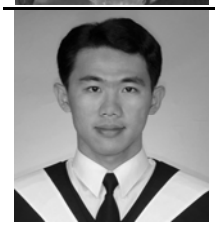
The Human Capacity of Signal Detection and Reaction in Multi-Screen Monitoring Tasks

Lin, Chiuhsiang Joe

Department of Industrial Engineering / Chung-Yuan Christian University / 200, Chung Pei Rd. / Chung- Li / Taiwan 32023, R.O.C.
E-mail: hsiang@cycu.edu.tw

Chang, Hsiao-Ching

Department of Industrial Engineering / Chung-Yuan Christian University / 200, Chung Pei Rd. / Chung- Li / Taiwan 32023, R.O.C.
E-mail: abby165019@hotmail.com



Yang, Chih-Wei

Department of Industrial Engineering / Chung-Yuan Christian University / 200, Chung Pei Rd. / Chung- Li / Taiwan 32023, R.O.C.
E-mail: wayne383@ms23.hinet.net

ABSTRACT

This study evaluated the human signal detection performance between single screen display and multiple screen display. A simulated experiment was used to test participants' performance with random signals and different number of split screens for a period of time. Results indicated that the response time and number of false alarms significantly increased with the number of split screens. The center screen was found to be paid more attention to as indicated by both the higher hit rates and false alarms. Designers should take into consideration the trade-off between the number of multiple split screens and the limited human capability.

Keywords

Multi-screen monitor, visual search, signal detection theory, attention

INTRODUCTION

Supervisory control has been widely implemented in various industries (Sheridan, 2006). Human operators under supervisory control condition are usually assigned to monitor several information sources continuously or periodically to looking for signals that require further actions (Moray, 2003). For each information source, there may be different types of information associated with a different signal frequency. Supervisory controllers must allocate their attention over these sources with appropriate sampling strategies and frequencies followed by state identification that will suggest appropriate action (Moray, 2003). Early studies (Tulga and Sheridan, 1980; Yeh and Wickens, 1988) tested the relationship between overload condition of workload and task performance. Their results indicated that subjective mental workload was correlated with the increasing task demands up to a point where task demands were equal to available total individual capacity.

Piechulla, Mayser, Gehrke, and Köonig (2003) found that the amount of information provided on the controller workstation displays influenced the controllers' ability to extract critical data. Other studies suggested that there existed an optimal display design in which appropriate arrangement of information and amount could improve controller's performance and reduce workload (Tsang and Wilson, 1997; Moray, 1979, 1988). A particular case of supervisory control is the security monitoring system where the human operator is brought to watch several displays containing video camera information. The advancement of technology has made multiple screens on a single display available to accommodate several camera views. It seems that the more camera views, the safer the system, assuming that the human operator is able to watch as many camera views as provided. However, the signal detection performance under these situations may degrade as can be predicted by the supervisory control theory, but how?

The term supervisory control in the strict sense indicates that one or more human operators are setting initial conditions for intermittently adjusting and receiving information from a computer that itself closes a control loop through external sensors, effectors, and the task environment (Sheridan, 2006). Very little experimental work has been conducted to understand the human performance in multi-screen monitoring tasks as to how the number of screens and frequency of signal occurrence affect the signal detection performance. An experimental study was therefore designed to test the signal detection performance of one screen compared to multiple screen display.

METHOD

The purpose of the present research was to examine the human capacity of signal detection and reaction within a multi-screen monitoring task. An empirical investigation was therefore conducted to assess the effect of the number of multiple screens and split screens arrangement on human performance and subjective vision assessment in a simulation of a dynamic control task. The results of the this research can be used to provide an indication of the relevant task aspects that may influence the success or failure of signal detection that seek to create an effective and safe multi-screen monitoring system.

PARTICIPANTS

Ten undergraduate students comprising four females and six males having normal or corrected to 20/20 visual acuity and full-color vision participated for monetary compensation (3 US dollar/hr) on a voluntary basis. Participants ranged in age from 23 to 29 years. The average age of the participants was 24.2 years. All participants were right-hand dominant. These participants were required to have personal computer (PC) experiences. These characteristics were considered to be relevant to performance of the experimental tasks; however, they were not used as control factors in data analyses.

TASK

The study was conducted using a computer dynamic control task called Interstellar Simulator Task 1.0 (IST 1.0). The IST 1.0 simulation presents a single screen or multiple screens depending on the experimental condition. On each screen, the signals were shown as red dots and noises as white dots on gray background, as shown in Figure 1. Both the white and red dots occurred following the exponential distribution of certain arrival intervals (30, 60, 180, 300, 480 secs). The locations of each dot were random on each screen. Each dot will disappear on the screen randomly after they appear. As soon as the participants detected any red dot on the

screen, they were required to move the cursor using the mouse to click on the screen in which the red dot appeared. Human performances in this simulation are computed and recorded as hit rate, number of false alarm, and reaction time in a 20 minute session, respectively. Hit rate is the ratio between the numbers of detected and appeared red dots. Number of false alarm is computed as false clicks when there was no red dot. Response time is recorded as the time period between red dot appearance and mouse click.

Generally, IST 1.0 is a signal detection task program involving targets monitoring and actions implementing on the temporal and spatial relations of signals. In order to optimize performance, participants need to (a) continually monitor the split screens, and (b) generate a strategy for processing targets, (c) select targets by pointing to them with the mouse, and (d) click on the screen in which the red dot appeared using the mouse as soon as possible. Each participant completed 20 trials (4 screen modes×5 signal frequencies) in random order. In each session, participants were inquired to do the experimental task three times.

APPARATUS

This study was conducted by running a split screens simulated task on Intel Pentium III 601 MHZ based personal computer (PC) linked to a 17-in cathode ray tube (CRT) monitor, a standard keyboard, and mouse. The monitor operated at 120Hz under 1024×768 resolution. To support observational data collection, this study developed a software application, called split screens simulated task, to simulate the monitoring task. The software application was developed by Microsoft Visual Basic 6.0.

EXPERIMENTAL DESIGN

A within-subjects experimental design was used in this study. The independent variables for the experiment comprised the screen mode and signal frequency. There were four settings of the screen modes, including one, four, nine, and sixteen split screens. The multiple screens were of the same size and shape (rectangle) and evenly distributed on the display. The five signal frequencies were one in every 30 seconds, 60 seconds, 180 seconds, 300 seconds, and 480 seconds. Each participant completed 20 trials (4 screen modes×5 signal frequencies) in random order. Each experiment session was followed by a 5-min rest period. In total, all participants completed twenty 20-min sessions for the experiment. The dependent variables for the experiment were hit rates, response times, and number of false alarms. The IST 1.0 simulated program records the hit rate (ratio between the numbers of detected and appeared red dots), number of false alarm (false clicks when there was no red dot), and response time (between red dot appearance and mouse click). Since there are a large number of white dots on each screen during each 20 minute session, the false alarm rate becomes too small to be recorded in valid decimal digit. The program was unable to calculate false alarm rate. Instead, number of false alarm was analyzed.

PROCEDURES

All participants participated in a 5-min familiarization period at the onset of the experimental session. The period was designed to acquaint participants with the procedures and equipment used in the experiment including the mouse and graphics display used in the simulated task. During data collection trials, subjects were required to pay attention on each split screen, which is called channel. Participants were informed in advance that signals might occur during the experiment period. This is also reflective of real-world scenarios where operators of supervised systems are instructed as to their responsibilities. The participant was required to touch the

channel where the signal occurred. If participants detected the signal and touched the right channel, it would be recorded as a hit. It would be recorded as a miss if the participant ignored the red dot and did nothing. False alarm occurred if participants touched the channel without any signals on it.

The simulated program starts to record the response time while the signal occurred. There were five kinds of time periods were recorded, including occurred time, hit time, response time, and missing time. Response time can be seemed as subtract occurred time from hit time. Missing time was the time period from signal occurred to signal disappeared. Subjective vision assessment scale needs to be filled in at the beginning and the end of the experiment. From the comparison of the content of subjective vision assessment scale, the effect of signal detection task on vision can be identified. Each experiment session was followed by a 5-min rest period. In total, each required session required approximately 500 minutes per participant.

RESULT

FOUR MAJOR PERFORMANCE INDEXES OF PARTICIPANTS

Hit rate

Results of an ANOVA on monitoring performance, as measured in terms of the ratio of the number of hits to the total number of signals. Based on the analytic results, hit rates are not significantly influenced by independent variables, those are number of split screens, signal occurred time period, and the interaction of splits screens and signal occurred time period.

Average response time

Results of an ANOVA on monitoring performance measured through the reaction time revealed significant main effects of number of split screens ($F_{3,\ 164}=15.62$, $p=.000$). However, different signal occurred periods are not significantly influenced by independent variables. Additionally, Tukey's HSD test was used to further examine the significant main effects. Orthogonal contrast revealed the average reaction time under single spilt screen to be significantly less ($p<.05$) than sixteen splits screens. In fact, reaction time will get extending while number of split screens increasing.

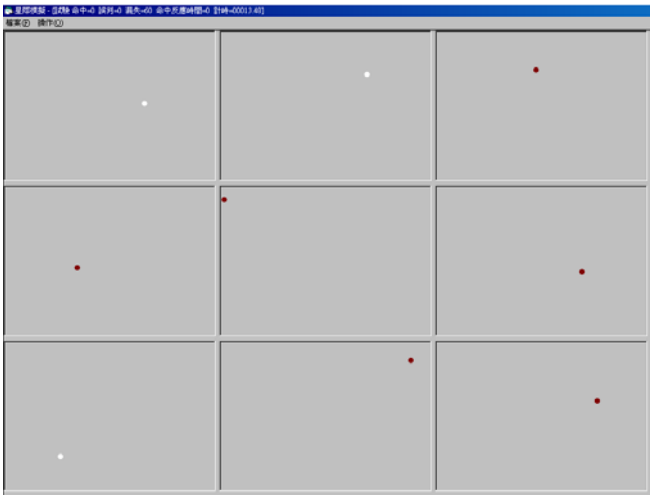


Figure 1. Interstellar Simulator Task 1.0 under nine split-screens

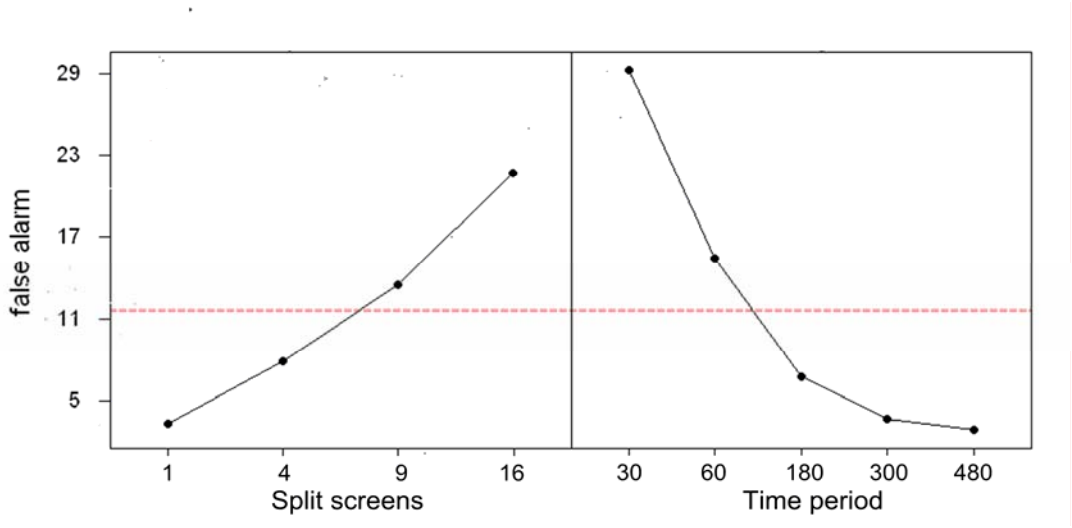


Figure 2. Plot of mean false alarms for the main effects versus the number of split screens (left) and signal occurred time period (right)

False alarm

Results of an ANOVA on monitoring performance measured through the number of false alarm revealed significant main effects of number of split screens ($F_{3, 164}$=63.22, p=.000), and signal occurred time period, ($F_{4, 164}$= 98.93, p=.000). The two-way interaction of split screens×signal occurred period ($F_{12, 164}$=16.08, p=.000) was also found to be significant. Tukey's HSD test was used to further examine the significant interaction and main effects. Orthogonal contrast revealed the main effects of average reaction time and signal occurred time period, as shown in Figure 2.

Subjective vision assessment scale

Results of an ANOVA on the subjective vision assessment scale indicated the following result. First, participants felt eye strain (p<.01) and numb (p<.01) after trials. The results meant that vigilance task will make people feel tired, especially to the eyes. Then, the satisfactions of participants are significantly different (p<.01) with different split screens. Finally, participants felt strange feeling around the eyes (p<.05) while they do experiment in different signal occurred time period.

PARTICULAR SPLIT SCREENS

From the results shown in the above section, split screens are found to be a significant variable while doing a monitoring task. While doing a monitoring task, the operators were trained to pay attention to the right target in the right time. However, operators always pay attention to the screen which can be notice easily. Thus, in order to remind operators to notice the important signals, it is important to find the region which is most easy to find in the screen.

Four rectangular areas

Each rectangular area (channel) in the split screens do not be present significant differences between each other in hit rate and response time. However, the signal occurred time period significantly (p<.01) influences the number of false alarms. Then, based on the main effect analysis, false alarm would be increased while the occurred time period decreased.

Nine rectangular areas

From the result of an ANOVA on the hit rate indicated the following result. Hit rate in different channel was significant different with the others ($P<.001$). Based on the main effects analysis, this study found that the hit rate of channel two, four, five, six and eight are higher than other channels. From the result of an ANOVA on the false alarm indicated the following result. False alarms in different channel was significant different with the others ($P=.033$).

Sixteen rectangular areas

From the results of an ANOVA on the hit rate indicated the following result. Hit rate in different channel was significant different with the others ($P<.001$). Based on the main effects analysis, this study found that the hit rate of channel three, five, six, seven, ten, eleven, twelve, fourteen and fifteen are higher than other channels. Then, the results of an ANOVA on the false alarm indicated the following result. False alarms in different channel was significant different with the others ($P<.001$). Also, based on the main effects analysis, this study found that the false alarms of channel two, six, seven, eight, nine, ten, eleven, and twelve are higher than other channels.

DISCUSSION

From the results shown in section 3.1, this study found that the response time and the false alarms were significant influenced by split screens. However, split screens did not influence the hit rates in this experiment. This study concludes that the criteria of participants is “hit rates are more important than false alarms”. Thus, participants focused on the signal occurrences base on the criteria in mind. That is the reason for the number of split screens got increased, the hit rates still kept high. For the reason, hit rates were not significantly influenced by the split screens. Actually, this result can used to validate the factors affecting response bias level and bias increment proposed by Wickens and Hollands (2000), that is, payoffs will influence response bias as in signal detection task (e.g., Davenport, 1968; See et al., 1997). Therefore, we concluded that the primary task measure, such as hit rates in monitoring task, is not suitable to this experiment.

Signal points and noise points have the same shape in this experiment. Participants were easily to make perception errors for this design of points. So, this study can gathered a significant result of perception errors. This study found that if the number of split screens increased, false alarms would significantly grow up. This result can also used to validate the factors affecting response bias level and bias increment proposed by Wickens and Hollands (2000). Target probability will influence bias. When signal occurrences probabilities getting higher, hits and false alarms getting more. When signal occurrences probabilities getting lower, correct rejections and misses getting more (Loeb and Binford, 1968; See et al., 1997; Williges, 1971).

While measure the participants’ mental workload, the average reaction time is an efficient performance index. In this experiment, the average response time significantly extending while the number of split screens increasing. We concluded the number of split screens significantly impact participants’ mental workload. From the results of subjective vision assessment scale, participants significantly felt uncomfortable with the 16 split screens. They also demonstrated that different screen layout significantly influenced the subjective feeling of monitoring task. Therefore, it is critical to discuss the region differences while monitor specific information. In the following paragraphs, this study discussed the difference influences to the number of split screens and regions while doing a monitoring task.

The signal occurred time period significantly influences the number of false alarms. From the results, this study concluded that while operators monitor four rectangular areas, the performance between different areas will be equal. The results show that when signal occurred frequently, operators would lower the criterion set to increase the hit rate. At the same time, false alarm will also get increasing. The result is the same as the conclusions proposed by Sanders and McCormick (2000), when the criterion is shifted to the left, beta will decrease and the operator will say signal more often and hence will have more hits, but more false alarms. If we mark the five channels, it looked like a cross. Thus, this study found that in nine rectangular areas, in the screen would gather the highest monitoring performance. Thus, if an operator need to monitor nine channels, the most important information should be put at the channels in the cross site.

Also, based on the main effects analysis, this study found that the false alarms of channel four, five, six and eight are higher than other channels. If we mark the four channels, it looked like a "T" letter. Compare the result of hit rate and false alarm, this study found that the channel four, five, six, and eight are different with channel two. Channel 2 has higher hit rate, but lower false alarms. Thus, if an operator needs to monitor nine channels and false alarm will make large loss, the most important information should be put at the channels in the central-upper site. If we mark all the channels, it looked like a cross. Thus, we found that operator always pay attention in the cross zone of the split screen. In the cross zone, hit rate and false alarms are always higher than other channel. Thus, we proposed that if the most important information was put in the cross site, the hit rate will be higher than others.

CONCLUSION

While the time period of signal occurrences increased, the average response time of operators will get decrease. When the number of split screens increased, operators will pay more attention on each channel. False alarms usually lower when the time period of signal occurrences decrease. Oppositely, when the time period of signal occurrences increased, false alarms get increase. Besides, the split screens also significant influences the operators. Operators will pay more attention in the screens while the split screens are getting increased. However, the average response time will get increased because operators need to pay attention on added screens. After operators monitor the screens for a long time, they usually feel uncomfortable, especially to the eyes and head.

To conclude the upper described. Performance of monitoring task is significant influenced by split screens. The performances include hit rate, average response time, false alarms, and subjective vision assessment scale. With the technology advance of technology, digital monitor can be divided into split screens. However, human capability is limited to human information processing. Monitoring system designers should put emphasis on the trade-off of split screens and manning strategy. Besides split screens, layout of screens is also an important factor. Therefore, this study proposed two research results. First, split screens could influence monitoring performance. The best split screens number base on the experimental results is four and nine. Second, in order to get the best monitoring performance, the screen layout should be arranged previously. The most important information should be arranged in the center on the screen. Then, operator will easily getting the critical target.

REFERENCES

DeGangi, G. and Porges, S. (1990). Neuroscience Foundations of Human Performance. Rockville, MD: American Occupational Therapy Association Inc.

Davenport, J. F. (1962). Relationships of intermittent noise, intersignal interval, and skin conductance to vigilance behavior. Journal of Applied Psychology, 46, 106-114.

De Waard, D. (1996). The measurement of drivers' mental workload. Ph.D. thesis, University of Groningen, Traffic Research Centre, Haren, The Netherlands.

Giambra, L. and Quilter, R. (1987). A two-term exponential description of the time course of sustained attention. Human factors, 29(6), 635-644.

Loeb, M. and Binford, J. R. (1968). Variation in performance on auditory and visual monitoring tasks as a function of signal and stimulus frequencies. Perception and Psychophysics, 4, 361-367.

Moray, N. (1979). Mental Workload: Its Theory and Measurement. Plenum Press, New York.

Moray, N. (1988). Mental workload since 1979. International Reviews of Ergonomics, 2, 123-150.

Moray, N. (2003). Monitoring, complacency, scepticism and eutactic behaviour. International Journal of Industrial Ergonomics, 31, 175-178.

Parasuraman, R. Warm, J., and Dember, W. (1987). Overview paper: Vigilance: Taxonomy and utility. In L. Mark, J. Warm, and R. Huston (eds.), Ergonomics and human factors: recent research. New York: Springer-Verlag.

Piechulla, W., Mayser, C., Gehrke, H., Köonig, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. Transportation Research Part F, 6, 233-248.

Sanders, M. S. and McCormick, E. J. (1992). Human factors in engineering and design. McGraw-Hill, New York, USA.

See, J. E., Warm, J. S., Dember, W. N., Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. Human Factors, 39, 14-29.

Sheridan, T. B. (2006). Supervisory control. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. John Wiley & Sons, New Jersey.

Tsang, P., and Wilson, G. (1997). Mental workload. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. Wiley, New York.

Tulga, M. K. and Sheridan, T. B. (1980) Dynamics decisions and workload in multitask supervisory control. IEEE transactions on Systems, Man, and Cybernetics, 10(5), 217-232.

Wickens, C. D. (1984), Engineering psychology and human performance, Columbus, OH: Merrill.

Wickens, C. D. (1992). Engineering psychology and human performance. New York: Harper Collins Publishers Inc.

Wickens, C. D., and Hollands, J. G. (2000). Engineering psychology and human performance. London: Prentice Hall.

Willings, R. C. (1971). The role of payoffs and signal ratios on criterion changes during a monitoring task. Human factors, 13, 261-267.

Yeh, Y.Y. and Wickens, C.D. (1988) Dissociation of performance and subjective measures of workload, Human Factors, 30(1), 111–120.