EVALUATION OF AN ADVANCED FAULT MANAGEMENT SYSTEM DISPLAY FOR NEXT GENERATION CREWED SPACE VEHICLES

Miwa Hayashi, Valerie Huemer, Joel Lachter, Dorion Liston
San Jose State University
Moffett Field, CA

Steve Elkins, Fritz Renema
QSS Group, Inc.
Moffett Field, CA

Brent Beutter, Jeffrey W. McCandless, Robert S. McCann, Lilly Spirkovska
NASA Ames Research Center
Moffett Field, CA

The next generation Crew Exploration Vehicle is planned to employ Integrated Systems Health Management (ISHM) technology to enhance crew safety and improve onboard operations. For example, the ISHM could assist crewmembers with real-time fault management operations by automatically identifying the root cause of complex system malfunctions. However, to implement such a system, several human-factors issues have to be addressed. For instance, human-machine functional allocations have to be made and supporting crew interfaces designed and evaluated. The paper describes a concept for crew-ISHM interactions called the Fault Management Support System (FAMSS) that addresses these human-factors issues. Simulator experiment results showed that a simulated FAMSS interface improved operators' situation awareness and fault-management performance while decreasing fault-management workload.

INTRODUCTION

Background

To achieve the goals set by the President's Vision for Space Exploration address in 2004, NASA is currently developing a Crew Exploration Vehicle (CEV) that transports astronauts to low Earth orbit, the Moon, and eventually Mars. During the development, CEV designers have an opportunity to infuse state-of-the-art artificial intelligence technology known as Integrated Systems Health Management (ISHM) to greatly increase the level of crew safety, improve crew performance, and reduce operations costs compared to those of the Space Shuttle (Exploration Systems Architecture Study Report, NASA, 2005).

For example, ISHM technologies automatically perform root-cause identification of system malfunctions from the clusters of off-nominal sensor inputs, using sophisticated pattern recognition and model-based reasoning techniques, and annunciate only the caution and warning (C&W) events associated with the root cause in the cockpit, inhibiting all other C&W events from the downstream subsystems. Furthermore, assuming a requisite level of integrated avionics, ISHM could automatically determine the appropriate sequence of procedures to handle the problem, execute the procedures, and verify the results. Many of the ISHM technologies needed to support such an "end-to-end" automated fault management process have already been implemented on unmanned spacecraft. However, for a crewed space vehicle, integration of the ISHM technologies raises important human-factors challenges, such as finding proper functional allocations between crewmembers and automation that improve the crewmembers' situation awareness (SA) and performance while minimizing their workload. Appropriate cockpit interfaces to support the given functional allocations also need to be provided.

In the present paper, first the human-factors issues in today's Space Shuttle fault management processes are examined to provide us lessons learned. Then, we propose a concept of a new fault management interface that addresses these human-factors issues utilizing ISHM technologies. Since the CEV is still under development and many system details are still fluid, our goal here is to derive a generic concept that is independent from the underlying vehicle system. A prototype interface design to support this concept was implemented in a Space Shuttle cockpit simulator, and a human-in-the-loop evaluation was conducted to evaluate the usability of the prototype interface. We report some of the results of the evaluation, followed by a discussion.

Human-Factors Issues in the Current Space Shuttle Fault Management Processes

Many fault management procedures of the Space Shuttle roughly break down into the following three steps. For each step, elements that cause human-factors issues are described.

1) Fault detection and identification. Detect visual/aural alarms when one or more parameters exceed preset limit values. Then, read the fault messages to identify the affected subsystem. The fault messages tell which page of the Flight Data Files (FDF) to read for the instructions (see the next paragraph). One of the human-factors concerns for this step is that, because of the highly interconnected nature of the subsystems, a failure of one component may trigger a cascade of alarms of all the related components. This can obscure the

root cause, leading to crewmember's distraction and/or impaired SA (McCandless, McCann, & Hilty, 2003).

2) FDF look-up. Based on the fault message, locate the appropriate fault management instructions in a paper FDF. The instructions are often written in an "If-Then-Else" style that leads to different branches through the instructions. Evaluate these conditionals and follow the proper logical path. Such FDF navigation is inherently complex. Therefore, a human-factors concern is that the FDF navigation could take up a significant portion of the crewmember's attentional and cognitive resources. Furthermore, from a purely psychomotor perspective, accessing the information in a paper FDF could be already problematic, particularly during ascent and entry phases when crewmembers are suited, restrained, and wearing helmets that restrict their field of view, and the cockpit is vibrating.

3) Manual Switch Throws. Locate the switches and manually execute the switch throws as instructed in the FDF. An obvious human-factors problem is that hundreds of similar switches are densely located in the cockpit; thus, remembering the location of each switch, locating it, and manually toggling it all demand additional attentional and cognitive resources. Moreover, the manual switch throw itself could cause potential "slips" (Reason, 1990), errors in executing a motor command that lead to an unintended action, such as inadvertently leaving the switch in the wrong position.

Concept for New Fault Management Interface

We propose a concept for an advanced fault management system, or Fault Management Support System (FAMSS), that assumes access to the new ISHM technologies and provides an interface that addresses the human-factors issues in the three steps described above.

For step 1), ISHM automatically identifies the root cause of the problem, eliminating the crewmembers' need to assess the root cause by themselves and, in turn, reducing potential distraction and improving the crewmembers' SA. For step 2), ISHM automatically retrieves the FDF instructions for the subsystem affected by the root cause and evaluates all the logical conditionals involved in the instructions. FAMSS, then, presents the appropriate sequence of procedures on its fault management (FM) display. That means, the crewmembers no longer need to flip through the paper FDF and evaluate the logic conditionals by themselves. (However, if they want to, they can still call up an electronic version of the FDF by pushing an "FDF" button on the FM display.) Right above the switch-throw instruction texts, the FM display shows schematic information of the affected subsystem to assist the operator in quickly understanding which subsystem has failed. Finally, for step 3), the crewmember pushes an "Accept" button on the FM display, provided next to each separate switch-throw instruction text, to permit FAMSS to automatically perform the corresponding switch throw. FAMSS will not execute the switch throw until a crewmember has given permission, so that the crewmember is always "in the loop." Automated switch throws can free up attentional and cognitive resources originally required for locating and throwing the switch, as well as prevent the possibility of slips during manual switch throws. If so desired, however, the crewmember is also allowed to throw the switches manually.

Simulator Evaluation

We designed a prototype FAMSS FM display and evaluated its usability by performing a human-in-the-loop simulation. In addition to the traditional measures such as performance accuracy, SA scores, and workload scores, the participants' eye-movement data were also measured to examine usage of the FM display. A Space Shuttle simulator was used in this experiment, and the Space Shuttle Cockpit Avionics Upgrade (CAU) display suite (Hayashi, Huemer, Renema, Elkins, McCandless, & McCann, 2005) was used as the baseline condition for the evaluation. The "FAMSS" condition, which was evaluated against the baseline, was nearly identical to the baseline condition except that the FAMSS FM display replaced the CAU C&W Annunciator light panel, which w rarely used by crewmembers. Hence, the differences in performance, etc., between the FAMSS and baseline conditions were most likely due to the presence or absence of the FM display interface.

METHOD

Simulator

The experiment was conducted in a fixed-base part-task Space Shuttle cockpit simulator. The simulator emulates key cockpit displays and switch panels accessible to the left-seat crewmember. The simulator contains 12 touch-panel LCD monitors, of which four 20-inch monitors represent the forward cockpit displays, seven 20-inch monitors represent the side and overhead panels, and a 12-inch monitor represents the keyboard. An audio system provides background engine noise and alarm annunciation. The entire system is driven by a distributed, multi-platform (SGI and PC) set of computers.

Display

Figure 1 shows the eight forward displays of the baseline condition. In the FAMSS condition, the FM display replaced the C&W Annunciator light panel. (The FM display was the same size as the other square-shaped displays, and, thus, in the FAMSS condition, the Fault Sum was shifted down to make room for the FM display).

In the FAMSS condition, the FM display normally indicated "SYSTEMS NOMINAL." When a system malfunction was detected, a page specific to the malfunction appeared here (see Figure 2 for example pages; brief descriptions of each malfunction are in the next section). If another malfunction occurred, a new tab appeared on the right side of the display indicating there was another page to view. A new malfunction page did not automatically replace the current page, because the operator might still be using the current page.

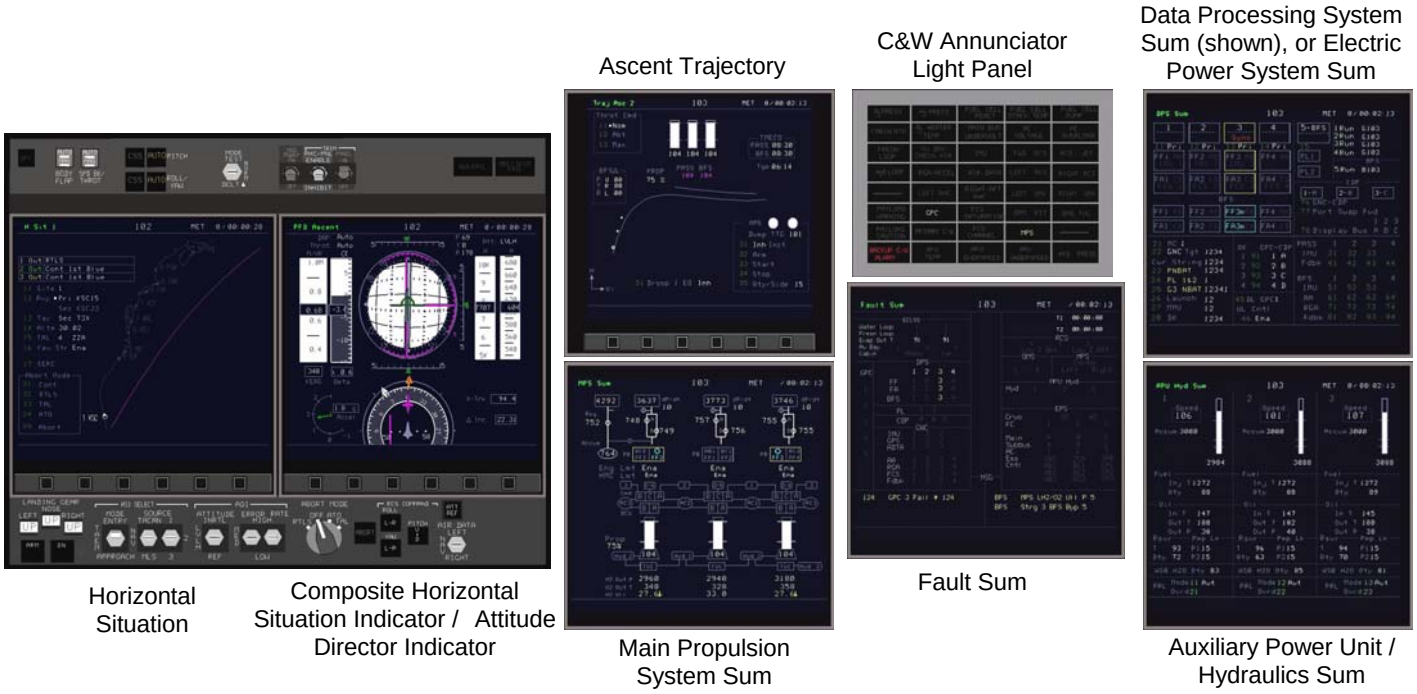

Figure 1. Eight front displays presented in baseline condition.

(a) Electrical subbus failure (b) GPC fail-to-sync failure (c) Non-isolatable helium leak

Figure 2. Examples of FAMSS FM display pages.

A magenta "Accept" button appeared next to a switch-throw instruction if the procedure was ready to be executed. If a procedure required, for instance, a wait until a variable reached to a certain value, a grey-colored countdown timer bar appeared in place of the "Accept" button, displaying the estimated remaining time (see Figure 2(c), lower right corner).

Scenarios

The Space Shuttle ascent-phase operation from launch to Main Engine Cutoff was simulated (about 8 min 30 sec). One nominal and two off-nominal scenarios were used in the experiment. During the nominal trials, no simulated system malfunctions occurred.

During the off-nominal trials, either one or three simulated system malfunctions were inserted. The single-malfunction (or *single-mal*) scenario contained a helium leak in the helium supply system of one of the three Main Engines at 1:50 Mission Elapsed Time (MET), where the leak location could be isolated by opening and closing the isolation valves for the two redundant helium supply lines in the proper order, such that both valves are not closed at the same time. Thus, this malfunction is referred to as the *isolatable* helium leak. The valve operations could be performed by a series of switch throws or a series of "Accept" button pushes on the FM display (the latter available in the FAMSS condition only).

The multiple-malfunction (or *multi-mal*) scenario contained three malfunctions. First, an electrical subbus failed at 0:30 MET, which generated multiple alarms for the related subsystems. Thus, in the baseline condition, the participants had to identify the root cause from the patterns of the alarms. In the FAMSS condition, the FM display indicated the root cause. The participants also had to remember that this malfunction left one of the helium isolation valves *failed*

closed in order to solve the third malfunction correctly (described below). Once the root cause was identified, there was no further operator action required for this malfunction. Second, one of the five General Purpose Computers (GPCs) went out-of-sync with the others at 0:50 MET and had to be brought to a halt by either a series of switch throws or a series of "Accept" button pushes (in the FAMSS condition only). Third, a leak in the helium supply system for one of the Main Engines occurred at 3:00 MET. The difference from the case in the single-mal trial was that one of the isolation valves was already *failed closed*; thus, this leak was *non-isolatable*. If the operator inadvertently attempted to isolate the leak location, he or she would close the isolation valves on both of the redundant helium supply lines at the same time, resulting in premature shutdown of the engine and possible mission abort. The correct procedures were to open the interconnect valve from the pneumatic helium tank to the affected helium tank when the affected tank pressure fell under 1150 psi, and/or shut down the affected engine when the vehicle velocity reached 23,000 ft/sec. Again, both procedures could be performed by switch throws or "Accept" button pushes (in the FAMSS condition only).

Participants

Fourteen Air Transport Pilots with an average of approximately 16,000 total flight hours participated in the experiment. All participants had experience participating in previous simulator experiments involving the CAU display formats (the baseline condition in the current experiment) (e.g., Hayashi et al., 2005). In addition, a one-day classroom lecture and a two-hour simulator training session were given to participants prior to the current experiment to refresh their memory about CAU operations and provide instructions in the use of the new FM display.

Data Collection

Each participant performed four trials in one cockpit condition (baseline or FAMSS), followed by four trials in the other cockpit condition. The cockpit condition order was counterbalanced between participants. Each block of four trials consisted of two nominal, one single-mal, and one multi-mal trial. The first and third trials were always off-nominal trials, and the second and fourth were always nominal. The order in which the single- and multi-mal trials were presented was counterbalanced between and within the participants.

The participants' eye-movement data were collected with a head-mounted eye camera (ISCAN ETL-500, ISCAN, Inc., Burlington, MA) and a head tracker (FasTRAK, Polhemus, Colchester, VT) at a sampling rate of 60 Hz. Their switch-throw and FM-display button-pushing activities were also recorded. After each trial, the participants rated their subjective workload levels using the Bedford workload scale (Roscoe & Ellis, 1990). In addition, after off-nominal trials, the participants rated their SA in terms of their ability to *diagnose* the malfunctions, and also in terms of their ability to *resolve* the malfunctions, on a continuous scale from 0 to 10.

RESULTS

In this section, only the data from off-nominal trials are discussed because the major baseline-FAMSS differences occur only on off-nominal trials.

Performance Accuracy

The numbers of participants who correctly completed the malfunction management procedures in each cockpit were as follows: for the isolatable helium leak, eight (8/14 = 57%) in the baseline condition, and 14 (100%) in the FAMSS condition; for the GPC fail-to-sync, eight (57%) in the baseline, and 12 (86%) in the FAMSS; for the non-isolatable helium leak, seven (50%) in the baseline, and 13 (93%) in the FAMSS. (The electrical subbus failure did not require any operator action.) Thus, for all three malfunctions, more participants successfully completed the procedures in the FAMSS condition than in the baseline condition ($p < 0.05$ for all three malfunctions assuming binomial distribution).

Subjective Ratings: SA and Workload

Three-way repeated-measures ANOVA with Cockpit Condition (FAMSS vs. baseline) and Malfunction Complexity (single- vs. multi-mal) as within-subject effects, and Order of Cockpit Presentation (baseline first vs. FAMSS first) as a between-subject effect was applied to the Bedford workload scales. The results showed that the participants reported significantly less workload in the FAMSS condition than in the baseline condition ($F(1,12) = 16.3$, $p < 0.01$), and also in the single-mal trials than in the multi-mal trials ($F(1,12) = 19.7$, $p < 0.01$). No other significant main or interaction effect was found.

The analogous ANOVA results on the SA scores showed that participants found it to be significantly easier to *diagnose* malfunctions in the FAMSS condition than in the baseline condition ($F(1,12) = 29.8$, $p < 0.01$), and also in the single-mal trials than in the multi-mal trials ($F(1,12) = 4.9$, $p < 0.05$). Similarly, they found it to be significantly easier to *resolve* malfunctions in the FAMSS condition than in the baseline condition ($F(1,12) = 31.4$, $p < 0.01$), and also in the single-mal trials than in the multi-mal trials ($F(1,12) = 9.4$, $p < 0.01$). No other significant main or interaction effect was found.

Malfunction Resolution Times and FM Display Usage

The eight participants who correctly completed the isolatable helium leak procedures in *both* cockpit conditions resulted in much shorter malfunction resolution time (RT) (i.e., the time from the Master Alarm to the completion of the last procedure step) in the FAMSS condition (mean = 38 sec) than in the baseline condition (mean = 128 sec). A two-way repeated-measures ANOVA with Cockpit Condition as within-subject effect and Order of Cockpit Presentation as a between-subject effect revealed the difference to be statistically significant ($F(1,6) = 8.8$, $p < 0.05$). The eye-movement data of the six of the eight participants were

subjected to an analogous ANOVA (two participants' data were excluded due to a large amount of missing eye-movement data during this malfunction). The results showed that the total fixation duration times on the FM display in the FAMSS condition were significantly shorter (mean = 20 sec) than those on the FDF and the fault message (necessary information to look up the correct FDF page) combined together in the baseline condition (mean = 35 sec) during the isolatable helium leak malfunction ($F(1,4) = 32.3$, $p < 0.01$). This suggests that the shorter total fixation durations on the FM display likely contributed to the shorter isolatable helium leak malfunction RT in the FAMSS condition.

The GPC fail-to-sync malfunction RTs tended to be shorter in the FAMSS condition than in the baseline condition (mean = 41 sec in the FAMSS, 60 sec in the baseline among the five participants who correctly completed the GPC fail-to-sync procedures in *both* cockpit conditions), but the difference did not reach statistical significance.

The cockpit condition did not affect the RTs for the non-isolatable helium leak as much (mean = 296 sec in the FAMSS, 303 sec in the baseline, among the six participants who completed the procedures correctly in *both* cockpit conditions) because the two long waits for the tank pressure and the vehicle velocity, respectively, to reach certain values compensated for the delay in the malfunction management processes in the baseline condition, resulting in similar RTs in the two cockpit conditions. Interestingly, however, their eye-movement data revealed that the total fixation durations on the FM display in the FAMSS condition were significantly longer (mean = 84 sec) than those on the FDF and the fault message combined together in the baseline condition (mean = 64 sec) during the malfunction ($F(1,4) = 32.3$, $p < 0.01$). The result suggests that the participants may have had attention capture on the FM display, possibly due to the countdown timer bar presenting the remaining time during the waits.

DISCUSSION

The results demonstrate that the FM display presented in the FAMSS condition generally improved the participants' SA, performance accuracy, and malfunction resolution speed while reducing the workload. Besides these positive effects, we also found some potential room for improvement in the FM display design. For instance, the eye-movement data suggested that the countdown timer bar on a FM display page might have caused attention capture. A possible remedy for this is to use an auditory alarm to notify the operator when the critical time is approaching, so that they do not have to be staring at the countdown timer bar. Also, two participants who failed to work on the GPC fail-to-sync in the FAMSS condition never switched to the GPC page. The current design of the FM display does not bring up a new malfunction page automatically over the current page so that the new page would not interrupt the operator's current malfunction management task. However, this design could be modified with a simple prioritization algorithm that allows a new page to be brought up over the current page if the operator is not working on any malfunction.

The level of automation (LOA) of the FAMSS system could be adjusted as well. For instance, on Sheridan & Verplank's LOA scale (1978), the current FAMSS system is level 5 ("computer executes if the human approves"). This level worked well in the malfunction situations evaluated in the present study. However, if the operators' workload were too high, then moving up to level 6 ("the computer allows the human a restricted time to veto before automatic execution") or higher with a potential risk of degrading the operators' SA might be an option. Or, if it becomes more critical to keep the operators "in the loop" all the time, then reverting to level 4 ("the computer suggests one alternative") or lower with the price of increased operator workload may be an option.

CONCLUSIONS

FAMSS, a new concept for a real-time fault management interface for the CEV cockpit, was proposed based on the review of the human-factors issues observed in the current Space Shuttle fault management processes. A human-in-the-loop simulator experiment study demonstrated the advantages of the FM display of the FAMSS, as well as some potential improvements of the prototype FM display format design. As more specifics of the CEV system fault management procedures are determined, the concept and interface design of the FAMSS should be re-examined and refined. The present study provides a starting point for these future efforts.

ACKNOWLEDGEMENT

This work was supported by the Engineering for Complex Systems (Resilient Systems and Operations) Program and by Grant #01-OPBR-07-0000-0164 from the NASA Space Human Factors Engineering project. The authors are grateful to Bruce Hilty of NASA Johnson Space Center, Dr. Leland Stone of NASA Ames, and Capt. Bob Lawrence of Battelle for their support. The authors also express great appreciation to the fourteen pilots who participated in the study.

REFERENCES

Hayashi, M., Huemer, V., Renema, F., Elkins, S., McCandless, J. W., & McCann, R. S. (2005, Sep. 26-30). *Effects of the space shuttle cockpit avionics upgrade on crewmember performance and situation awareness.* Paper presented at the Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL.

McCandless, J. W., McCann, R. S., & Hilty, B. R. (2003, Oct. 13-17). *Upgrades to the Caution and Warning System of the Space Shuttle.* Paper presented at the Human Factors and Ergonomics Society 47th Annual Meeting, Denver, CO.

NASA. (2005). *NASA's Exploration Systems Architecture Study* (No. NASA-TM-2005-214062).

Reason, J. (1990). *Human Error*. Cambridge, UK: Cambridge University Press.

Roscoe, A. H., & Ellis, G. A. (1990). *A subjective rating scale for assessing pilot workload in flight* (No. TR90019). Farnborough, UK: Royal Aeronautical Establishment.

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