adaptive systems aim to enhance joint human-machine performance by having the system invoking varying levels of automation support in real time during task execution. As such it will have the following aspects (Scerbo, 2001):

- An adaptive system must make timely decisions on how best to use varying levels of (adaptive) automation to provide support in a joint human-automation system
- The CAMMI system can make independent judgment and has the authority to initiate changes.
- Adaptations typically happen in real-time
- Adaptation strategies can be turned on or off, depending on cognitive workload of the pilot.
- In order to make appropriate decisions, the CAMMI system must be aware of the world around it

The level of automation in a joint human-automation system can vary from completely manual, where the whole task is performed by human, to fully autonomous. (Sheridan & Verplank, 1978).

There have been several taxonomies that describe various levels of automation. Sheridan (1992) describes 10 levels of automation ranging from no assistance to the computer ignores the human and decides everything. Similarly, Proud et al (2003) describes eight levels of automation. Adaptive automation can either provide adaptive aiding, which makes a certain component of a task simpler, or can provide adaptive task allocation, which shifts an entire task from a larger multitask context to automation (Parasuraman, Mouloua, & Hilburn, 1999). Additionally, automation can change the mode of interaction by adapting either the type, format, or amount of information presented to the human. In a general framework of CAMMI four types of mitigation will be considered:

1) Task scheduling (e.g. direct the operator to higher priority tasks, defer lower priority tasks, or assist in task switching),
2) Modify the interaction with the system (e.g. de-clutter displays, highlight important information, change the modality of incoming information),
3) Task offloading (e.g. automate lower priority tasks),
4) Task sharing (e.g. provide automation assistance on tasks – simplify tasks).

Proud and collegues (2003) illustrates the mapping of the eight autonomy levels onto the mitigation taxonomy used in the discussed domain.

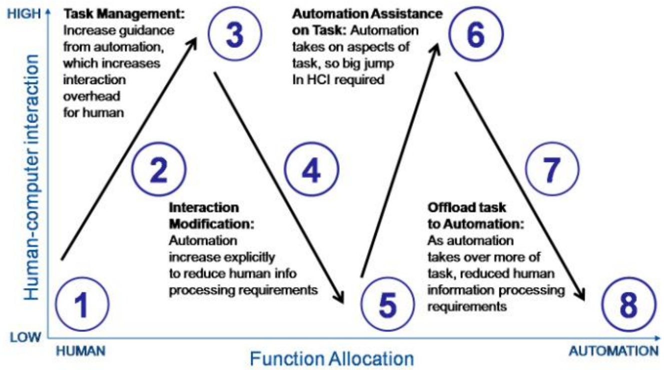

Figure 1. Possible mapping of the 8 levels of automation onto Mitigations Taxonomy.

Currently, adaptive systems can derive their inferences about the cognitive state of the operator from mental models, performance on the task, or from external factors related directly to the task environment (Wickens & Hollands, 2000). Recent research has explored deriving cognitive state from direct sensor-based measurements (Wilson & Eggemeier, 1991; Makeig & Jung, 1995; Gevins & Smith, 2000; Dorneich, Mathan, Ververs & Whitlow, 2008). In CAMMI the adaptive automation will be triggered, in part, by a real-time assessment of cognitive state.

The next section will introduce a specific flight deck mitigation developed as part of CAMMI. The subsequent section will describe the design of an aviation-based multi-attribute task battery to serve as the basis of an evaluation of the mitigation. An evaluation plan is outlined utilizing the Shared Aviation Task Battery (SAT-B); the evaluation is in progress and results will be presented at a later date.

THE CREW WORKLOAD MANAGER

Problem description

Experience levels between two pilot crews are typically asymmetric. Often crew rotation pairs experienced pilots with less experienced co-pilots; likewise, the workload between the pilot flying (PF) and pilot monitoring (PM) is often asymmetric, as the tasks for each role can be quite different. The PF concentrates on flying the aircraft, engaging the flight management systems, and maintaining situation awareness of the route, traffic, terrain, and the external environment conditions. The PM is responsible for monitoring the flight management and aircraft controls of the PF, as well as supporting tasks such as monitoring communications, systems management, and procedure checklists. Ideally, pilots work together to balance workload. For instance, the PM could reduce PF workload by assuming greater task responsibilities during periods where his or her workload was much less than the workload of the PF. Anecdotal evidence, however, suggests that it is difficult to assess the cognitive workload of a fellow pilot, and often pilots are unaware, or areonly marginally aware of the cognitive workload state of the other pilot. High workload conditions can cause a pilot to inappropriately focus their attention on the high workload task and become less attentive and less responsive to other tasks. Often it is the lack of cues that alert the other pilot that there is an issue. Even though some airlines have instituted explicit protocols that allow pilots to communicate their workload to each other (e.g., “I’m red – how about you?”) some pilots might be reluctant to acknowledge that they are overloaded or fatigued and forego re-allocation of tasks which could maintain a more optimal workload balance.

The Mitigation

The premise of the Crew Workload Manager (CWLM) is that by making the workload of pilots visible to each other, the crew can better detect workload imbalances and take steps to address it sooner, resulting in less time spent overloaded. Balancing workload and reducing the time spent in high workload should lead to improved performance, fewer errors and less fatigued pilots

The CWLM objectively measures, compares, and displays the workload between pilots, and can recommend task sharing. The automation acts as an objective, non-threatening third party that senses and communicates the state of each pilot. By acting as an “honest broker,” the state assessment might be better received and responded to than if the one of the pilots insinuates that the other pilot is overloaded or drowsy. The workload for each pilot is displayed on the instrument panel such that it is visible to both.

For this study, the CWLM will depict estimated workload for both the PF and PM. Workload for the left operator depicted left of the display’s centerline; workload for the right operator is depicted right of the display’s centerline. The CWLM will depict three workload states: Low, Nominal, and High. Low workload is indicated by a narrow band closest to the centerline while high workload is indicated by wide band furthest from the centerline. Examples of the CWLM display formats are shown in **Error! Reference source not found.**.

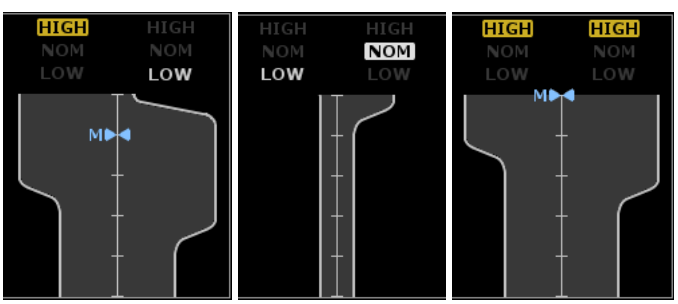

Figure 2. Three different states of the CWLM.

The CWLM is a research prototype display that depicts a pilot’s real-time cognitive workload based on his or her neurophysiological data. EEG (brain) and ECG (heart) sensors can be used as the basis of a cognitive state classifier to identify when assistance is most needed. Research has shown that EEG activity can be used to assess a variety of cognitive states that affect complex task performance, such as working memory (Gevins & Smith, 2000), alertness (Makeig & Jung, 1995), engagement (Pope, Bogart & Bartolome, 1995), executive control (Garavan, Ross, Li, & Stein, 2000), and visual information processing (Thorpe, Fize, & Marlot, 1996).

In previous work, we achieved an overall classification accuracy into the 90% range (Dorneich et al., 2006).

When workload is out of balance between operators, or if workload for one or both of the operators is determined as “High” an advisory notification will trigger the appropriate alert message will appear in the crew alerting system (CAS) window (see Figure 3).

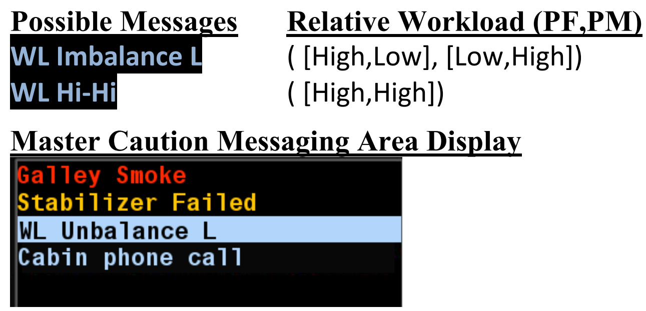

Figure 3. Example alert messages associated with the CWLM.

CWLM as an Open-loop Solution

Typically adaptive systems are thought of as “closed-loop” systems, where output of the sensing (of the human, the system, and the environmental context) is used to drive automated responses that then in turn effects the human, system, and environment. It is closed loop in the sense that the system both senses the initial trigger and initiates the mitigation. The CWLM is an open-loop mitigation, where it is up to pilots to address the situation by adapting their workload distribution. Also, the automation is not the initiating agent in any changes to the task environment. The automation simply detects the event, and it is up to the human operator to initiate any changes to mitigate the condition of concern. The CWLM acts as detection and awareness tool. The emphasis lies on the decision of the pilots to deal with the situation at hand. The expected pilot-initiated mitigation would be one of task management (see Figure 1), where the pilots would dynamically re-assign tasks between themselves to better balance workload.

THE SHARED ATTRIBUTE TASK BATTERY

The evaluation plan discussed in the next section will utilize the Shared Aviation Task Battery (SAT-B) to create workload in an aviation-like setting. The SAT-B is based on Multiple Attribute Task Battery (MAT-B) (Comstock & Arnegard, 1992), which is a well established experimental test bed. The original MAT-B was designed for single person operation. In contrast, the SAT-B is designed to allow two people to each have a (mirrored) screen where tasks are shared between the two operators. The SAT-B simulates four simple cognitive tasks running in parallel, designed to resemble cockpit operations. The four cognitive tasks are:

- *System monitoring* – This task simulates the demands of monitoring gauges and warning lights. The participant monitors two indicators and four dials, and reacts to abnormal/emergency condition**.**
- *Communications*. This task simulates receiving audio messages from Air Traffic Control. Participants respond

only to messages preceded by the operators' call sign, and respond to the subsequent command.

- *Resource management.* This task is considered as a analog of a fuel balancing task. The participant must monitor and control levels in two tanks via system of tanks and pumps with different flow rates.
- *Tracking.* The task represents direct control of aircraft attitude and the demands of manual control, simulated by the tracking task. The task is to monitor and control the random deviations of the attitude direction indicator from central position via a joystick.

The SAT-B interface (Figure 4), is an abstraction of tasks a pilot is required to perform during a typical flight. The attitude direction indicator (ADI) is shown in the upper left hand corner (A) and is used to perform the tracking task. The dial indicators (B) in the upper right hand corner of the display are used to perform the monitoring task. The systems display (C) is shown in the lower left area of the display and is used to perform the resource management task. Finally, the radio and navigation channel indicators (D) are shown in the lower right hand display and are used to perform the communications task.

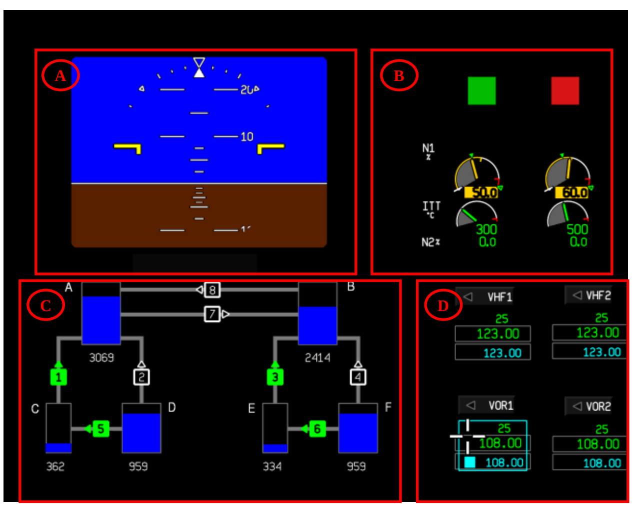

Figure 4. The SAT-B is designed for dual operation between two participants.

Participants are taught that if they feel their performance on a task is deteriorating, they may off-load a task to their partner. Likewise, if a participant feels his or her partner is overwhelmed or performance is deteriorating, the participant can also help their partner by taking over a task.

METHOD

This section describes the evaluation plan for assessing the CWLM, utilizing the SAT-B. The evaluation is in progress and the results will be reported at a later date.

Objective

Demonstrate improved operator performance and crew resource management (CRM) with the CWLM mitigation. It is our hypothesis that, with assistance from the CWLM, the participant will better recognize high workload conditions and will respond faster by either on- or off-loading tasks to/from the confederate; and thereby, balance workload more evenly between the operators.

Participants and Tasks

The goal of the participants is to balance workload so as to optimize performance of both operators. Balancing workload requires the participant to recognize when he/she is overloaded and pass off tasks to the other pilot. Conversely, the participant must also be able to recognize when the other pilot is overloaded and actively take on tasks. For each scenario participants will be assigned individual tasks within the SAT-B environment each with initial task load and the participants must balance overall crew workload by on- or off-loading tasks depending on the crew member real-time workload assessment and self assessment and distribution of task load in the crew members . Participants will on- or off-load tasks by pressing "Accept" or "Off-load" buttons displayed in the SAT-B task windows.

Use of a Confederate

SAT-B can be run with two participants simultaneously or with one participant and a confederate. For this experiment, we chose to use a conferderate in order to exert more control on the task load manipulation of the remaining participant. Use of a confederate as the co-pilot also simplifies the experimental design. Since tasks will be divided among participants it will be necessary to vary the workload of the individual SAT-B tasks (e.g., tracking, monitoring, etc) by increasing or decreasing the difficulty or frequency of the individual tasks.

Experimental Design

There are two independent variables: 1) CWLM Mitigation: On/Off, and 2) Workload Conditions: nine combinations of low, nominal, and high workload for two pilots (i.e., LL, LN, LH, NL, NN, NH, HL, HN, HH).

In the unmitigated condition, the participant and the confederate will conduct their tasks on the SAT-B without the aid of the CWLM. In mitigated trials, the CWLM will provide displays and alerts. Conditions will be quasi-random in that the workload patterns will be random However levels of workload will be contiguous (e.g., workload will not jump from Low to High without transitioning through Medium)

Dependent variables will be: 1) performance on SAT-B tasks, 2) time spent in unbalanced workload, 3) detection time of unbalanced workload, and 4) workload as measured by NASA TLX (Hart & Staveland, 1988), will corroborate the workload detected by the EEG and ECG sensors.

Procedure

Participants will be introduced to the study and briefed on their role and responsibilities. Participants will be trained on the CWLM and the SAT-B tasks. Subjective workload measures will be measured via the NASA TLX. Participants will experience every combination of the two independent variables, in trial runs of at least two minutes. The order of the trials will be randomized for each participant. One of the two people performing the SAT-B will be a confederate and the other will be the participant. After the trials are completed, the participant will fill out a survey to give subjective feedback on the CWLM.

DISCUSSION

The Crew Workload Manager (CWLM) is a research prototype that objectively measures, compares, and displays the workload between pilots, and can recommend task sharing or automate lower order tasks as necessary. The evaluation outlined above is in progress, and will be presented at the conference and documented at a later date. It is expected that the CWLM will minimize the time pilots spend in unbalanced workload conditions, and thereby reduce errors and pilot fatigue, and improve crew resource management. We expect to see less time spent in unbalanced workload in the mitigated condition when one of the two pilots is under high workload before and task-swapping has occurred. We expect participants to detect an imbalanced earlier in the mitigated trials vs. the unmitigated trials. We expect performance to be improved or at least maintained in the mitigated trails when the workload is high for one of the two people. The goal is to keep pilot workload balanced and allow the pilot to focus on the most safety critical tasks, resulting in increased safety, higher capacity, and increased efficiency of operations. This approach will mitigate automation-related shortcomings, leverage human strengths, and augment human performance specifically when assessed human cognitive capacity falls short of the demands imposed by task environments.

ACKNOWLEDGMENTS

This research has been performed with support from the EU ARTEMIS JU project CAMMI SP-8-GA No. 100008. Any opinions, recommendations, findings, or conclusions expressed herein are those of the authors and do not necessarily reflect the views of the ARTEMIS JU and/or the European Commission.

REFERENCES

ARTEMIS Annual Work program (2008). Retrieved from https://www.artemisju.eu/attachments/20/ARTEMIS_Annual_Work_Programme_2008.pdf

Comstock, J. R., & Arnegard, R. J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research (Tech. Memorandum No. 104174). Hampton, VA: NASA Langley ResearchCenter.

Dorneich, M.C., Mathan, S., Ververs, P.M, & Whitlow, S.D. (2008). Cognitive State Estimation in Mobile Environments. In D. Schmorrow and K. Stanney (Eds.), Augmented Cognition: A Practitioner's Guide (pp. 75-111). Santa Monica, CA: HFES.

Dorneich, M.C., Ververs, P.M., Mathan, S., Whitlow, S.D., Carciofini, J., and Reusser, T. 2006. Neuro-physiologically-driven adaptive automation to improve decision making under stress. *Proceedings of the Human Factors and Ergonomics Society Conference 2006*. San Francisco, CA.

Dorneich, Michael C., Whitlow, Stephen D., Ververs, Patricia May, Rogers, William H. (2003). "Mitigating Cognitive Bottlenecks via an Augmented Cognition Adaptive System," Proceedings of the 2003 IEEE International Conference on Systems, Man, and Cybernetics, Washington DC, October 5-8, 2003 (Invited), pp. 937-944.

Garavan, H., Ross, TJ, Li, S.-J., and Stein, E.A. 2000. A parametric manipulation of central executive functioning using fMRI. *Cerebral Cortex* 10:585-592.

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

Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Human Mental Workload, (Eds. P. A. Hancock and N. Meshkati), Elsevier Science Publishers, North-Holland, 139-184.

SESAR Consortium. (2006, July 20). D1 Air Transport Framework: The current situation. Retrieved from http://www.sesarju.eu/gallery/content/public/DLM-0602-001-03-00.pdf

Makeig, S., & Jung, T -P. (1995). Changes in alertness are a principal component of variance in the EEG spectrum. NeuroReport, 7(1), 213–216.

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

NextGen. Joint planning and development office, (2007). Concept of operations for the Next Generation Air Transportation System. Retrieved from _http://jpdo.gov/library/NextGen_v2.0.pdf.

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

Proud, R. W., Hart, J. J., & Morzinski, R. B. (2003). Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach. Conference on Performance Metric for Intelligent Systems.

Scerbo, M.W. 1996. Theoretical perspectives on adaptive automation. In Automation and human performance: Theory and applications, eds. R. Parasuraman and M. Mouloua, 37-63, Mahwah, NJ: Lawrence Erlbaum Associates.

Sheridan, T. B., & Verplank, W. (1978). Human and Computer Control of Undersea Teleoperators. Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.

Sheridan, T. B. Telerobotics, Automation, and Human Supervisory Control. The MIT Press. 1992.

Thorpe S., Fize D., and Marlot C. 1996. Speed of processing in the human visual system. *Nature*, 381: 520-522.

Wickens, C.D. and Hollands, J. 2000. *Engineering Psychology and Human Performance*. Prentice Hall, 3rd edition..

Wilson, G. F., & Eggemeier, F. T. (1991). Physiological measures of workload in multi-task environments. In D. Damos (Ed.), Multiple-task performance (pp. 329–360). London: Taylor & Francis.