Regular Paper

# Predictive Mental Workload Modeling for Semiautonomous System Design: Implications for Systems of Systems

John M. Colombi,[1,*] Michael E. Miller,[1] Michael Schneider,[2] Major Jason McGrogan,[3]

Colonel David S. Long,[4] and John Plaga[5]

[1]*Department of Systems and Engineering Management, Air Force Institute of Technology, Wright-Patterson* AFB, OH 45433
[2]*Aeronautical Systems Center/Engineering Directorate, Wright-Patterson* AFB, OH 45433
[3]*Aeronautical Systems Center/B-52 Program Office, Wright-Patterson* AFB, OH 45433
[4]*Air Force Research Laboratory/Sensors Directorate,* 2241 *Avionic Circle, Wright-Patterson* AFB, OH 45433
[5]711*th Human Performance Wing/Human Performance Integration Directorate,* 2215 *First Street, Wright-Patterson* AFB, OH 45433-7028

Received 26 April 2011; Revised 9 January 2012; Accepted 9 January 2012, after one or more revisions
Published online 11 July 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/sys.21210

## ABSTRACT

Predictive mental workload modeling is one established tool within the broad systems engineering activity of Human Systems Integration (HSI). Using system architecture as the foundation, this paper explores the use of Multiple Resource Theory to create representative workload models for evaluating operational system-of-systems (SoS) concepts. Through careful consideration of task demands, conflict generated between tasks, and workload mitigation strategies, informed design decision can improve overall human-system performance. An example involving a single pilot controlling multiple remotely piloted aircraft (RPA) is presented to illustrate the use of workload modeling. Several observations are made that drive measurably excessive workload: multitasking, communications, continuously updating situational awareness and mission planning. In addition, three metrics are proposed for incorporating human workload analysis during system design. This technique has applicability across a wide range systems-of-systems and operational concepts involving complex human-system interactions. © 2012 Wiley Periodicals, Inc. Syst Eng 15: 448–460, 2012

Key words: system-of-systems; workload; modeling and simulation (M&S); human systems integration (HSI)

## 1. INTRODUCTION

Advances in communication and networking technology have contributed to the development of complex, distributed systems, and System-of-Systems (SoS). These systems can involve large social and technical networks and interactions with these system-of-systems can impose significant stress on human operators. One method to reduce the demands these

*Author to whom all correspondence should be addressed (e-mail: john.colombi@afit.edu; john.colombi@incose.org; michael.miller@afit.edu; Michael.schneider2@wpafb.af.mil; Jason.McGrogan@wpafb.af.mil; david.long@wpafb.af.mil; john.plaga@wpafb.af.mil).

SoS place on the human operator is to automate a portion of the tasks typically allocated to a human. This trend is present in multiple industries; including commercial aviation [Ramirez, 1997; Byrne and Kirlik, 2005], automotive [Ramirez, 1997]; and power and chemical facility control rooms [Allender, 2000; Chen and Joyner, 2009]. The use of automation has the potential to not only reduce the workload imposed on a human operator but can reduce the number of operators necessary to run and support a system, which can reduce manpower and system life cycle costs.

In one specific Defense and Aerospace application, Unmanned Aerial Vehicles (UAVs) or Remotely Piloted Aircraft (RPA), are being deployed to provide capability needs previously addressed through manned aircraft. These UAVs are deployed within an SoS that involves possibly several unmanned vehicles, multiple ground stations and their crews, other manned surveillance aircraft, ground forces and their equipment, Command and Control operations centers, and intelligence processing systems. The need for increasing the number of pilots to enable long duration missions and the reduction in pilot workload due to an increase in automation has motivated an operational concept that involves a single pilot simultaneously controlling multiple RPA. This concept of operations has been documented in the US Air Force UAV flight Plan [USAF, 2009]. While automation can eliminate tasks that would be performed by the human operator [Dixon et al., 2005], the need to control multiple aircraft, potentially with different missions, can add complexity and result in increased pilot workload. In fact, an increase in pilot workload beyond human limits has the potential to limit the viability of this concept of operations. The original goal of this research was to investigate the impact of this operational concept on human mental workload to determine the viability of this operational concept in the context of human workload constraints within an operational SoS.

## 2. BACKGROUND

As system complexity increases, demands on the human operator can also increase. This observation has led the US Department of Defense (DoD) to require a Human Systems Integration (HSI) plan during major acquisition programs [Mercer, 2008]. This requirement is intended to improve the integration of human considerations from the nine domains of manpower, personnel, training, human factors engineering, system safety, environmental safety, occupational health, habitability and survivability during the system development life cycle [Wagenhals, Liles, and Levis, 2009; Booher, 2003] with the aim to reduce the life cycle cost of the system. This concept requires tradeoffs between the complexity of the system, the number of operators required to operate the system, the difficulty of system operation, and the necessary knowledge skills and abilities of the operators. Unfortunately, design and acquisition is frequently performed individually for each system within an SoS while any human operator will often interact with multiple systems within a complex SoS, each of which can impose a significant level of workload. Therefore, human performance and limitations, including mental workload, should be evaluated within the context of the SoS with which the operator is engaged.

### 2.1. Mental Workload

Mental workload expresses the task demands placed on an operator [Beevis et al., 1999]. Task demand, or task load, often considers the goals to be achieved by the operator, the time available to perform the tasks necessary to accomplish the goals, and the performance level of the operator [Hardman et al., 2008]. Therefore, workload increases when the number or difficulty of tasks necessary to perform a goal increase, or when the times allotted to complete these tasks decrease. Assuming that the operator has a given amount of mental resources (e.g., attention, memory, etc.) that he or she can utilize to complete the necessary tasks, mental workload corresponds to the proportion of the operator's mental resources demanded by a task or set of tasks. Several methods have been employed to measure and quantify mental workload over the past 4 decades and have been summarized in numerous publications [Beevis et al., 1999; Gawron, 2008; Stanton et al., 2005]. Truly objective measures of workload have not been developed; instead measures such as subjective ratings of perceived workload, measurement of performance degradation of secondary tasks, measurements of performance when conducting a primary task, or changes in physiology have often been developed and used as a proxy measure to estimate mental workload. The most common of these approaches is to obtain subjective ratings of perceived workload through methods such as the NASA Task Load Index [Hart and Staveland, 1988] or the Subjective Workload Assessment Technique [Reid and Nygren, 1988].

Mental workload influences operator performance such that an optimal level of workload exists. Very low workload levels do not provide appropriate stimulation and result in low human performance [O'Hanlon, 1981; Young and Stanton, 2002]; conversely, very high workload levels provide excessive demands, also resulting in reduced human performance [Deutsch and Pew, 2004; O'Donnell and Eggemeier, 1986: 42-3]. Additionally, when operating systems that require an excessive number or difficulty of tasks to be performed within limited time, the human operator will become overloaded and will be incapable of performing all of the tasks, potentially foregoing tasks that are critical for safe system operation.

Mental workload can impact considerations across many HSI domains. For example, as an individual's skill level is increased, his or her proficiency at performing a task increases and his or her perception of workload decreases [Law, 2006]. Therefore, training and personnel selection can impact the workload of the operator. Further, if the operator is incapable of performing system tasks during critical mission segments, survivability can be compromised. Manpower should involve workload considerations as an individual's workload can be decreased by dividing the necessary tasks among multiple individuals, allocating a portion of the workload to each of the multiple individuals. The tasks of multiple operators can alternately be combined and assigned to a single individual, theoretically improving human performance, if each operator is undertasked.

## 2.2. Predicting Mental Workload

Although an operator's mental workload can be measured or estimated as he or she operates a system, this type of measurement requires a working prototype of a system and trained operators. Further this measurement is expensive and time-consuming during early conceptual design, in which numerous, very diverse system concepts might be considered. Instead, various modeling tools have been developed that permit the estimation of human mental workload. These tools permit mental workload to be estimated in the absence of an operational facsimile of the operator interface. Generally, two complementary tools are required, including a theory, which permits the estimation of human mental workload, and a computational engine to facilitate the necessary calculations.

Several theories of human mental workload exist. These theories can be classified as either temporally focused or demand difficulty focused. Temporally focused theories, including Single Channel Theory (SCT) [Welford, 1967] and Information Processing/Perceptual Control Theory (IP/PCT) [Hendy, East, and Farrell, 2001], focus on task time as a measure of workload, assuming that tasks must be performed asynchronously and multitasking is not possible. IP/PCT includes the concept of task difficulty, as a modulator of task time, which is used to predict workload [Parasuraman and Rovira, 2005]. Demand difficulty models include Single Resource Theory (SRT) [Kahneman, 1973], Visual-Auditory-Cognitive-Perceptual (VACP) theory [McCracken and Aldrich, 1984], Malleable Attentional Resource Theory and Multiple Resource Theory [Wickens, 2008]. These theories allow for multitasking and differ in their model of resources and resource interactions to predict workload. The dynamic workload model merges concepts from temporal or demand-focused theories alone; modeling workload as a function of a 3-dimensional vector that includes time to act, perceived distance until goal completion, and the effort required to accomplish the goal [Hancock and Caird, 1993]. The appropriateness of any of these theories is domain-specific, and theory selection will depend upon the likelihood of the system to impose multiple tasks on the individual, the ability to estimate the necessary inputs, and the desired output in terms of available times or estimates of perceived workload.

In the current investigation, Multiple Resource Theory (MRT) was selected as it has been shown to provide robust estimates of mental workload for tasks that require the operator to receive and respond to information through multiple information channels and perform concurrent tasks. This characteristic is important in the system under investigation as it is common for a pilot to listen to an auditory input while providing fine motor control to navigate the RPA in response to a video feed; simultaneously receiving input and providing output through auditory, visual, and fine motor channels. According to MRT, human operators can employ multiple mental resources simultaneously, permitting a degree of parallel processing. However, mental workload can become excessive when a given task requires more of a mental resource than the human operator can provide or when the operator is asked to perform multiple tasks simultaneously that require the same mental resource and together the multiple tasks demand more mental resources than are available. Further, tasks that require the same mental resource must compete for this resource which results in mental conflict. This competition for a common resource contributes to additional workload. The MRT algorithm sums the task demand workload on each channel and the conflict workload, which is a weighted summation of intra- and cross-channel task demand to arrive at a final estimate of workload.

The professional version of Improved Performance Research Integration Tool (IMPRINT PRO) was used as a modeling tool for the current investigation [Parasuraman and Rovira, 2005; Keller et al., 2002]. This tool, which was developed by the US Army Research Laboratory's Human Research and Engineering Directorate, permits discrete event models to be integrated with embedded tools for calculation of mental workload according to MRT. This modeling tool permits mental workload values to be assigned to defined operator tasks, the frequency or duration of which can be stochastically determined. Alternate workload modeling tools are commercially-available, including the Man-Machine Integration Design and Analysis System [Stanton et al., 2005], the Queuing Network-Model Human Processor [Boles and Adair 2001], and the Integrated Performance Modeling Environment [Law and Kelton, 2000]. However, IMPRINT was selected since it facilitates the use of MRT and has been applied extensively by the US Army to evaluate human interface designs and operator configurations for military vehicles [Mitchell and Samms, 2007, 2009; Mitchell 2009] and unmanned systems [Pomranky et al., 2007]. This and similar tools have previously been discussed within a systems design framework for use during early design or acquisition [Glenn et al., 2004; Handley and Smillie, 2010], and so the application of this tool within a systems design framework is not unique. This tool was selected, however, to investigate the feasibility of designing a ground control station to permit a pilot to simultaneously control multiple RPA as part of an early systems engineering evaluation. Through this process the utility of these tools to the design of an SoS and automation aids is illustrated.

# 3. MODEL DEVELOPMENT AND SoS ANALYSIS

The connectivity of modern information systems requires operators, including RPA pilots, to interact with a broadly distributed array of people and systems. This fact complicates HSI analysis as workload is not only influenced by the System of Interest (SOI) but also by the array of systems composing the SoS. In this investigation, the SOI is nominally the GCS which provides the controls necessary to command one or more RPA during one or more missions. However, the pilots' workload is influenced not only through the interface to control the RPA but additionally through communications with external systems within a larger SoS. Without the networked array of communications, the pilot might be able to aviate and navigate the RPA but will be unable to fulfill mission goals.

The RPA Pilot's mental workload was modeled during concept definition for the MAC RPA GCS using the method shown in Figure 1. As shown, one or more potential operational concepts are defined, and a system(s) architecture and

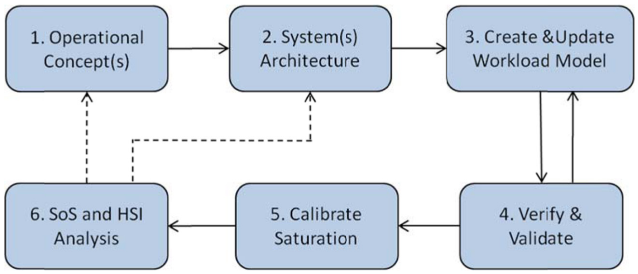


**Figure 1.** Method to investigate system-of-system and system-induced mental workload. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

design is created for each concept. This architecture ensures that the analysis is germane and appropriately scoped. A complimentary task structure is then created to enumerate the tasks for which the operator will be responsible. A simulation model is constructed to represent the architecture and task analysis, which are then verified and validated.

The resulting model produces workload values in dimensionless units. Calibration is then necessary to lend meaning to the resulting mental workload values. Analysis includes evaluating representative operational scenarios and identifying conditions resulting in unacceptable workload. The operational concept, architecture, or human task description is then modified to address conditions that result in unacceptable workload. The process is iteratively applied to arrive at an acceptable SoS concept and an SOI definition that is sufficient to meet the overall project goals. The workload model can be updated to improve the fidelity of model components as the SoS concept is refined. Each of these steps is discussed in more detail within the following sections.

## 3.1. Operational Concept

Semiautonomous remotely piloted aircraft have an increasing role in all forms of surveillance. RPA provide a significant paradigm shift from traditional manned flight, removing the man from the aircraft and permitting the pilot to remain safely in a ground station. This modification reduces the risk to the pilot and enables long endurance, as well as highly maneuverable flight. Existing RPA can remain airborne for over 24 h without refueling, and multiple pilots can control the RPA during different segments of a long endurance mission. Simultaneously, increases in automation in both manned and remotely piloted aircraft provide the potential to leave the pilot undertasked during large portions of a mission. The combination of a need for an increased number of pilots from a constrained pilot pool to man a single long endurance flight and the increased reliance on automation has suggested systems in which a single pilot is responsible for controlling multiple RPA. This concept, commonly referred to as Multi-Aircraft Control (MAC), provides the application domain for this research. This concept provides a desirable application area as the pilots are given responsibility for multiple RPA, potentially creating situations in which the human operator cannot appropriately control all of the vehicles. Among the system design considerations within the concept of MAC are questions of how many RPA can a single pilot reasonably control, which aircraft control functions will require further automation, and can changes to the SOI, i.e., the GCS, provide sufficient workload reduction to enable MAC.

## 3.2. Relevant System(s) Architecture and Task Descriptions

Relevant system architecture guides design and couples any model-based systems engineering analysis efforts to design implications. Architectural views specific to human interaction were developed according to the guidelines in the UK Ministry of Defence Architecture Framework (MoDAF) *Handbook of Human Views* [Systems Engineering & Assessment, 2009]. The HV-C, Human Interaction Structure, in Figure 2, synthesizes the information from the Operations Concept into a human-focused view that centers on the pilot and pilot interactions. This permits the pertinent information for this analysis to be collected and presented in a single comprehensive view. As shown, the pilot does not physically control components on the RPA, but interacts with the controls and displays of the Ground Control Station.

The interactions between the pilot and other personnel both within and external to the GCS may be more complex than the connection between the pilot and each RPA. The pilot has multiple means of communication with actors in several external nodes across the SoS. Changes to the system configuration in the SoS, such as increased number of RPA, can result in increased workload and interactions across the SoS. Creation of this architectural view (HV-C) prompted the addition of a refined communication component within the workload model.

Other architectural views can also be insightful. The Human View-E (HV-E), Human Functions and Tasks view, provides a general overview of pilot tasks, a structure for the model, and is useful in validation efforts. However, this view generally does not contain adequate information to enable the development of a workload model. To create the workload model, it is necessary to provide a description of the tasks the pilot will perform within the modified or future system. This task description can be derived from the allocation of steps within a UML/ SySML activity diagram if the system does not exist, or it can be derived from an existing task description for a legacy system. As the system description is evolved, updates to the architecture, model elements, or function allocation can be flowed through to the activity diagrams and to future versions of the workload model. In this example, a task analysis was previously conducted for a pilot in a legacy GCS controlling a single RPA [Eaton et al., 2006]. This study was leveraged to define the initial task structure and difficulty levels for multiple RPA control.

A typical operational scenario could include a pilot assuming control of an RPA through either a handover (i.e., the transfer of control between pilots in geographically separated GCSs) or a changeover (i.e., the transfer of control when one pilot replaces another pilot in a GCS at shift completion). The aircraft may have to transit some set of waypoints to reach the desired location. Once the RPA is at the desire location, it performs a mission—generally Intelligence, Surveillance, and Reconnaissance (ISR) or some form of offensive strike. The ISR missions may be "benign" such as loitering or flying

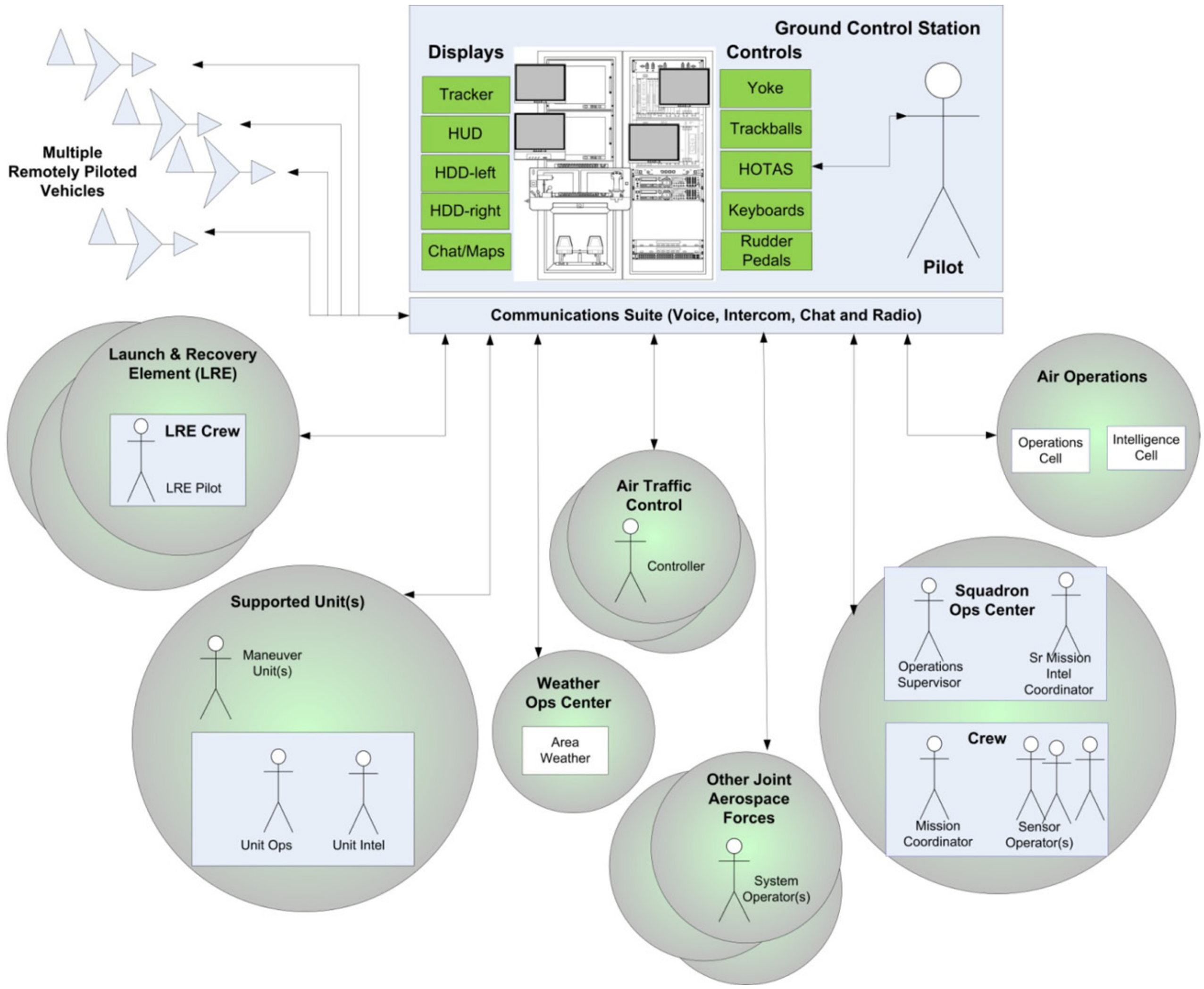


**Figure 2.** Human Interaction Structure (HV-C) for multiple RPA operations within a system-of-systems. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

a predetermined fixed route, or more "dynamic" such as surveilling a moving object through a complex terrain. If the mission is completed before the pilot's shift is over, the RPA is retasked, or it may return to base. When the shift is complete, the pilot relinquishes control through a losing changeover or handover. Each of these mission segments place different demands upon the user as the necessary tasks change within each segment. Each segment was then decomposed into a task list, which was used to construct the IMPRINT model. This decomposition of the operational concept into mission segments and tasks necessary to perform each mission segment bears a resemblance to the Task Analysis Workload (TAWL) methodology [Bierbaum, Fulford, and Hamilton, 1990], which has been previously applied by the US Army for workload estimation. The use of a discrete event simulation environment, such as IMPRINT PRO, to construct a dynamic model of the mission segments permits events and tasks to be stochastically generated, permitting the understanding of not only workload but also the variability of workload within any mission segment.

## 3.3. Workload Model

The top level of the IMPRINT PRO model is shown in Figure 3. Workload is generated in one of five blocks, including the communication block and four Aircraft Control (AC) blocks. The AC blocks are identical except for unique vehicle identification (tail number). Several mission segments are simulated in each aircraft control block:

- gaining handover
- losing handover
- changeover
- transit
- Intelligence, Surveillance and Reconnaissance (ISR)—both benign and dynamic
- strike
- emergencies (can occur anytime).

The need to model two types of ISR (benign and dynamic) was derived from discussions with active pilots. These pilots

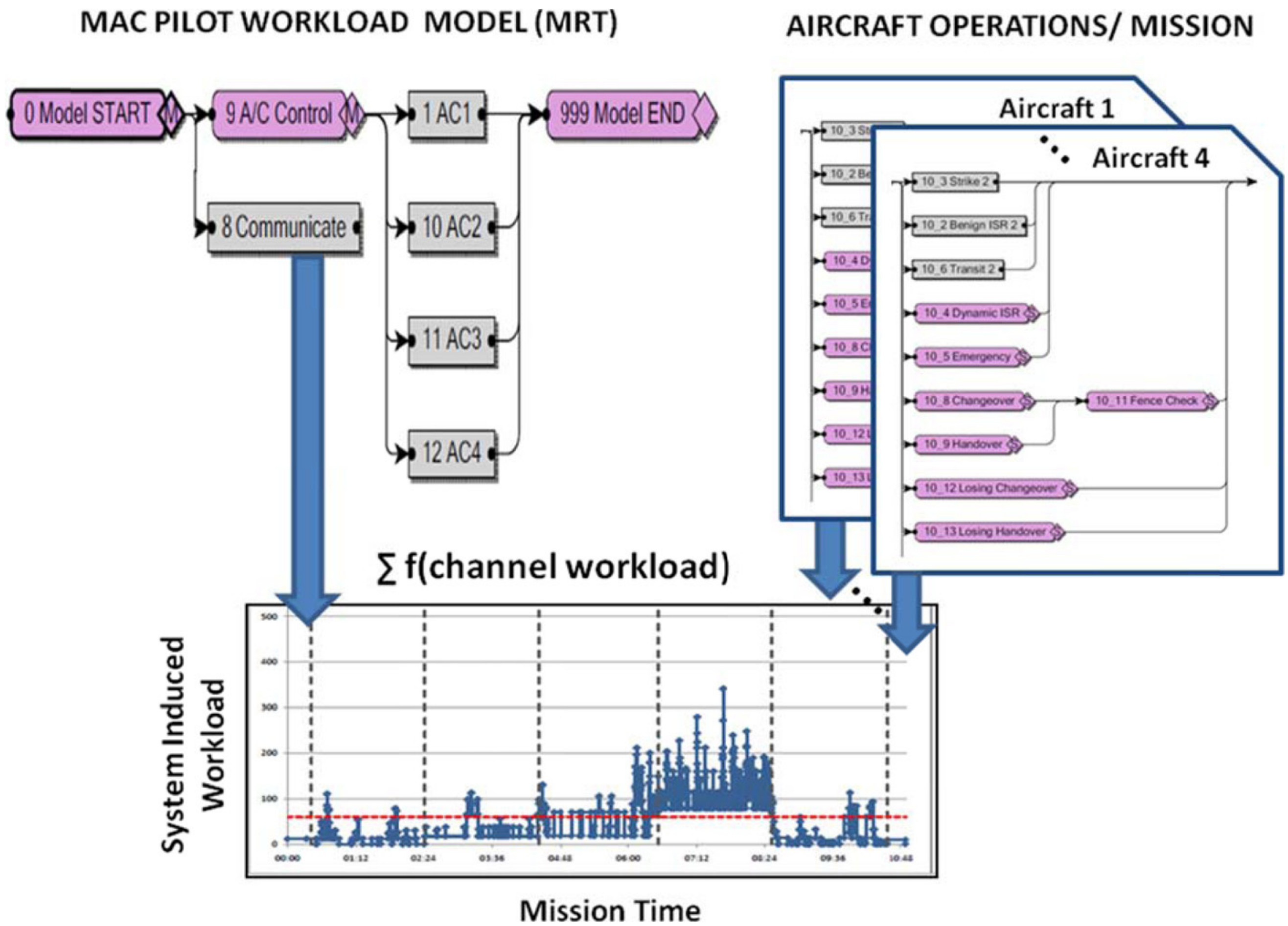


**Figure 3.** Top-level multiple aircraft control IMPRINT PRO model layout. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

observed that ISR can include preplanned navigation for area surveillance, which is a fairly low workload activity. Alternately ISR can include navigation based on sensor feedback, such as surveillance of a moving target, which creates substantially increased workload. Each of these seven missions is represented by a segment of the model that changes the pilot's goals, as well as the tasks and task frequencies that must be performed to accomplish the goals.

The model is executed through a combination of deterministic and stochastic processes. Specifically, selection of the mission mode to represent within a simulation segment and the duration of the simulation segments were controlled in a deterministic fashion to permit the analysis of workload generated by multiple RPA when each RPA was operated in a specified mission mode. Control of mission mode and simulation segment duration were deemed to be important since workload and strategies differ within each mission mode and the system concept might be modified differently for each mission mode. Therefore, each RPA was deterministically assigned a specific sequence of mission modes to execute within any given model run. During execution, most task lengths and occurrences are stochastic with distributions appropriate to the mission mode. For example, the frequency of communication events during a benign ISR mission is lower than the frequency of communication events during a dynamic ISR mission. The tasks which make up a mission modes are taken directly from the task analysis for the legacy GCS with demands and conflict values drawn from MRT and the McCracken and Aldrich Workload Demand Values, an accepted and validated rating system ranging from 0.0 to 7.0 [McCracken and Aldrich, 1984]. Times and frequencies were derived from estimated values provided by active RPA pilots. Workload is calculated at each event which produces a trace of workload over the mission time specified through the selection of submodules and their corresponding duration.

A significant decision during the development of the model was to either (1) predict the workload an operator would truly experience or (2) predict the workload the system would induce if the user could perform all of the tasks imposed by the system. Traditionally, models of mental workload have attempted to create values which would correlate to subjective ratings of workload. However, the mental workload experienced by operators will approach an asymptote as the human operator has a limited capacity and will, by necessity, begin to employ workload mitigation strategies when exposed to excessive workload levels. Mitigation strategies can include delaying, avoiding, transferring, or simplifying tasks. IMPRINT provides facilities to mimic human workload mitigation strategies and early versions of this model attempted to augment these facilities. However, two disadvantages of this approach became clear. First, sufficient research does not exist to permit one to objectively create robust features to mimic human workload strategies. Second, early experimentation demonstrated that the primary effect of implementing workload strategies is to smooth the workload function within the temporal domain to limit peak workload, which obscures the relationship between task events and the resulting workload. As a result, the ability to trace peak workload events to system-generated events that produce these peaks is diminished. For this reason, the model was constructed to calculate mental workload, assuming that the user has unlimited capacity with the goal of identifying system generated events that result in unacceptably high mental workload levels. This value, which we term system induced workload, attempts to estimate the workload that results from tasks that are imposed on the pilot by the SoS to successfully fulfill the mission, regardless of whether the pilot is capable of performing all of these tasks.

### 3.4. Model Verification and Validation

Full formal verification, validation, and accreditation of preliminary models of future or conceptual systems is extremely difficult, due to the inaccessibility or impracticality of data collection, and/or due to the cost and time demands associated with these activities. The model verification and validation step applied recommended practices [Modeling and Simulation Coordination Office, 2006] which have been successful in other similar efforts to model future systems [Wong, 2010]. The practices for this research effort included Desk Checking, Face Validation, Reviews, and Walkthroughs to insure consistency of the model logic with the architecture and task structure. Periodic reviews with Subject Matter Experts (SMEs) ensured that the model architecture and task structure, as well as the distribution parameters for the stochastic processes, were appropriate.

### 3.5. Model Calibration

The model creates mental workload values; however, the interpretation of these values to effect system design requires interpretation. Early research on workload measurement [Reid and Colle, 1988] recognized the need to establish a maximum permissible mental workload value to enable a system designer to avoid excessive mental workload. This research involved increasing task demand until a decrease in human performance was noted. Workload values were measured for this task demand and the corresponding workload value was termed a "redline." This criterion was applied to pilot ratings of workload during simulated flight to determine the acceptability of the man-machine interface [Rueb, Vidulich, and Hassoun, 1992] with the assumption that workload values at or above the redline would potentially threaten system safety due to a higher likelihood of human error. Methods for establishing this criterion or the precise definition of this criterion are not agreed upon within the human factors or psychology domains [Grier et al., 2008; de Waard, 1996]. IMPRINT PRO supports the concept of a redline as a workload limit when using workload mitigation, providing a default value of 60.

For the current application, it was desirable to establish a criterion corresponding to a mental workload value that would be challenging for the pilot to achieve and maintain. Interviews with experienced RPA pilots revealed that the tasks that are currently performed during dynamic ISR can be highly challenging from a mental workload perspective. Specifically, the pilots indicated that, during some dynamic ISR segments, they are task-saturated and find it necessary to employ workload mitigation to accomplish their goals. To establish a numerical correlate to the pilots' task saturation threshold, a model run of a 12 h long dynamic ISR was performed. The 90th percentile of the mental workload values for this run was between 58 and 59. This value is very near the default redline value of 60 in IMPRINT PRO. Therefore, a criterion of 60 workload values was chosen and applied throughout the analysis to assess the pilot's level of task saturation. This value, referred to as the "task saturation threshold," is included in subsequent plots of mental workload values as a horizontal dotted red line. Events having a mental workload value near or above 60 are considered to present the operator with more work than the pilot can perform effectively. This criterion can be applied to determine mission segments in which the workload levels are higher than this task saturation threshold and are very likely to result in workload that is beyond the operator's capacity. As such, effort during early system design can be focused to transferring, simplifying, or eliminating human tasks during these mission segments to improve the likelihood of mission success.

## 4. ANALYSIS

To evaluate the initial SoS concept, a set of operational concept model runs were structured to explore the workload imposed by the system or SoS that the pilot interacts with over the entire trade space by exploring 75 different combinations of the number of RPA per pilot and mission modes. This included every possible combination of available, physically realizable mission modes. These combinations were designed to explore the interactions and conflicts of different mission modes on workload. Typical workload graphs for a single pilot controlling two RPA are provided in Figure 4. Analogous conditions were simulated and data collected for a pilot controlling one, three, and four RPA. See the analogous graph for a pilot controlling four aircraft in Figure 5.

The sharp spikes in workload throughout these graphs indicate communication events, which are generated at different rates based on the RPA mission. The communication spikes are higher with a MAC ratio of 2 than when the pilot controls a single aircraft due to overlapping communications, producing greater conflict between the separate communication events, as well as conflict between the communication events and the other ongoing mental tasks necessary to fly the RPA.

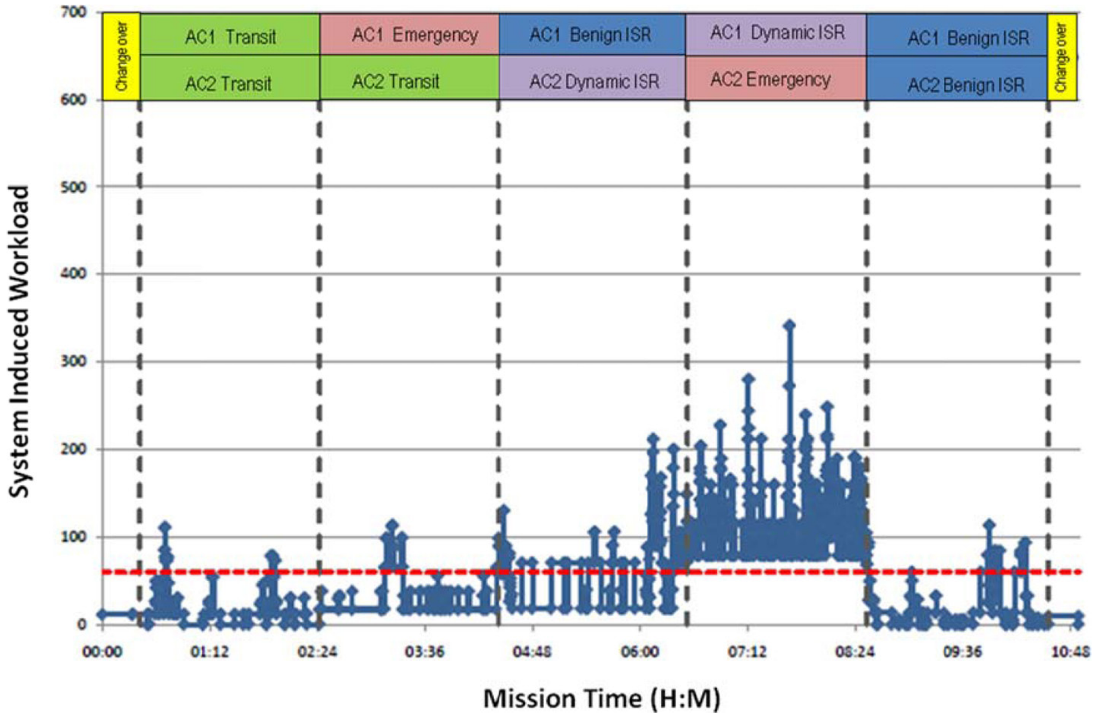


**Figure 4.** Mental workload as a function of time for a pilot controlling two aircraft. Red dashed horizontal line represents the task saturation threshold; vertical black dashed lines represent changes in mission segments. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

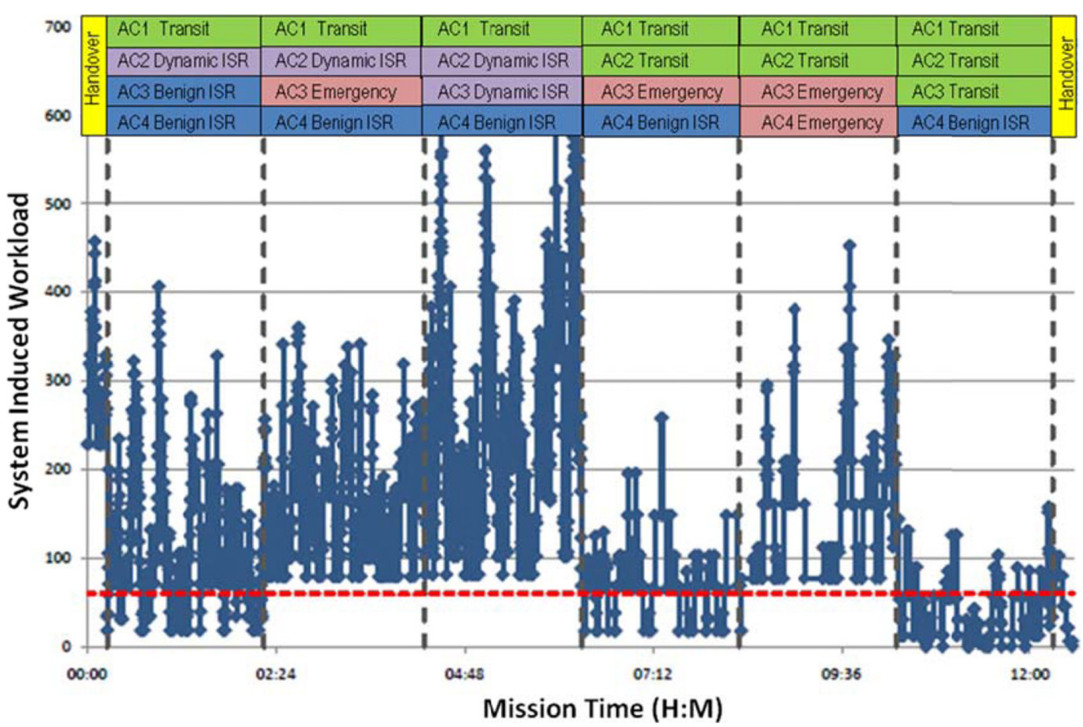


**Figure 5.** Extended workload model run with MAC ratio of 4 depicting workload as a function of time within various flight segments. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

When both RPA are operated in highly automated mission segments, which individually induce relatively low workload (i.e., transit or benign ISR), workload infrequently spikes above the saturation threshold. This suggests that workload is manageable with some minor task sequencing or workload mitigation when the RPA are all operated in these mission segments. However, in mission segments where automation cannot be relied upon, such as emergency or dynamic ISR events, the workload is frequently above the task saturation threshold for sustained periods of time. It is unlikely that workload mitigation or task sequencing will be successful in these modes and the operational concept, or system architecture, will likely require modification. Communication, often with systems external to the GCS drives the workload spikes during the dynamic events, but the primary tasks associated with controlling the RPA also conflict with one another, amplifying the workload values in the presence of communication. Due to this high workload, the current concept of operations for multi-RPA control does not appear viable during these mission segments. Further, piloting one RPA during a dynamic event and a second RPA in benign surveillance would be difficult and will increase the possibility of mission degradation.

In the fourth segment of Figure 4, one vehicle (AC1) enters dynamic ISR while the other vehicle (AC2) has an emergency. As shown, the lowest points on the workload graph are above the task saturation threshold, with spikes over four times the task saturation threshold value. Pilots in this situation would be forced to handover one of the vehicles (defer workload) to another operator, or quickly change the dynamic mission to a less workload intensive transit (loiter) or benign mission.

In comparison, Figure 5 depicts a condition including a pilot controlling four RPA, which results in a workload trace that is almost entirely at or above the saturation threshold. This plot clearly shows that under the existing operational concept and system design, and within the rules of the SoS, the control of four RPA may exceed the bounds of human performance under any scenario with even a single dynamic event. The only section of this graph that is consistently below the saturation threshold is the rightmost phase in which all of the RPA are in transit or benign mission modes. Even with all of the RPA in highly automated, low workload, mission segments, the workload trace is comparable to a single RPA during dynamic ISR and approaches the task saturation threshold. While the model indicates a MAC ratio of 4 with all RPA in benign mission modes can be successfully managed by a single pilot under the current concept of operations, sustained operations at this level is very near the task saturation threshold and in the event of an unplanned emergency or dynamic event, the pilot will be overloaded.

Based on the results of the first phase of the analysis given the current SoS operation task structure, only the highly automated transit or benign ISR conditions would appear to be amenable to a concept of operations with multiple pilots controlling an individual aircraft. However, the viability of this condition may be dependent upon operating conditions. To explore the statistics of workload, realistic shifts were simulated in which pilots receive control of an RPA, transit the RPA, perform a benign ISR with a potential short dynamic event, transit the aircraft to a new location, and finally hand control to another pilot at the end of their shift.

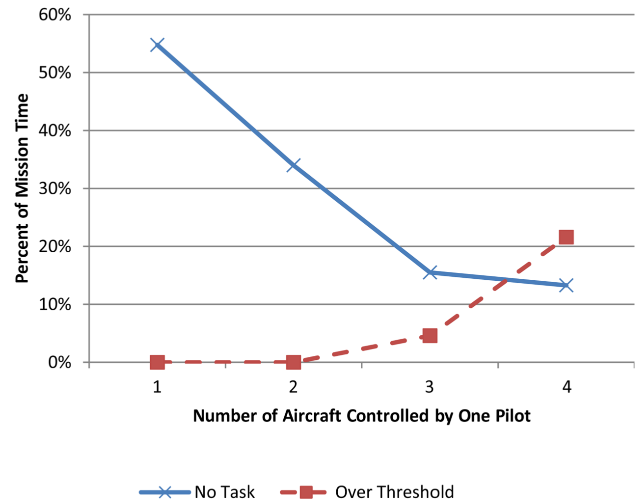


**Figure 6.** Percent of untasked time and percent time over saturation threshold as a function of the number of aircraft under control of a single pilot. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 6 characterizes the pilot untasked time and the time over the saturation threshold for RPA in benign ISR to observe the cumulative effects of increasing the MAC ratio in these highly automated conditions. As noted earlier, the belief motivating the MAC concept of operations is that the pilot's untasked time can be consumed to minimize task overlap and thus minimize the time spent over the task saturation threshold. This belief is supported when transitioning from a pilot controlling a single RPA to a pilot controlling two aircraft as a 20% decrease in untasked time is observed without increasing the time above the task saturation threshold.

However, when piloting three aircraft, time over threshold increases with further reductions in the pilot's untasked time, reducing untasked time to about 15%. This increase in the time over threshold could indicate a potential for degradation of mission effectivness if the pilot is required to control three RPA, all of which are operting in the relatively low workload benign ISR mission segments. Increasing the number of RPA beyond three further increases the percent of time over the task saturation threshold (over 20%) and results in a slightly further decrease in untasked time.

## 5. OPPORTUNITIES FOR OPERATIONAL CONCEPT MODIFICATION

Four broad observations can be made about the SoS under investigation as listed below. While these observations are made in the context of the RPA MAC concept, similar observations will likely apply to many complex systems that include multiple semiautonomous components controlled or supervised by a single operator.

*Observation 1—Multitasking:* An increase in multitasking requirements can have a substantial impact on operator performance, especially when tasks are not predictable. While the data demonstrated that the highest consistent workload values occur when a pilot is required to attend to coincident piloting tasks, sporadic communication events tended to drive extreme spikes in mental workload. Therefore, it is necessary not only to consider the workload imposed by a primary task imposed by the SOI, such as piloting an RPA, but also to consider the workload imposed by seemingly lesser tasks imposed by external systems, such as communications, and how these different types of events impact perceived workload. A modification of the system concept might explore modification of the boundary of the SOI to include aspects of the communications system. These concept modifications could involve changing the communications protocol to reduce the amount of direct person-to-person communication, utilizing additional operators to reduce the amount of communications to the pilot, or facilitating the persistence of communications that are not time critical so that they can be attended to during periods of reduced workload.

*Observation 2— Emergencies/Dynamic Events:* It is important to consider the workload imposed by demanding tasks, even if these tasks are infrequent. The results indicate that a pilot should be able to control multiple aircraft as long as these aircraft operate in transit or benign ISR conditions. Emergencies and unplanned Dynamic Events may occur infrequently during these mission segments. However, they also occur unexpectedly and the results indicate that these mission segments impose a high degree of mental workload upon the operator. The workload corresponding to an emergency or dynamic event is high enough that it will be very difficult, if not impossible, for the pilot to attend to second RPA during an emergency or dynamic event. The system concept either might be modified to redefine the tasks performed by the operator during emergencies and dynamic events to significantly reduce the workload demand during these mission segments or it will be necessary to design the system to rapidly transition a subset of the RPA to another pilot near the moment that an aircraft enters an emergency or dynamic event.

*Observation 3—Situational Awareness:* Not all task demands impose workload that is directly observable and yet mental capacity must be maintained to permit these tasks to be accomplished. As the data show when performing MAC, mental workload can be high enough that the pilot has decreased untasked time, even controlling three or more RPA in benign ISR. This fact becomes a concern when the pilot requires any untasked time to maintain situational awareness and perform mission planning, two necessary tasks that are not represented within the model. These continuous activities are not directly observable nor are they discrete (start/stop) events. Never the less, time must be allocated for the pilot to perform these unobservable activities.

*Observation 4—Automation:* Automation of the tasks that are most amenable to machine control will not necessarily reduce peak operator workload. Existing RPA systems have highly automated the benign ISR and transit operations as aviation and simple navigation operations can be performed reliably by autopilots. Automation has not been successfully applied in the circumstances that arguably present the highest workload conditions (i.e., complex aviation, navigation, and communication tasks during emergency and dynamic events). In fact, during emergency and dynamic events, the current concept does not involve any additional automation. The potential exists to modify the concept to provide aids to the pilot during these flight segments which will reduce mental workload during these mission segments to enable multiaircraft control. These aids might automate some of the lower level functions but leverage the higher level decision-making processes of the human operator.

Based upon these observations, it is now conceivable to begin a first iteration through the process shown in Figure 1. A following iteration might consider changes to the operational concept, such as the addition of operators to assume a portion of the pilots' communication or other duties, or limiting the control of multiple RPA to low-workload mission modes, such as transit or other architecture, task allocation, or automation changes.

## 6. CONCLUSION AND FUTURE RESEARCH

A method has been proposed and demonstrated for applying predictive workload modeling to the refinement of operational concepts and early system design within an SoS. Depicted in Figure 1, the method permits the integration of systems architecture, human task analysis, and workload theory to permit the assessment of human mental workload early in SoS design. Further, by calibrating the output of the model to create an understanding of the task saturation threshold, and by depicting this entity graphically with time-stamped workload values that are traceable to generated events, one can readily understand the system-generated events and configurations that produce unacceptable workload levels. This understanding is useful in targeting future development to reduce workload during the times that it would be most beneficial to the operator instead of automating the functions that are easiest to automate with the hope that this automation will reduce operator workload. This approach was exercised to evaluate multiaircraft control and was successful in identifying several concerns with this operational concept.

The human component within a complex system-of-systems must be understood as it provides greatest significant source of emergent behavior. A change to one system in the SoS, modification to support MAC in the ground control station, is expected to increase communications loads throughout the SoS directed back toward the pilot. The conflict of this new traffic with existing tasks has challenging secondary effects, which will have to be mitigated and managed, even though they were originally considered external to the SOI. Success and performance of the SoS can be significantly impacted by the workload of key operators in one of the interconnected systems. Future research is targeting how to shape, and potentially optimize, the system-imposed workload by designing the system to dynamically allocate tasks to human input and output modalities in order to reduce workload.

## REFERENCES

L. Allender, Tools for modeling human performance in systems through green-colored glasses: An army perspective, Proc XIVth

Triennial Cong Int Ergonomics Assoc 44th Annu Meet Hum Factors Ergonomics Assoc, Ergonomics for the New Millennium, July 29–August 4, 2000, San Diego, CA, pp. 717–720.

D. Beevis, R. Bost, B. Dring, E. Nordo, F. Oberman, J.P. Papin, H. Schuffel, and D. Streets, Analysis techniques for human-machine systems design, CSERIAC SOAR 99-01, Wright-Patterson AFB, OH, 1999.

C.R. Bierbaum, L.A. Fulford, and D.B. Hamilton, 1990. Task Analysis/Workload (TAWL) user's guide—version 3.0, ARI Research Product 90-15, US Army Research Institute Aviation Research and Development, Fort Rucker, AL, 1990.

D.B. Boles and L.P. Adair, Validity of the multiple resources questionnaire (MRQ), Proc Hum Factors Ergonomics Soc 45th Annu Meet, Minneapolis, MN, October 8–12, 2001, pp. 1795–1799.

H.R. Booher, "Introduction: Human systems integration," Handbook of human systems integration, H.R. Booher (Editor), Wiley, Hoboken, NJ, 2003, pp. 4–8.

M.D. Byrne and A. Kirlik. Using computational cognitive modeling to diagnose possible sources of aviation error, Int J Aviation Psychol 15(2) (2005), 135–155.

J.Y.C. Chen and C.T. Joyner. Concurrent performance of gunner's and robotics operator's tasks in a multitasking environment, Mil Psychol 21(1) (2009), 98–113.

S. Deutsch and R.W. Pew, Agent-based Modeling and Behavior Representation (AMBR), AFRL-HE-WPTR-2004-0191, BBN Labs, Cambridge, MA, 2004.

D. de Waard, The measurement of drivers' mental workload, Doctoral Dissertation, University of Groningen, Groningen, The Netherlands, 1996.

S.R. Dixon, C.D. Wickens, and D. Chang, Mission control of multiple unmanned aerial vehicles: A workload analysis, Hum Factors 47(3) (2005), 479–487.

G. Eaton, P. Geier, S. Kalita, J. Nagy, B. Palmer, A. Phillips, A. Probst, and C. Sanchez, U.S. Air Force unmanned aircraft systems performance analyses: Predator pilot Front End Analysis (FEA) report, SURVIAC-TR-2006-XXX, Survivability/Vulnerability Information Analysis Center (SURVIAC), Wright-Patterson AFB, OH, 2006.

V.J. Gawron, Human performance, workload, and situational awareness measures handbook. 2nd edition, CRC Press, Boca Raton, FL, 2008.

F. Glenn, K. Neville, J. Stokes, and J. Ryder, 2004. Validation and calibration of human performance models to support simulation-based acquisition, Proc 2004 Winter Simulation Conf, Washington, DC, December 5–8, 2004, pp. 1533–1540.

R. Grier, C. Wickens, D. Kaber, D. Strayer, D. Boehm-Davis, J.G. Trafton, and M. St John, The red-line of workload: Theory, research, and design, Hum Factors Ergonom Soc Annu Meet Proc, New York, 2008, pp. 1204–1208.

P.A. Hancock and J.K. Caird, Experimental evaluation of a model of mental workload, Human Factors 35 (3) (1993), 413–429.

H.A.H. Handley and R.J. Smillie, Human view dynamics-the NATO approach, Syst Eng 13(1) (2010), 72–79.

N. Hardman, J. Colombi, D. Jacques, and J. Miller, Human systems integration within the DOD architecture framework, IIE Annu Conf Expo 2008, Vancouver, BC, May 17–21, 2008, pp. 840–845.

S.G. Hart and L.E. Staveland, "Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research," Human Mental Workload, P.A. Hancock and N. Meshkati, (Editors), Elsevier (North-Holland), Amsterdam, 1988, pp. 139–183.

K.C. Hendy, K.P. East, and P.S.E. Farrell, "An information processing model of operator stress and performance," Stress, workload and fatigue: Theory, research, and practice, Peter A. Hancock and Paula A. Desmond (Editors), Lawrence Erlbaum Assoc. Inc., Mahwah, NJ, 2001, pp. 34–38.

D. Kahneman, Attention and effort, Prentice Hall, Englewood Cliffs NJ, 1973.

J. Keller, E. Yucesan, C.-H Chen, J.L. Snowden, and J.M. Charnes, Human performance modeling for discrete-event simulation: Workload, Proc 2002 Winter Simulation Conf, San Diego, CA, 2002, pp. 157–162.

A.M. Law (Editor), Simulation modeling and analysis, 4th edition, McGraw-Hill, Boston, 2006.

A. Law and W.D. Kelton, Simulation modeling and analysis, 3rd edition, McGraw-Hill, Boston, 2000.

J. McCracken and T. Aldrich, Analyses of selected LHX mission functions: Implications for operator workload and system automation goals, ADA232330, Anacapa Sciences, Fort Rucker, AL, 1984.

B. Mercer, Enabling executable architecture by improving the foundations of DoD architecting, Int Symp Collaborative Technol Syst, Irvine, CA, 2008, pp.558–559.

D.K. Mitchell, Successfully changing conceptual system designs using human performance modeling, Hum Syst Integration Symp, Annapolis, MD, 2009. www.navalengineers.org/publications/symposiaproceedings/2009proceedings/

D.K. Mitchell and C.L. Samms, Please don't abuse the models, Proc Hum Factors Ergonomics Soc 51st Annu Meet, Baltimore, MD, 2007, pp. 1454–1457.

D.K. Mitchell and C.L. Samms, Workload warriors: Lessons learned from a decade of mental workload prediction using human performance modeling, 53rd Hum Factors Ergonomics Soc Annu Meet, San Antonio, TX, 2009, pp. 819–823.

Modeling and Simulation Coordination Office. Verification, validation, and accreditation recommended practices guide, 2006, Alexandria, Virginia, http://vva.msco.mil/.

R.D. O'Donnell and F.T. Eggemeier, "Workload assessment methodology," Handbook of perception and human performance, K.R. Boff, L. Kaufman, and J.P. Thomas (Editors), Wiley, New York, 1986, Vol. II, pp. 42-1–42-49.

J.F. O'Hanlon, Boredom: Practical consequences and a theory, Acta Psychol 49 (1981), 53–82.

R. Parasuraman and E. Rovira, Workload modeling and workload management: Recent theoretical developments, ARL-CR-0562, US Army Research Laboratory's Human Research and Engineering Directorate, Aberdeen Proving Ground, MD, 2005.

R. Pomranky and J. Wojciechowski, Determination of mental workload during operation of multiple unmanned systems, ARL-TR-4309, US Army Research Laboratory, Aberdeen Proving Ground, MD, 2007.

T.L. Ramirez, Modeling military task performance for Army and Air Force personnel wearing chemical protective clothing, Mil Psychol 9(4) (1997), 375–394.

G.B. Reid and H.A. Colle, Critical SWAT values for predicting operator overload, Proc Hum Factors Soc 32nd Annu Meet, Anaheim, CA, 1988, pp. 1414–1418.

G.B. Reid and T.E. Nygren, "The subjective workload assessment technique: A scaling procedure for measuring mental workload,"

Human mental workload, P.A. Hancock and N. Meshkati (Editors), Elsevier (North-Holland), Amsterdam, 1988, pp. 185-218.
J. Rueb, M. Vidulich, and J. Hassoun, Establishing workload acceptability: An evaluation of a proposed KC-135 cockpit redesign, Proc Hum Factors Soc 36th Annu Meet, Atlanta, GA, 1992, pp. 17–21.
N. Stanton, P. Salmon, G. Walker, C. Baber, and D. Jenkins, Human factors methods: A practical guide for engineering and design, Ashgate, Burlington, VT, 2005.
Systems Engineering &. Assessment Ltd., The human view handbook for MODAF, Bristol, UK, 2009.
US Air Force, Headquarters, US Air Force Unmanned Aircraft Systems Flight Plan 2009–2047, Washington, DC, 2009, pp. 502–511.
L.W. Wagenhals, S.W. Liles, and A.H. Levis, Toward executable architectures to support evaluation, Int Symp Collaborative Technol Syst, Baltimore, MD, 2009, pp. 502–511.
A.T. Welford, Single channel operation in the brain, Acta Psychol 27(1) (1967), 5–21.
C.D. Wickens, Multiple resources and mental workload, Hum Factors 50(3) (2008), 479–487.
D.T. Wong, Validating human performance models of the future Orion crew exploration vehicle, 54th Annu Meet Hum Factors Ergonomics Soc, San Francisco, CA, 2010, pp. 1002–1006.
M.S. Young and N.A. Stanton, Malleable attentional resources theory: A new explanation for the effects of mental underload on performance, Hum Factors 44(3) (2002), 365–375.

John Colombi, Ph.D., is an Assistant Professor of Systems Engineering at the Air Force Institute of Technology, Wright-Patterson AFB, OH. His research interests include systems and enterprise architectural analysis, complex adaptive systems, and Human Systems Integration. Before joining the faculty, Dr. Colombi led command and control systems integration including systems engineering for the E-3 (AWACS) aircraft at Hanscom AFB, Bedford, MA. He has served at the National Security Agency developing biometric and information security technologies and analyzed communications networking at the Air Force Research Laboratory.

Michael E. Miller, Ph.D., is an Assistant Professor of Human Systems Integration at the Air Force Institute of Technology (AFIT), Wright-Patterson AFB, OH. His research interests include human systems integration, modeling and measurement of human performance, and human interface design, particularly human-display integration. Prior to joining AFIT, Dr. Miller spent more than 15 years with Eastman Kodak Company as a systems and human factors engineer. He is a member of INCOSE, a member of the Human Factors and Ergonomics Society, and a senior member of the SID.

Michael Schneider received a B.S. in Mechanical Engineering from the University of Dayton, Dayton, OH, and an M.S. in Systems Engineering from the Air Force Institute of Technology, Wright-Patterson AFB, OH, with a focus in Human Systems Integration. Prior to attending AFIT, he was an air transportability engineer for the Air Transportability and Test Loading Activity (ATTLA), in the Aeronautical Systems Center's Engineering Directorate. He ensured the safe airlift of oversized cargo to support military operations.

Major Jason McGrogan, USAF, received an M.S. in Systems Engineering from the Air Force Institute of Technology, Wright-Patterson AFB, OH. Prior to AFIT, he served as a program manager doing rapid acquisition at the Air Force Tactical Exploitation of National Capabilities. He has served as a test director for B-2 operational test and evaluation certifying new systems and capabilities and an avionics systems engineer for the C/KC-135 working on fleet sustainment.

Colonel David Long, USAF, is currently assigned as Chief of the Advanced Plans and Programs Division at the Air Force Research Laboratory. He is primarily an acquisition officer, and his previous assignments include flight test operations, aircraft maintenance, and instructing graduate school. His research interests include system architectures, unmanned aircraft systems, and human-machine interfaces.

John Plaga is a Senior Research Aerospace Engineer in the Human Performance Integration Directorate of the 711th Human Performance Wing (711 HPW/HP), Air Force Research Laboratory, Wright-Patterson AFB, OH. He is currently involved in Human Systems Integration (HSI) activities including advising the Predator/Reaper Systems Group on HSI aspects of Ground Control Station design, developing HSI tools for systems acquisition, and developing HSI recommendations for virtual game-based training tools. He received a B.S. in Aeronautical/ Astronautical Engineering from The Ohio State University, Columbus, in 1989.