

EXPERIENCE AND GROUPING EFFECTS WHEN HANDLING NON-NORMAL SITUATIONS

Anna C. Trujillo
NASA Langley Research Center
Hampton, VA

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. The results indicated that operators quickly modify their display preferences to the display configuration most recently used.

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 holdout 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 monetary savings due to a decrease in operating expenses.

Previous research has developed some guidelines on collocating two of these elements but none has considered how collocating all three of these elements would affect the operator (Andre & Wickens, 1992; Sanders & McCormick, 1987; Stokes & Wickens, 1982; Mann & Morrison, 1986; Wickens, 1984; Francis & Reardon, 1997). Other results from this particular research on which elements should be collocated determined that combining all the information on to a display or combining the status and alert/procedure information onto one screen and placing the controls in another area were best when considering time to complete checklist(s), tracking task accuracy, and subjective preferences (Trujillo, 2001). Other factors, such as the experience of the operator and whether his initial grouping of these elements matched the element configuration he saw, might affect these results. Since these factors were thought to affect performance significantly, they were also designed into the experiment described in reference (Trujillo, 2001). These factors and their effects on performance are detailed in this paper. In particular, this paper examines *how the way people naturally group these elements affect their performance when using element configurations that match or mismatch their groupings and how experience affects their performance.*

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 later 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 detecting, diagnosing, and controlling times for non-normal events.

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

There were five experimental variables: display configuration, number of systems with an alert, whether the subject was a pilot, whether the display configuration matched the subjects preferences, and the individual faults. The first four were of primary concern.

Display configuration. The display configurations were: all functions separated on three displays (status/alerts/controls or s/a/c); all functions collocated on one display (sac), status and

alerting collocated with control separate (sa/c); alerting and controls collocated with status separate (s/ac) and status and controls collocated with alerting separate (sc/a). These configurations were modified between subject variables. Each subject saw the s/a/c configuration plus one of the collocated display configurations. The s/a/c configuration is considered the baseline condition since it is typical of most operator stations.

For the sac configuration, (fig. 1) the status information, in the form of a bowtie with the parameter value indicated at the top, was to the far left. The control switch for 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.

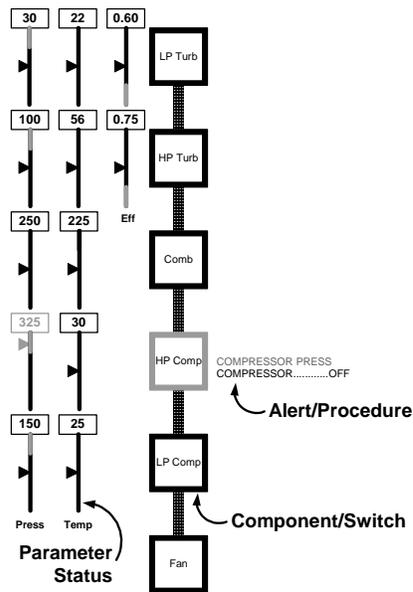


Figure 1 – Combined power plant display

For the 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. The number of systems with an alert, a within subject variable, was 1, 2, or 3. Therefore, 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 (exp). The other half of the subjects had no piloting experience (inexp). This was manipulated so that the possible bias of pilots towards using three separate displays (as is the case in most flight decks) could be measured against a population that supposedly had no formal experience on the separate display configuration.

Grouping match or mismatch. The fourth experimental variable, also between subjects, was whether the subject saw a

combined display configuration that matched his preference in grouping the information (match) or whether it mismatched his grouping preference (mismatch).

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.

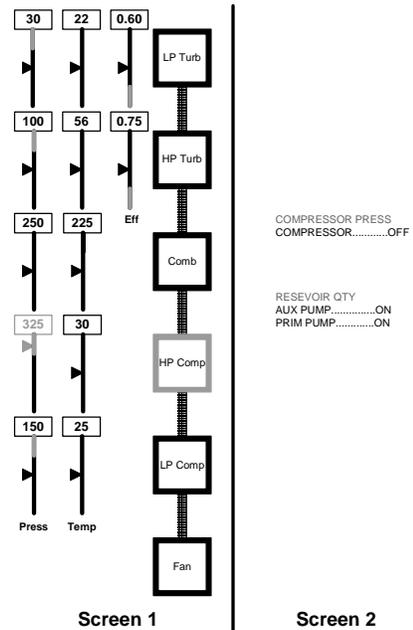


Figure 2 – Power plant sc/a display

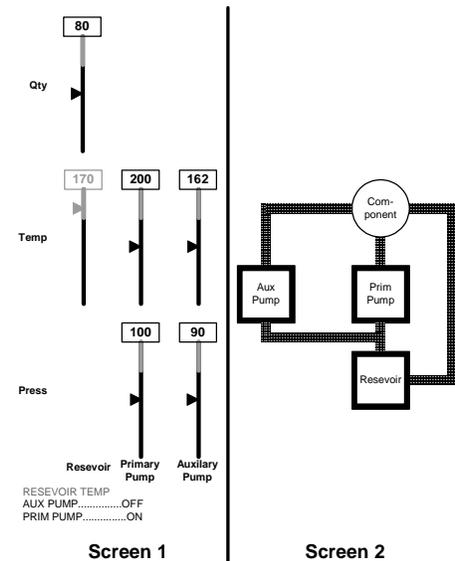


Figure 3 – Heat exchanger sa/c display

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

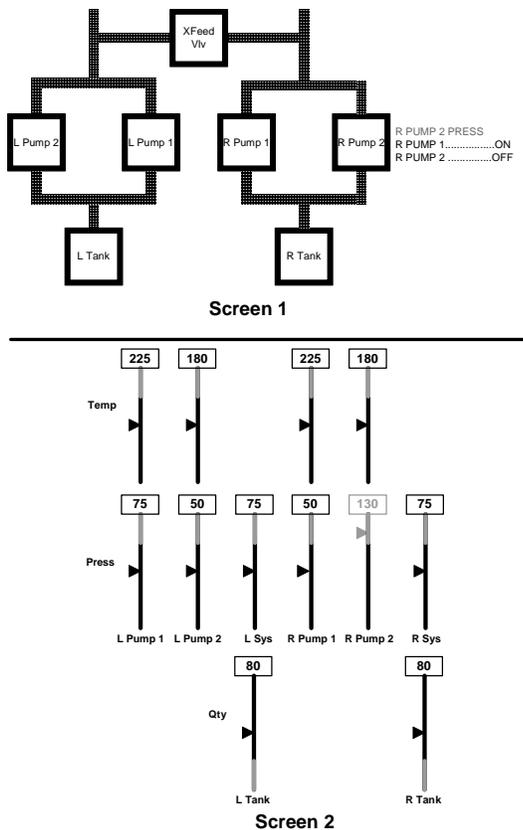


Figure 4 – Fuel feed s/ac display

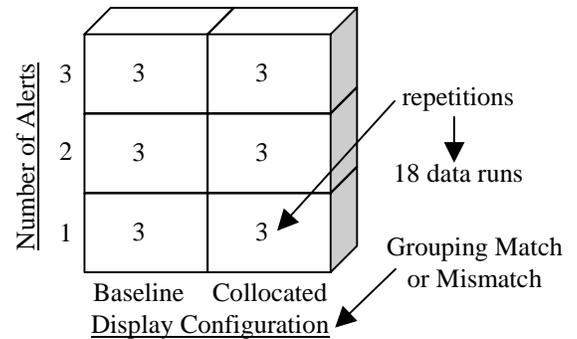


Figure 5 – Test design for each subject

Dependent Measures

The objective dependent measures were whether the checklists were completed, whether the status page was brought up, and the accuracy of the tracking task. The subjective independent measures consisted of subjective workload ratings and subject grouping preferences about the display configurations before and after using a particular display configuration.

Procedure

A questionnaire that determined subject grouping preferences was distributed well before the simulator portion of this experiment. Subjects were then assigned a display configuration based on the questionnaire's results.

When a subject first arrived for the simulator portion of this experiment, 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) and access other pertinent screens 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 (Human Performance Research Group). Finally, at the end of all data runs, subjects completed a questionnaire asking them about their display preferences and they redid the grouping preference

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 (PP) (fig. 1), fuel feed (FF) (fig. 4), and heat exchanger (HE) (fig. 3). The parameters for the components consisted primarily of pressure, temperature, and quantity measurements. Each of these systems was independent of 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, pilot status, and grouping match or mismatch 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).

questionnaire in order to see if working with a particular display configuration affected their grouping preferences.

Data Analysis

Data was analyzed using SPSS®, Statistical Product and Service Solutions (SPSS, 1999). The subject accuracy in the tracking task, 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 Chi² test. In all cases, significance was set at p≤0.05.

RESULTS

Tracking Task

Experienced subjects did better on the tracking task during alerts. This effect was significant for experience (p≤0.04), experience X alerts (p≤0.01), and experience X repetition X display (p≤0.01). Also for the tracking task, there was an experience X group X repetition X display interaction (p≤0.01). In general, experienced and matched subjects did better on the tracking task (fig. 6).

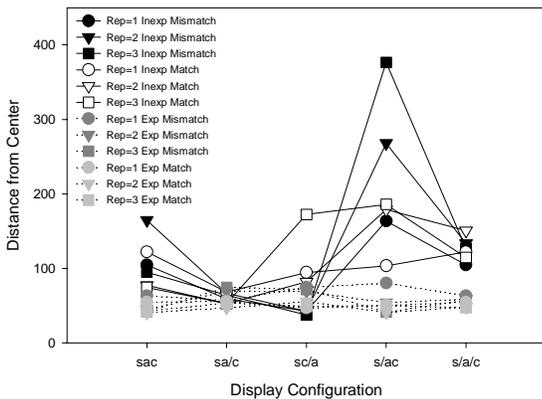


Figure 6 – Tracking Task Accuracy

Workload

The experience X group X alerts X repetition X display interaction was significant (p≤0.05). Workload was the least for experienced subjects with no match and for inexperienced subjects that matched (fig. 7).

Problem Recognition

Experienced and matched subjects were more likely to bring up the status page (significant for HE at p≤0.01, and for FF and PP at p≤0.02) (fig. 8).

Furthermore, experienced subjects were more likely to complete the checklists (significant for PP at p≤0.01) (fig. 8), and remember what the non-normal situations were (significant for PP and FF at p≤0.01) (fig. 9).

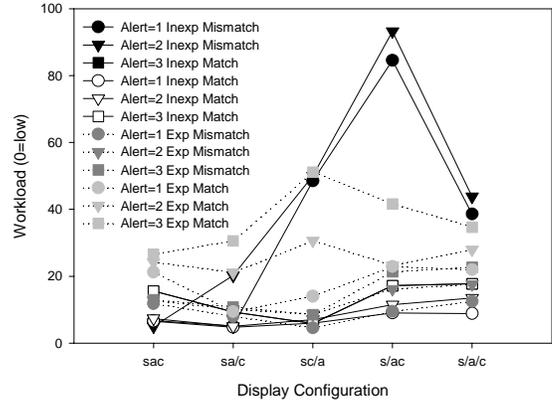


Figure 7 – Subjective Workload Ratings

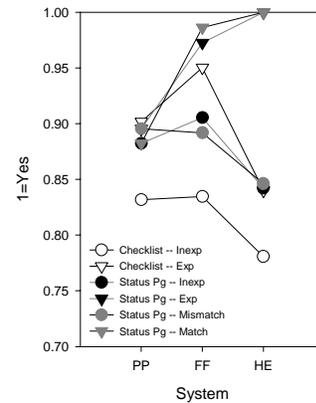


Figure 8 – Checklist Completion and Status Page Viewed

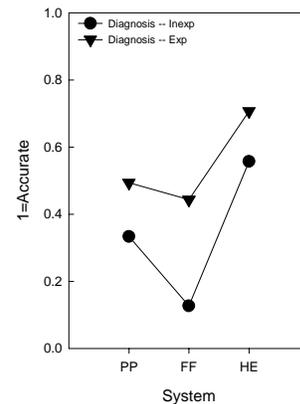


Figure 9 – Non-Normal Situation Recall

Grouping Change

In general, experienced subjects did not change their preferred grouping when they had a mismatch whereas inexperienced subjects did change their preferred grouping to match one of the display configurations they used (fig. 10).

DISCUSSION

As can be seen in the figures, inexperienced subjects with a match basically performed as well as experienced subjects for the tracking task. These groups also subjectively rated workload about the same. As for the best display

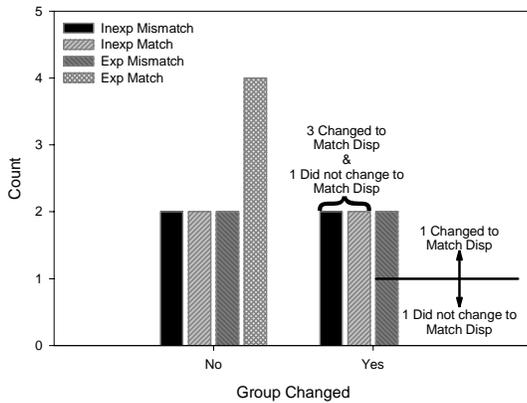


Figure 10 – Grouping Changes

combination, the tracking task accuracy and the subjective workload ratings indicate that either the combined display (sac) or combining the status and alert/procedure information on one screen and separating out the controls to another area (sa/c) are the best combined displays. These display combinations coincide with the best display combinations reported in Trujillo, 2001.

Not surprisingly, experienced subjects were better at checklist completion and non-normal situation recall. Also note that experienced and matched subjects were about as likely to bring up the status page.

Lastly, inexperienced subjects tended to change their grouping to match one of the display configurations they worked with. None of the experienced–matched subjects changed their grouping preference.

These results suggest that while experienced subjects with a match did the best overall, inexperienced subjects with a match may quickly perform as well after some experience. The results also indicate that performance on problem recognition might also improve dramatically in inexperienced subjects.

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. The results also indicated that with training, using one of these two displays is feasible even if there is initially a mismatch in grouping the information because of the grouping change to a match after use of the display.

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

- Andre, A. D. & Wickens, C. D. (1992). Compatibility and Consistency in Display-Control Systems: Implications for Aircraft Decision Aid Design. *Human Factors*, 34(6), 639-653.
- Francis, G. & Reardon, M. J. (1997). *Aircraft Multifunction Display and Control Systems: A New Quantitative Human Factors Design Method for Organizing Functions and Display Contents* (pp. 1-40). Fort Rucker, AL: US Army Aeromedical Research Laboratory.
- Human Performance Research Group. *NASA Task Load Index (TLX) v. 1.0: Paper and Pencil Package*. Moffett Field, CA: NASA Ames Research Center.
- Mann, T. L. & Morrison J. G. (1986). *Effects of Display Density and Format Type on Control Display Unit Format Design* (pp. 330-337). Paper presented at the IEEE/AIAA 7th Digital Avionics Systems Conference, Forth Worth, TX.
- Sanders, M. S. & McCormick, E. J. (1987). *Human Factors in Engineering and Design* (pp. 380-384). USA: McGraw-Hill Publishing.
- SPSS Inc. (1999). *SPSS® User's Guide*, SPSS Inc., Chicago, IL.
- Stokes, A. F. & Wickens, C. D. (1988). Aviation Displays. In Wiener, E. L. & Nagel, D. C. (eds.), *Human Factors in Aviation* (pp. 420-421). San Diego, CA: Academic Press.
- Trujillo, A. C. (2001, June) *Response Times In Correcting Non-Normal System Events When Collocating Status, Alerts and Procedures, and Controls*. Paper presented at the People in Control 2001 Conference, Manchester, UK.
- Wickens, C. D. (1984). *Engineering Psychology and Human Performance* (pp. 73-118). Glenview, IL: Scott, Foresman.