PREDICTIVE INFORMATION: STATUS OR ALERT INFORMATION?

Anna Trujillo, NASA Langley Research Center, Hampton, VA
Dr. Daniel Bruneau, Georgia Institute of Technology, Atlanta, GA
Hayes Press, Lockheed Martin, Hampton, VA

Abstract

Previous research investigating the efficacy of predictive information for detecting and diagnosing aircraft system failures found that subjects like to have predictive information concerning when a parameter would reach an alert range. This research focused on where the predictive information should be located, whether the information should be more closely associated with the parameter information or with the alert information. Each subject saw 3 forms of predictive information: (1) none, (2) a predictive alert message, and (3) predictive information on the status display. Generally, subjects performed better and preferred to have predictive information available although the difference between status and alert predictive information was minimal. Overall, for detection and recalling what happened, status predictive information is best; however for diagnosis, alert predictive information holds a slight edge.

Introduction

In the aviation community, the early detection of a possible aircraft system failure may increase the safety of flight by allowing flight crews to rectify the problem before an alert occurs. One method to aid early detection of a problem is to present information on the predicted state of the system. Previous research found that subjects liked to have predictive information indicating when a parameter would reach an alert range [1-7]. This earlier research focused on how predictive information affected pilot decision making because some of the benefits of predictive information are in the realm of improved decision making [8-10]. Previous research conducted by Bartolone and Trujillo [11] and Trujillo [1-3] examined the format predictive information should take; in particular, the update rates and whether pilots wanted to know the predicted time to an alert or whether they preferred to know when a parameter was abnormally increasing or decreasing in value.

Objectives

The primary objective of this experiment was to understand whether the predictive information should be located with the parameter information or with the alert information. It was postulated that because a parameter had yet to reach an alert range, subjects might prefer it to be associated more closely with the affected parameter. Further, because no procedure was associated with the predictive information, it was acting as, and thus depicted as, status information. On the other hand, a parameter was trending towards an alert range, so some subjects might prefer it to be associated with an alert, and therefore the alert's associated corrective procedure. This is apparent for subjects wanting to take proactive actions [5]. Therefore, in this experiment subjects saw both forms of predictive information in order to obtain objective and subjective preferences.

Method

Subjects

Twelve people participated as subjects. Six were certificated pilots with a current Private Pilot license [12]. The rest of the subjects were non-pilots. The average age of the pilots was 47 years and the average age of the non-pilots was 42 years. The pilots had an average of 16 years experience and an average of 780 hrs of flight experience.

Independent Variable – Predictive Information

Each subject saw 3 forms of predictive information: (1) none, (2) predictive alert, and (3) predictive status. The predictive information was an estimate of when an alert range would be reached given the current system configuration. The prediction was in the format of minutes and tens of second (*e.g.*, 1:20). All predictions were accurate for this evaluation *i.e.*, if the predictive time to an alert was 1 min, then in 1 min the affected parameter would reach an alert range for the given system configuration.

Predictive Alert

For the predictive alert, the time to an alert range was at the end of the appropriate alert message. The time to an alert message was presented as an advisory alert (cyan in color) because the alert range had yet to be reached (Figure 1).

FF ADP TEMP 1:20

Figure 1. Predictive Alert

Predictive Status

For the predictive status, the time to an alert range was at the beginning of the alert range for the corresponding parameter display (Figure 2).

Figure 2. Predictive Status

Independent Variable – Display Format

Subjects also saw two display formats, the baseline display and one of the collocated displays. The collocated displays were the Dial-on-Control (DoC) and MultiDimensional Object (MDO) display formats [13]. The collocated displays were designed so that all three types of information were located on one screen [11, 14-16]. All the display formats modeled the same 3 systems – fuel system, power plant, and heat exchanger.

Baseline Display

On the baseline display format, status information was separate from the alert/procedure and control screens. Status information was presented with standard dial formats whose normalized area of parameter movement was 1. When all the parameters were at their expected values, the dial pointers were horizontal (Figure 3). This aspect of the display encouraged check reading because pattern matching could be employed [8]; any parameter deviation displayed a dial pointer departing from horizontal, which entailed large parameter movement. Therefore, this display incorporated pattern matching and large parameter movement.

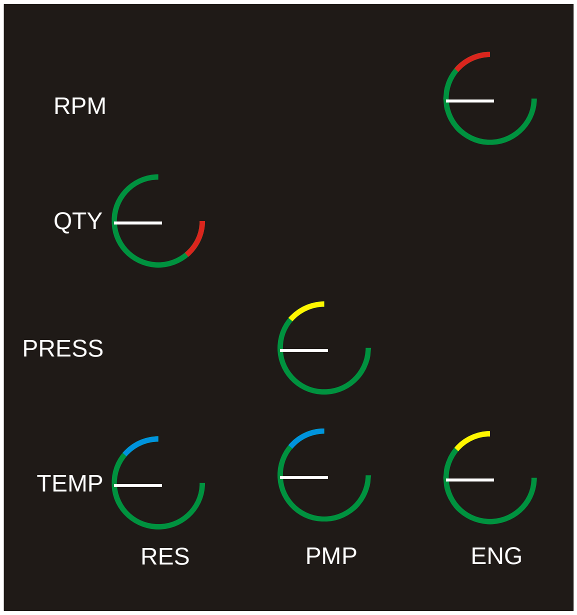

Figure 3. Baseline Display Dial Format

The control screen mimicked the functional layout of the generic system (Figure 4). Components with no change of state, such as the RES, were shown with white squares. Components that could change state (*i.e.*, turn on and off), such as the PMP, were shown with circles. A single outlined circle indicated a component that was on

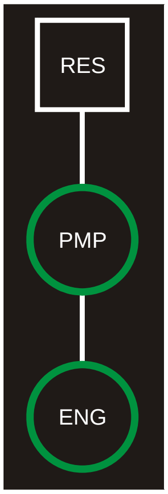

Figure 4. Baseline Display Control Screen

while a double circle denoted a component that off. The outline color of the component depicted the highest alert range of the component's parameter values. A failed component was shown with a red outline and a red X.

Dial-on-Control (DoC) Display

The dial-on-control format was a collocated display with parameter information integrated into the control display (Figure 5). It had a normalized area of parameter movement of 0.4. This display shared some of the conventions employed in the baseline display. Components with no change of state were displayed as square while components that could change state were displayed as circles. Also, a single outlined circle indicated a component that was "on" while a double circle designated a component that was "off."

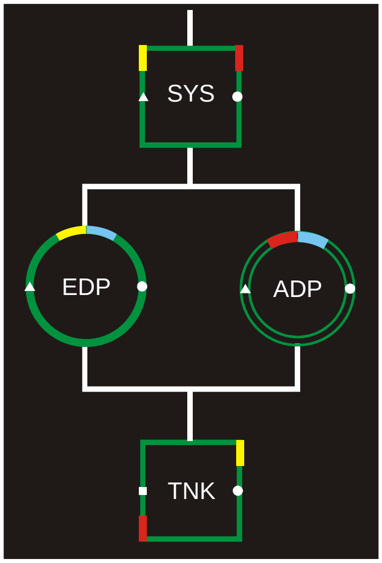

Figure 5. DoC Display

Each component symbol was split in half vertically. The left half of the component registered either pressure or quantity while the right half of the component indicated temperature. Pressure was shown with a triangle icon, quantity with a rectangle, and temperature with a circle. The icons that indicated the parameter values traveled around the component outline. When all the parameters were at their expected values, the icons were displayed at the horizontal middle of the component outline. Therefore, this display incorporated collocation and pattern matching but with limited movement.

The appropriately color-coded alert range was indicated at either the top or bottom of the component outline. The rest of the outline of the component symbol, not including the alert ranges, was green.

When a parameter reached an alert range, the icon changed from white to black and the component name was displayed in the same color as the component's parameters highest alert classification reached; otherwise the component's name was displayed in white. A failed component was displayed with a red X through the component and the component's name was displayed in red, indicating a warning.

Multi-Dimensional Object (MDO) Display

As with the DoC display, the MDO display collocated the parameter information with the control display but the parameter information was integrated pictorially into the control display [17]; pressure was shown by size, temperature was indicated by fill color, and quantity was depicted by fill level. Therefore, this display supported collocation with no pattern matching because subjects were unfamiliar with this display (Figure 6) but it did have large parameter movement – a normalized area of parameter movement of 2. The additional incorporation of the parameter information was thought to enhance visual processing of the display in a glance such as was found with polar-star displays [14]. As with the

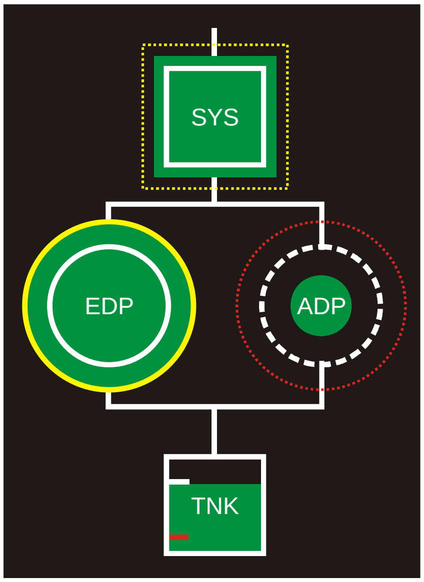

Figure 6. MDO Display

other two displays, components with no change of state were square while components that could change state were circles. For the components with a change of state, a solid white outline indicated a component that was “on” while a thick, long-dash white outline indicated a component that was “off.” A failed component was shown with a red X through it.

Pressure was indicated by size. When pressure increased, the amount of the component symbol’s filled area grew proportionally. When pressure decreased, the colored fill shrank proportionally. The beginning of a pressure alert range was shown with a dotted colored outline indicating the alert level. When the pressure alert range was reached, the dotted colored outline turned to a solid red, amber, or cyan indicating the alert level and the component name turned black in color.

Temperature was indicated by fill color. When the temperature increased, the fill color changed from green to the alert range color from the center out. When the temperature decreased, the fill color changed from green to the alert range color from the outside in. The beginning of the high temperature alert range was indicated by the outside edge of the colored component fill and the beginning of a low temperature alert range was the center of the colored component fill. When a high temperature alert were reached, the fill color was displayed in the same as the alert range color with a dotted green outline at the edge. When a low temperature alert range were reached, the fill color was displayed in the same as the alert range color with a small black circle in the middle. Also, the component name was displayed in black.

Quantity was indicated by fill level. When the quantity increased, the fill level rose and when the quantity decreased, the fill level fell. A small white horizontal line on the side of the component outline indicated normal fill level. A small color-coded line on the side of the component outline showed the beginning of an alert range. This color-coded line was the same color as the alert category. When an alert range was reached, the component name turned black and the top of the fill level changed to the color coded alert range.

Dependent Variables

The dependent variables consisted of the subjects’ ability to track a randomly moving object, to detect, diagnose and mitigate the system failure, and their recollection of the problem. Also measured was the subject’s accuracy in indicating the system with the failure and the subject’s diagnostic accuracy of the component parameter affected.

At the end of each display format, subjects completed the NASA-TLX workload measure questionnaire [18, 19] and a Cooper-Harper (CH) controllability scale rating [20, 21].

At the end of the experiment, subjects completed a final questionnaire. This questionnaire asked subjects to rate on a continuous scale how easy it was to determine system status with the collocated and predictive displays. The questionnaire also asked for subject preferences, and likes and dislikes by display type.

Procedure

When subjects first arrived, they signed a consent form before being given a verbal briefing on the experiment tasks.

Subjects then moved to the simulator where they completed two practice runs with the first display format. After the practice runs, subjects completed 12 data runs. During each run, subjects had to keep a randomly moving target centered using a left-handed joystick. They also had to monitor for a single failure that would occur in one of the systems. Once they identified the system with the failure, subjects then corrected the failure by following a checklist. At the end of each run, subjects answered questions about what failure occurred, and complete the NASA-TLX and CH controllability rating scale. At the end of the 12 data runs with the first display, subjects completed two practice runs with the second display and then the 12 data runs with that display.

At the end of the simulation runs and questions, subjects completed the final questionnaire.

Apparatus

The simulation ran on two SGI Crimson computers with IRIX 5.2. The redraw refresh rate and graphics update was 30 Hz.

The tracking task was on a 27-inch CRT screen in the upper middle of the layout. For the baseline display configuration, the dials were on the left screen below the tracking task display, the alerts and checklists were on the right screen below the tracking task, and the controls were on the right-most screen. For the collocated displays, the display was on the middle screen below the tracking task. The questions asked after each run were presented on the middle screen below the tracking task.

These three touch screens were 27-inch CRT screens with an Elo Touchsystems IntelliTouch overlay for touch-screen capability. The Saitek Cyborg evo joystick was on the subject's left side.

The tracking task was modeled in STAGE v9.2 from Presagis. VAPS v6.4.1, also from Presagis, was used for generating the displays.

Data Analysis

Data was analyzed using SPSS® for Windows v15. The data was analyzed using a 3-way ANOVA with predictive information, display type, and pilot status (pilot vs. non-pilot) as the independent variables.

The time to detect, diagnose, and mitigate the system failure was defined as the difference in time between when the failure started and when subjects started remedial action. The subject's diagnostic accuracy of the system with the failure was graded by comparing their answer to the correct answer. The CH ratings were on an integer scale. For the NASA-TLX ratings, the six dimensions of workload were normalized to a 100-point scale and then averaged together. The final questionnaire was on a continuous 100-point scale.

Results

Time to Act

Not surprisingly, subjects noticed failures earlier – that is, they identified the system with the failure earlier – with the predictive information ($F_{(2, 17)}$=21.16, p<0.01) by approximately 30 s. Predictive information was also significant for when corrective checklists were started ($F_{(2, 17)}$=11.50, p<0.01). Again, the difference between having no predictive information and having status and alert predictive information was approximately 30 s.

Display type by predictive information ($F_{(4, 17)}$=4.51, p<0.01) was significant for how long it took subjects to indicate system setup at the end of the run. As seen in Figure 7, type of predictive information does not matter much except for displays without pattern recognition (MDO). In this case, it is best to have alert predictive information available.

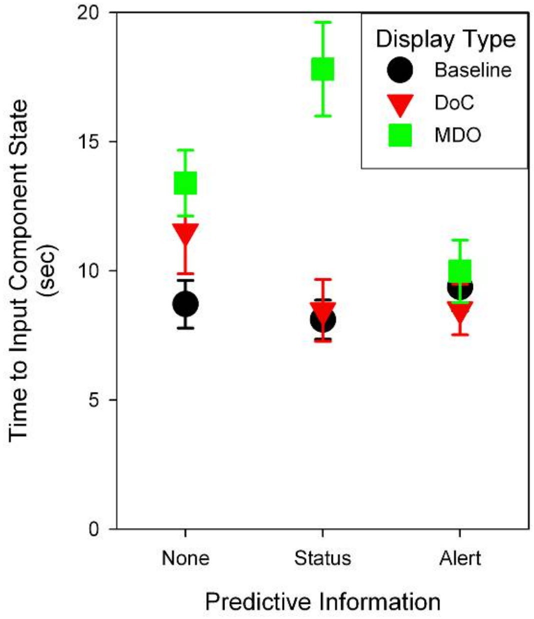

Figure 7. Time to Indicate System Configuration at End of Run

Diagnostic Accuracy of the System Failure

Pilot status by predictive information ($F_{(2, 17)}$=7.04, p<0.01), display type by predictive information ($F_{(4, 17)}$=4.08, p<0.01), and pilot status by display type by predictive information ($F_{(4, 17)}$=3.21, p<0.02) were significant for the accuracy with which subjects recognized failed systems. But as can be seen in Figure 8, pilots with the MDO display appear to drive these statistically significant results. In general, pilot subjects preferred either no predictive information or status predictive information with pattern recognition

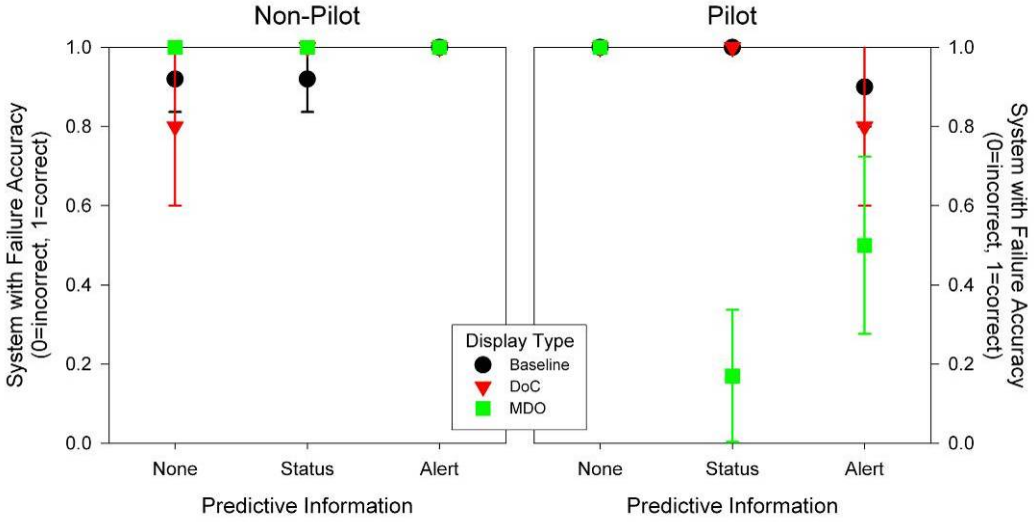

Figure 8. Accuracy in Recognizing System with Failure

displays and non-pilot subjects preferred either status or alert predictive information.

Display type by predictive information ($F_{(4, 17)}$=2.99, p<0.03) and pilot status by display type by predictive information ($F_{(4, 17)}$=2.63, p<0.04) were significant for the accuracy of diagnosis of the failure. Results were similar to that of the accuracy that subjects recognized which system had the failure; non-pilots had more accurate diagnoses with none and status predictive information with a collocated and large area of parameter movement display (MDO) and with alert predictive information with a collocated and pattern recognition display (DoC). For pilots, a collocated pattern-recognition display (DoC) was the best for diagnosis irrespective of predictive information.

Display type by predictive information ($F_{(4, 17)}$=3.53, p<0.01) was significant for accurately recalling component values at the end of the scenario (Figure 9). Here, type of predictive information only matters for collocated, large area of parameter movement displays (MDO).

Lastly, predictive information ($F_{(2, 17)}$=4.59, p<0.02) was significant for the subjects remembering the time to an alert when they first noticed the failure. In general, subjects were the most accurate in remembering the initial time to an alert with status predictive information (Figure 10).

Subjective Preferences

Overall, subjects preferred to have the predictive information compared to no predictive information. With predictive information, subjects found it easier to determine the system (≈51% easier; $F_{(2, 11)}$ = 31.33, p<0.01), component (≈62% easier; $F_{(2, 11)}$ = 62.21, p<0.01), and parameter (≈59% easier; $F_{(2, 11)}$=46.49, p<0.01) with the failure. They also found it easier to determine which checklist to use with predictive information (≈37% easier; $F_{(2, 11)}$=7.10, p<0.01).

Subjects also reported that workload was lower for determination of system status (≈43% less;

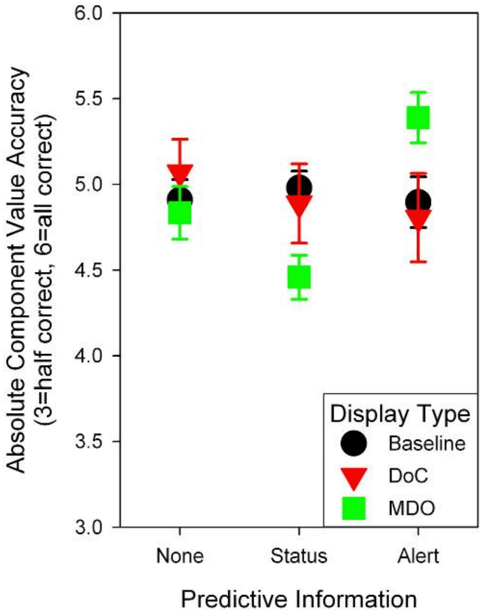

Figure 9. Component Value Accuracy at End of Run

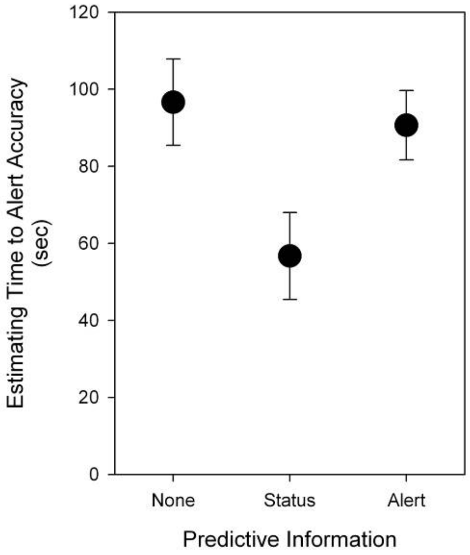

Figure 10. Remembering the Initial Time to an Alert

$F_{(2, 11)}$=24.29, p<0.01), the component with the failure (≈53% less; $F_{(2, 11)}$=50.23, p<0.01), and the parameter with the failure (≈52% less; $F_{(2, 11)}$=39.22, p<0.01) with the predictive information.

Lastly, subjects found predictive information useful ($F_{(2, 11)}$=22.12, p<0.01) and they preferred predictive information ($F_{(2, 11)}$=19.66, p<0.01) (Table 1).

Table 1. Subjective Usefulness and Preference for Predictive Information

Predictive Information	**Average**	**SE**
Usefulness		
None	50.0	0.0
Status	72.8	4.9
Alert	76.4	5.7
Preference		
None	50.0	0.0
Status	75.2	5.7
Alert	77.8	6.3

0=low, 100=high; Compared against No Predictive Information; SE=standard error of the mean

Discussion

Previous research found that subjects like to have information predicting when a parameter would reach an alert range. This current research looked at whether the information should be more closely associated with the parameter information or with the alert information.

Overall, subjects performed better and preferred to have predictive information available although the difference between status and alert predictive information was minimal. Subjects noticed failures with predictive information in 33 s as opposed to 64 s without predictive information. They also started the checklists 153 s before the failure with predictive information as opposed to 122 s before the failure without predictive information. All failures took 39 s to 199 s from manifestation to alert if system configuration was not changed. Therefore, with predictive information, subjects saw the failures before an alert was reached and started the checklist well before any alert range would be reached. This led to lower workload levels.

Previous research found that for detection, subjects wanted a display with a large area of parameter movement [13]. For detecting a failure with predictive information, pilot subjects preformed better with either no predictive information or status predictive information but non-pilot subjects performed better with alert predictive information.

Pilot subjects probably did better with the status predictive information when detecting a failure because they also preferred displays with large areas of parameter movement for this task. This means the pilot subjects were using the parameter displays to look for failures and thus would do better with status predictive information which is on the parameter display. Pilots are accustomed to dividing attention between tasks, in this case keeping a target centered and detecting failures which involve different processing stages and processing codes [8, 10].

For non-pilot subjects who were not used to scanning for changes while timesharing with a tracking task, an alert predictive information message was easier to use in failure detection than status predictive information. The non-pilot subjects were, most likely, employing selective attention [8, 10]. Their attention was directed to the alert screen from the tracking task where the alert predictive information message was much easier to spot than the status predictive information. Furthermore, the alert message was much more rule-based knowledge (similar to “if – then” constructs) for the non-pilot subjects because the alert message directed them to a checklist to perform [10]; *i.e.*, they did not have to understand the system in order to address the abnormality.

For diagnosis with predictive information, the pilot subjects did not have a preference on type of predictive information. Non-pilot subjects performed better with alert predictive information on a pattern recognition display. Non-pilot subjects also preferred no predictive information or status predictive information on a display with a large area of parameter movement. Previous research found that pattern recognition helps diagnosis [13]; therefore, alert predictive information may be best for diagnosis. This is because the alert message contains the diagnosis, which allows for quicker rule-based behavior during a possibly time-critical situation [10].

Subjects did best with status predictive information when they were asked to recall the system configuration and the time to an alert when they first detected the failure. This is most likely due to the simple message located on the affected system. This simple message had to be remembered was, for example, only the time to an alert rather than the time to an alert appended to the end of an alert message. Lastly, subjects preferred to have predictive information with no subjective preference between status and alert predictive information.

Therefore, for detection and recalling what happened, status predictive information is best. For diagnosis, alert predictive information holds a slight edge. In general, predictive information allows for operators to remedy problems before they become critical. Operator decisions would, most likely, be better because operators would have more time to troubleshoot and plan, which would help lower workload and stress during non-normal events. This may also lead to better decision making during knowledge-based related non-normal events where considering all factors, rather than quickly focusing on a solution, is necessary to optimally solve the problem.

References

[1] Trujillo, Anna C., 1994, Effects of Historical and Predictive Information on Ability of Transport Pilot to Predict an Alert, Hampton, VA, NASA Langley Research Center, NASA-TM-4547.

[2] Trujillo, Anna C., 1996, Airline Transport Pilot Preferences for Predictive Information, Hampton, VA, NASA Langley Research Center, NASA-TM-4702.

[3] Trujillo, Anna C., Oct 12-15, 1997, Pilot Performance With Predictive System Status Information, in *1997 IEEE International Conference on Systems, Man, and Cybernetics*, Orlando, FL, pp. 1972-1977.

[4] Trujillo, Anna C., March 22-24, 1998, Pilot Mental Workload with Predictive System Status Information, in *4th Annual Symposium on Human Interaction with Complex Systems*, Fairborn, OH, pp. 73-80.

[5] Trujillo, Anna C., May 2-6, 1999, Changes in Pilot Behavior With Predictive System Status Information, in *Tenth International Symposium on Aviation Psychology*, Columbus, OH.

[6] Trujillo, Anna C., 2002, Vertex Movement for Mission Status Graphics A Polar-Star Display, Hampton, VA, NASA Langley Research Center, NASA-TM-2002-211414, p. 18.

[7] Trujillo, Anna C., 2004, Pilot Preferences for Information Provided and Its Format for Status, Alerts, and Controls, Hampton, VA, NASA Langley Research Center, NASA-TP-2004-213498, p. 95.

[8] Sanders, Mark S., Ernest J. McCormick, 1987, Human Factors in Engineering and Design. USA: McGraw-Hill Publishing Co.

[9] Weiner, Earl L., Barbara G. Kanki, Robert L. Helmreich, 1993, Cockpit Resource Management: Academic Press.

[10] Wickens, Christopher D., 1984, Engineering Psychology and Human Performance. Glenview, IL: Scott, Foresman and Co.

[11] Bartolone, Anthony P., Anna C. Trujillo, 2002, Glass-Cockpit Pilot Subjective Ratings of Predictive Information, Collocation, and Mission Status Graphics: An Analysis and Summary of the Future Focus of Flight Deck Research Survey, Hampton, VA, NASA Langley Research Center, NASA-TM 2002-211419, p. 53.

[12] Federal Aviation Administration, August 28, 2008, Electronic Code of Federal Regulations - Title 14: Aeronautics and Space Subpart E-Private Pilots Section 61.103. vol. 2008: FAA.

[13] Trujillo, Anna C., Hayes N. Press, April 23-26, 2007, Collocation and Pattern Recognition Effects on System Failure Remediation, in *2007 (14th) International Symposium on Aviation Psychology*, Dayton, OH, pp. 707-713.

[14] Mahaffey, D. L., R. L. Horst, R. C. Munson, 1986, Behavioral Comparison of the Efficacy of Bar Graphs and Polar Graphics for Displays of System Status, in *1986 IEEE International Conference on Systems, Man and Cybernetics*, pp. 1514-1519.

[15] Trujillo, Anna C., 2001, Experience and Grouping Affects When Handling Non-Normal Situations, in *Human Factors and Ergonomics Society 45th Annual Meeting*, pp. 302-306.

[16] Trujillo, Anna C., 2001, Response Times in Correcting Non-Normal System Events When Collocating Status, Alerts and Procedures, and Controls, in *2nd International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*, Manchester, England, pp. 8-13.

[17] Albert, Robert, Noah Syroid, Yinqi Zhang, Jim Agutter, Frank Drews, Dave Strayer, George Hutchinson, Dwayne Westenskow, October 19-24, 2003, Psychophysical Scaling of a Cardiovascular Information Display, in *IEEE Visualization 2003*, Seattle, Washington, pp. 35-42.

[18] Byers, James C., Alvah C. Bittner, Susan G. Hill, 1989, 1989, Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary?, in *Advances in Industrial Ergonomics and Safety*, pp. 481-485.

[19] Hart, S. G., L. E. Staveland, 1988, Development of a NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in *Human Mental Workload*, Hancock, P. S., N. Meshkati, Eds. Amsterdam: Elsevier Science Publishers B. V., pp. 139-183.

[20] Cooper, George E., Robert P. Harper, 1969, The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities, Technical Report 567, AGARD, p. 52.

[21] Harper, Robert P., George E. Cooper, 1986, Handling Qualities and Pilot Evaluation (Wright Brothers Lecture in Aeronautics), in *Journal of Guidance, Control, and Dynamics*, vol. 9, no. 6, pp. 515-529.

Acknowledgements

The authors would like to thank Messrs. Lucas Hempley and John Barry of Lockheed Martin for programming the simulation and helping with data collection and preliminary analysis.

Disclaimer

The use of trademarks or names of manufacturers in the report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.