RESPONSE TIMES IN CORRECTING NON-NORMAL SYSTEM EVENTS WHEN COLLOCATING STATUS, ALERTS AND PROCEDURES, AND CONTROLS

Anna C. Trujillo

NASA Langley Research Center, USA

ABSTRACT

Currently, most of the displays in control rooms can be categorized as status, alerts/procedures, or control screens. With the advent and use of CRTs and the associated computing power available to compute and display information, it is now possible to combine these different elements of information and control onto a single display. An experiment was conducted to determine which, if any, of these functions should be collocated in order to better handle simple anticipated non-normal system events. The results indicated that there are performance benefits and subject preferences to combining all the information onto one screen or combining the status and alert/procedure information onto one screen and placing the controls in another area.

INTRODUCTION

Currently, most of the displays in control rooms can be categorized as status screens, alerts/procedures screens (or paper), or control screens (where the state of a component is changed). This is likely a holdback from the steam-gauge era when one instrument had one use. With the advent and use of CRTs and various types of input devices, and the associated computing power available to compute and display information, it is now possible to combine these different elements of information and control onto a single display. This will result in space and weight savings. One domain where this is especially relevant is in airplane flight decks where these space and weight savings will translate into fuel savings and ultimately a saving of money due to a decrease in operating expenses.

An overriding question that needs to be answered is *Which of these information and control elements should be grouped?* Previous research has developed some guidelines on collocating two of these elements but none has considered collocating all three of these elements (table 1).

Objectives

This experiment was conducted to determine which functions should be collocated in order to better handle simple anticipated non-normal system events. The functions were status information, alert/procedure messages, and control screens. Simple was defined as no propagation between or within a system and anticipated was defined as having a checklist available in order to remedy the problem. Simple anticipated failures were used to first see if collocating functions was beneficial and what display combinations were best. More complicated and realistic failures will be used in latter experiments using the promising display combinations in order to fully document the effects of collocating these functions. Finally, better was characterized as improving workload and reducing the detection, diagnosing, and controlling of non-normal events.

Although applicable to most any type of control room (*e.g.*, nuclear power plant control rooms, assembly line

TABLE 1 – Previous research results on grouping functions

Combine Functions	Separate Functions
• Combine stimulus and response	• Separate displays and controls
• Reduce crew workload by collocating displays and controls [Control room design – Sanders and McCormack (1)]	• Better performance with less cluttered displays [Display automation and decluttering – Stokes and Wickens (2) Control display unit guidelines – Mann and Morrison (3)]
• Reduce complexity of data search [Decision making – Wickens (4)]	• Related data should be grouped and separated from unrelated data [Control room design – Sanders and McCormack (1) Multifunction display and control systems design – Francis and Reardon (5)]
• Command decision aids should be augmented with status information [Aircraft decision aids design – Andre and Wickens (6)]	

control rooms, and aircraft flight decks), this experiment looked at collocating these functions using systems found in most power plants or aircraft. The generic systems used were power plant, fuel feed, and heat exchanger systems. Subjects monitored these systems and controlled the system configuration when a failure occurred. As a secondary task, subjects had to keep a randomly moving target centered on a display using a side stick before, during, and after the failures.

EXPERIMENT DESIGN

Experimental variables

Of the four experimental variables, three were of primary concern: display configuration, number of systems with an alert, and whether the subject was a pilot (which was to determine whether pilots had a bias towards the display configuration that is currently on aircraft). The fourth independent variable was the 18 independent faults each subject encountered.

The display configuration, a modified between subject variable, had five gradations: (1) the three functions on separate screens, s/a/c; (2) the three functions grouped on one screen, sac; (3) status and alert/procedure information grouped on one screen and controls on a separate screen, sa/c; (4) status and controls grouped on one screen and alert/procedures on a separate screen, sc/a; and (5) alert/procedures and controls grouped on one screen and status on a separate screen, s/ac. The number of systems with an alert, a within subject variable, was 1, 2, or 3 systems with an alert. The third experimental variable was whether the subject was a pilot or had no piloting experience.

Display configuration. For the five display configurations, each subject saw the baseline configuration plus one of the collocated display configurations. The baseline condition was defined as the three functions on separate displays, *i.e.*, s/a/c, which is the current configuration in most control rooms.

When all three functions were collocated, sac, (fig. 1) the status information, in the form of a bowtie with the parameter value indicated at the top, was to the far left. Next, the control switch of the component was to the right of the status information. Lastly, the associated alert message with the procedure below it was located to the far right.

When two functions were collocated on the same screen with the third one on a separate screen, the collocated functions were grouped on a single screen with one function pulled out and located elsewhere. For the s/a/c and sc/a display configurations, the alerts for each system were displayed together (fig. 2). The other two display combinations, sa/c and s/ac, are shown in figures 3 and 4.

Number of systems with an alert. During each scenario, there would be up to three alerts but only one alert per system. Six scenarios had 1 alert, six scenarios had 2 alerts, and six scenarios had 3 alerts. None of the alerts propagated within or between systems; therefore, only one component per system was the root of the failure. When there were multiple failures, they were timed such that all alerts occurred at the same time. This eliminated the order of alerts between systems factor during data analysis.

Pilot status. Half of the subjects were commercial glass-cockpit line-pilots. The other half of the subjects had no piloting experience. As mentioned earlier, this was done so that the possible bias of pilots towards using three separate displays could be measured against a population that supposedly had no formal experience on the separate display configuration.

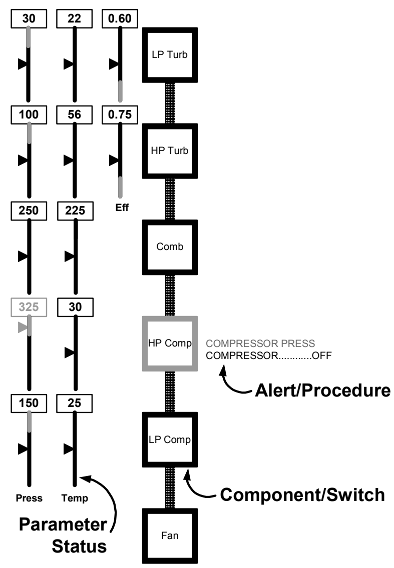

Figure 1: Combined power plant display

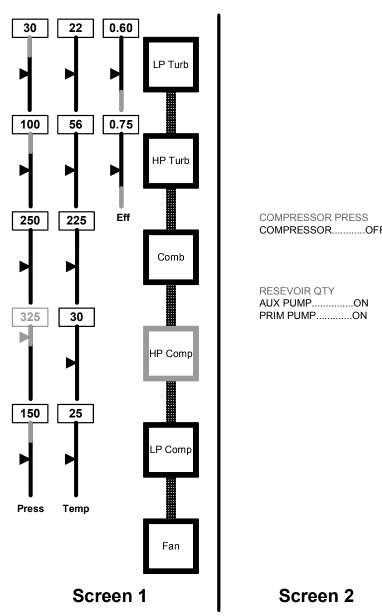

Figure 2: Power plant sc/a display

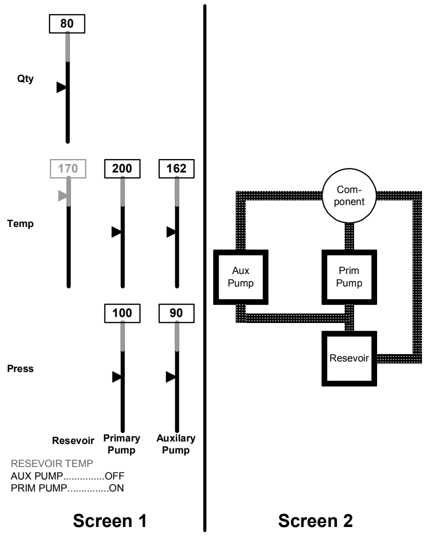

Figure 3: Heat exchanger sa/c display

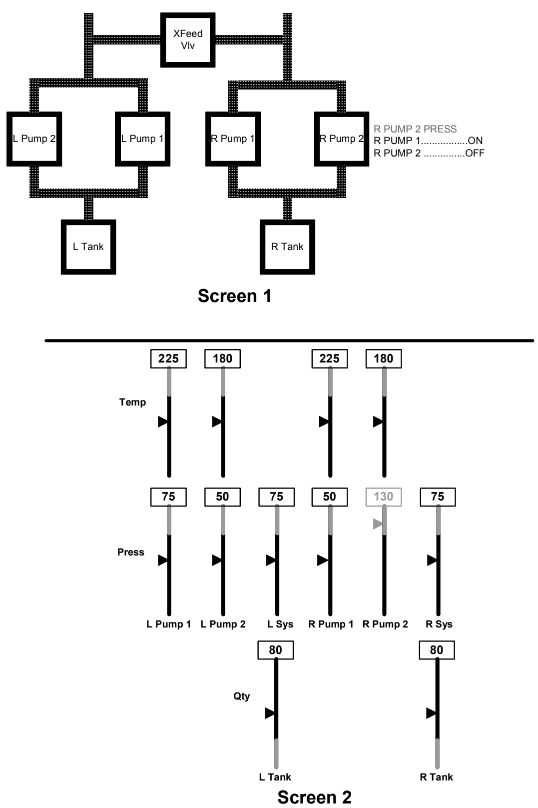

Figure 4: Fuel feed s/ac display

Faults. Each subject saw 18 faults where up to three alerts would occur. Each fault was accompanied with an alert message and procedure, the parameter values were in the indicated alert range, and the failed component was outlined in amber.

Subjects

Sixteen people participated as subjects; eight were commercial glass-cockpit line-pilots and eight had no piloting experience. Of the pilots, 4 were currently first officers and the remaining 4 were captains. The average age of the pilots was 37 years old with an average commercial airline flight experience of 14 years. The average age of the non-pilots was 44 years old.

Test design

The experiment was conducted in the Intermediate Design and Evaluation Simulation Lab at the NASA Langley Research Center. This lab allows for simulation of various systems. In this experiment, the systems simulated were a power plant (fig. 1), fuel feed (fig. 4), and heat exchanger (fig. 3). The parameters for the components consisted primarily of pressure, temperature, and quantity measurements. Each of these systems was independent from one another and subjects were notified of this.

As mentioned earlier, the faults and number of systems with alerts were within subject variables while the display configuration and pilot status were between subject variables. Since subjects could only see each failure once, each subject had 18 data runs in addition to four training runs (2 before the baseline display condition and 2 before the collocated display condition). Thus, all subjects saw each of the 18 faults once and each of the number of systems with an alert three times with the baseline display configuration and with a collocated display configuration (fig. 5).

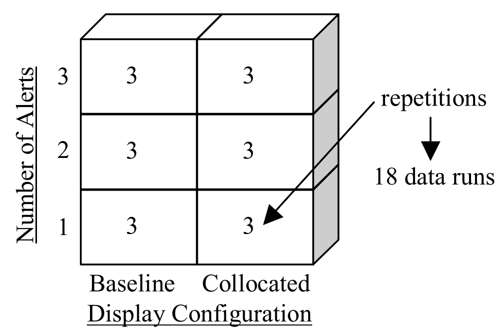

Figure 5: Test design for each subject

Dependent measure

The objective dependent measures were the amount of time taking care of the failures (*i.e.*, checklist completion time), whether the checklists were completed, and the accuracy of the tracking task. The subjective independent measures consisted of subjective workload ratings and subject preferences about the display configurations.

Procedure

When a subject first arrived, he was given an overview of this experiment. He then received instruction on each of the systems, how to bring up each system with their related screens after an alert occurred, and how to perform the checklists through the touch screens.

After this briefing, the subject went to the simulator where he was able to familiarize himself with the tracking task and each of the systems with the first display configuration he would be using. Before any data runs, the subjects had two practice runs that behaved the same as the data runs. After the two practice runs, the nine data runs with the first display configuration were completed.

During the initial part of the data run, the subject kept a randomly moving target centered using a sidestick on his left side. This task continued throughout the data run. Two to four minutes after the beginning of the tracking task, one to three alerts occurred. At this time, the subject had to access the checklist(s) in order to remedy the failure. Once the subject reported that he had completed the checklist(s), the data run ended. When the data runs for the first display configuration were finished, the display configuration was changed and the subject had two practice runs with the new configuration before the nine data runs began.

At the end of each data run, subjects recorded their workload ratings using the NASA-TLX (7). Finally, at the end of all data runs, subjects completed a questionnaire asking them about their display preferences.

Data analysis

Data was analyzed using SPSS®, Statistical Product and Service Solutions (8). The subject accuracy in the tracking task, time it took the subject to complete the checklists data, and workload ratings were analyzed using a repeated measures test in SPSS®. The repeated measures for these analyses were the number of systems with an alert and the repetition number (rep). The questionnaire data was analyzed using a Chi^2 test. In all cases, significance was set at $p \leq 0.05$.

RESULTS

Checklist completion time

The time to complete the checklists was dependent on the repetition number by display type ($p=0.03$). As seen in figure 6, checklists were completed the fastest using the sac and sa/c displays. Furthermore, although not statistically significant, checklists were completed just as or more accurately using the sac and sa/c displays as they were for the baseline display, s/a/c (see table 2).

Tracking task

The distance from center for the tracking task while the subjects tended to the alerts was dependent on the repetition number by pilot status by display type interaction ($p<0.01$). As shown in figure 7, pilots were able to keep the target centered better than non-pilots were. The overall best collocation displays for both groups appear to be sac, sa/c, and sc/a.

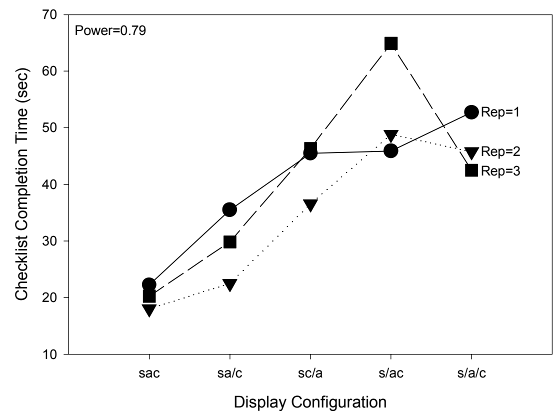

Figure 6: Display time by display configuration

TABLE 2 – Checklist completion accuracy

Display Config.	Power Plant	Fuel Feed	Heat Exchanger
sac	0.89	0.91	0.89
sa/c	0.85	0.93	0.88
sc/a	0.81	0.84	0.78
s/ac	0.84	0.84	0.94
s/a/c	0.89	0.91	0.75
p-value	0.84	0.64	0.28
estimated power	0.13	0.20	0.39

Note: 1=completed, 0=not completed

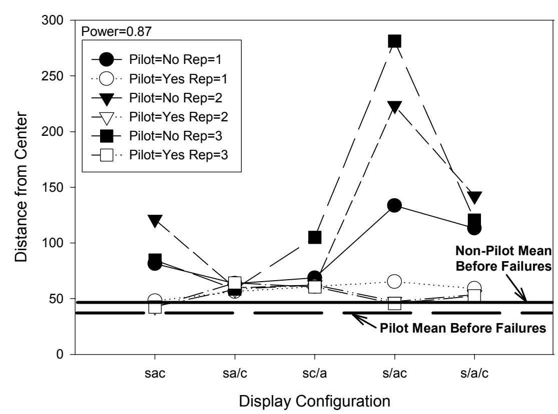

Figure 7: Tracking task by display configuration, pilot status, and number of systems with alerts

Subjective preferences

The following data were not significant but indicated a trend toward particular display configurations (table 3).

Workload. Workload was rated subjectively lower for the sac and sa/c displays. The other two collocated displays, sc/a and s/ac, were given similar subjective ratings as the baseline configuration, s/a/c.

Clutter of the collocated display. Clutter was measured by having subjects rate how more or less cluttered the combined displays were compared to the baseline display of s/a/c. Clutter was rated subjectively lowest for the sa/c and s/ac collocated displays. The other two collocated displays were given similar subjective ratings as the baseline display.

Display preference. When compared to the baseline display, subjects preferred the collocated displays overall. Their subjective ratings were highest for the sac and sa/c collocated displays.

TABLE 3 – Trend data

Display Config.	Workload	Clutter	Preference
sac	12.8	47.8	85.0
sa/c	12.8	22.5	90.7
sc/a	23.5	57.5	56.1
s/ac	35.7	31.8	77.1
s/a/c	25.1	50.0	50.0
p-value	0.41	0.16	0.24
estimated power	0.24	0.42	0.30

Note: 0=low, 100=high

DISCUSSION

To explore the benefits collocating functions would have on the ability of an operator to handle simple anticipated non-normal system events, a simulator experiment tested five different display configurations. The display configurations differed in which information (status, alert/procedure, and control) was grouped on the display. The number of systems with an alert and whether the subject was a pilot or not were the other important independent variables.

The data indicated that the sac and sa/c displays were best. The former display collocated all the information onto one display. The latter collocated the status and alert/procedure information onto one display and the controls onto a separate display.

The objective data showed that the amount of time to complete the checklists and the loss of accuracy of the tracking task were the least with these displays. The data from the tracking task did show a difference between subjects who were trained (pilots) and who were untrained (non-pilots) for handling system failures while attending to another continuous task.

Also note that while pilots performed well on the tracking task with all displays, non-pilots had a much greater difficulty keeping the target centered with the s/ac display. The checklist completion time also showed a decrease in performance for all subjects when using the s/ac display. Therefore, it was eliminated from consideration.

Although not statistically significant, the accuracy of completing the checklists was also high using the sac and sa/c displays. The checklist completion accuracy data trends deserve consideration since they show a constant pattern across systems and display configurations even with the estimated observed power being relatively low.

Furthermore, subjective data pointed to a preference of the sac and sa/c displays. As mentioned earlier, this data was not statistically significant but the trend does hold across the subjective measures especially since the estimated observed power was also low.

Therefore, the above results indicate that that the sac and sa/c displays are better than the baseline display, s/a/c, for handling simple anticipated system failures. This finding holds when considering the time to complete the checklist(s) and the accuracy of the tracking task. Subject preferences and subject accuracy in completing the checklist(s) also corroborated this finding although these particular results were not statistically significant with the small sample size run.

CONCLUSIONS

Computers and CRTs have enabled designers to combine different types of information onto one display. This is an attractive proposition because savings in space, weight, and materials result in cost savings. Even though the ability exists to combine displays and functions, it must be determined whether this will adversely affect the operator of a system.

This experiment began answering this question by collocating status, alert/procedure, and control information. The results indicated that combining all three pieces of information onto a single screen or combining the status and alert/procedure information on one screen and separating out the controls to another area may improve performance over the current display configuration of keeping this information separate from one another.

These results are promising not just because of the performance increases but also because of possible cost savings to various industries. However, before it can be definitively said that a totally collocated display is best, further research must be done using a more real-world simulation; *i.e.*, interaction between systems needs to be considered before this information is combined onto one or two displays.

REFERENCES

1. Sanders, Mark S. and McCormick, Ernest J., 1987. Human Factors in Engineering and Design, McGraw-Hill Publishing Co., USA, 380-384.

2. Stokes, Alan F. and Wickens, Christopher D., 1988. “Aviation Displays,” Human Factors in Aviation, Wiener, Earl L. and Nagel, David C. eds., Academic Press, Inc., San Diego, CA, 420-421.

3. Mann, Teresa L. and Morrison Jeff G., 1986. “Effects of Display Density and Format Type on Control Display Unit Format Design,” in IEEE/AIAA 7th Digital Avionics Systems Conference, Forth Worth, TX, 330-337.

4. Wickens, Christopher D., 1984. Engineering Psychology and Human Performance, Scott, Foresman and Co., Glenview, IL, 73-118.

5. Francis, Gregory and Reardon, Matthew J., 1997. “Aircraft Multifunction Dipslay and Control Systems: A New Quantitative Human Factors Design Method for Organizing Functions and Display Contents,” US Army Aeromedical Research Laboratory, Fort Rucker, AL, 1-40.

6. Andre, Anthony D. and Wickens, Christopher D., 1992. “Compatibility and Consistency in Display-Control Systems: Implications for Aircraft Decision Aid Design,” Human Factors, 34(6), 639-653.

7. Human Performance Research Group. “NASA Task Load Index (TLX) v. 1.0: Paper and Pencil Package,” NASA Ames Research Center, Moffett Field, CA, 1-19.

8. SPSS Inc., 1999. SPSS® User’s Guide, SPSS Inc., Chicago, IL.