D44

N79-15632

A MODEL OF HUMAN EVENT DETECTION IN MULTIPLE PROCESS MONITORING SITUATIONS*

Joel S. Greenstein and William B. Rouse

Department of Mechanical and Industrial Engineering
Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

SUMMARY

It is proposed that human decision making in many multi-task situations might be modeled in terms of the manner in which the human detects events related to his tasks and the manner in which he allocates his attention among his tasks once he feels events have occurred. A model of human event detection performance in such a situation is presented. An assumption of the model is that, in attempting to detect events, the human generates the probabilities that events have occurred. Discriminant analysis is used to model the human's generation of these probabilities. An experimental study of human event detection performance in a multiple process monitoring situation is described and the application of the event detection model to this situation is addressed. The experimental study employed a situation in which subjects simultaneously monitored several dynamic processes for the occurrence of events and made yes/no decisions on the presence of events in each process. Input to the event detection model of the information displayed to the experimental subjects allows comparison of the model's performance with the performance of the subjects.

INTRODUCTION

In many systems, the human operator spends much of his time monitoring subsystems for events which call for action on his part. Aircraft, power stations, and process control plants are examples of such systems. As the complexity of these systems increases, the operator becomes responsible for more subsystems of greater variety. There is consequently a greater probability that the operator will encounter situations in which there are more tasks than he can acceptably perform.

One means of maintaining the operator's workload at a satisfactory level is the introduction of automation capable of performing some of the operator's tasks. Models of the operator's task performance would be of use

*This research was supported by the National Aeronautics and Space Administration under NASA-Ames Grant NSG-2119.

in predicting the performance gains to be expected from the introduction of such aids. Further, in systems in which the responsibilities for some tasks are shared by the operator and an automated decision maker, these models might also be used within the system to coordinate the actions of the two decision makers.

Senders [1] and Smallwood [2] have modeled human decision making in multiple process monitoring tasks. Senders postulated that the human monitor samples his displays in a manner which allows reconstruction of the displayed signals. An information theory approach is employed to determine how often and for what duration the human must sample each display. Smallwood proposed that the human operator forms an internal model of the processes he is monitoring and of the environment relevant to his task as a result of his past perceptions of them. A situation is considered in which the operator seeks to detect excursions of instruments beyond threshold values. The operator is modeled as directing his attention to the instrument whose current probability of exceeding threshold (based on the operator's internal model) is greatest. It might be noted, in passing, that the internal model concept discussed by Smallwood is perhaps as appropriate to the design of automated decision makers as it is to modeling the human decision maker. If the automated decision maker is to interact appropriately with the human, it would seem that its internal model of the relevant environment should include a model of the human.

Carbonell [3,4] and Senders and Posner [5] have proposed queueing theory approaches to the modeling of human decision making in multiple process monitoring tasks. Carbonell uses a priority queueing discipline. He assumes that the human operator attempts to minimize the risk involved in not observing other instruments when he chooses to monitor a particular instrument. Senders and Posner employ a first come first served service discipline. They suggest two models which might be used to estimate the inter-observation intervals for an instrument (i.e., the time between arrivals of the instrument to the queue of instruments awaiting observation by the human monitor). The first model involves the degree of the observer's uncertainty about the value of the variable displayed on the instrument. The second model involves the probability that the displayed variable will exceed an acceptable limit.

The models cited above emphasize the monitoring of displays, rather than the decisions or actions that result from the human operator's perception of the displayed values. The operator's motivation for monitoring the displays is the possibility that an event which requires his action will occur. The multi-task decision making problem addressed in this paper concerns the event detection and action selection decisions the operator makes on the basis of the information he gains through monitoring.

Human decision making in such multi-task situations, then, might be modeled in terms of the manner in which the human detects events related to his tasks and the manner in which he allocates his attention among his tasks once he feels events have occurred. Gai and Curry [6] have developed a model of the human monitor in a failure detection task. The model has two stages, the first being a Kalman filter which estimates the states and observations

of the monitored process and the second a decision mechanism which operates on the Kalman filter residuals using sequential analysis concepts. The model can be used to describe the human monitor's detection of additive failures in stationary random processes.

Sheridan and Tulga [7] have modeled the manner in which the human operator allocates his attention among various tasks. They address a situation in which events present themselves unequivocally and use a dynamic programming approach to determine the action sequence which maximizes the operator's earnings. This action sequence is begun, but can be superceded by a new sequence calculated in response to the appearance of additional tasks.

Rouse [8] has investigated the issue of allocation of decision making responsibility between a human operator and an automated decision maker. He presents a mathematical formulation of the multi-task decision making situation appropriate to the modeling of either decision maker. Based on displayed information, the decision maker is assumed to generate probabilities that events have occurred in his tasks. He also generates density functions which characterize his perceptions of what might occur in his tasks while his attention is diverted to a particular task and how long his attention will be diverted should he decide to take a given action. Combining estimates of the probabilities events have occurred with the density functions of time between events in the tasks and action times with respect to the tasks, the decision maker chooses his actions to minimize an appropriate cost criterion. In this paper, we present a model of the human's event detection performance consistent with this mathematical formulation, describe an experimental study of event detection performance in a multiple process monitoring situation, and address the application of the model to the process monitoring situation.

THE EVENT DETECTION MODEL

The event detection model assumes that, in attempting to detect events, the human generates the probabilities that events have occurred. A discriminant analysis approach [9,10] is used to model the human's generation of these probabilities. Our use of discriminant analysis to model the human's generation of event probabilities is motivated by the fact that this approach does not require explicit models of the systems the human is monitoring. An understanding of the systems is certainly helpful in determining the features to extract from the observations. But explicit models of the systems' structures are not required.

For each task i, various features x_{ij}, $j=1,2, \ldots,m_i$, are extracted from the human's task related observations z_i. These features are properties of the observations that characterize (or are believed to characterize) the presence or absence of events related to the task. Following the extraction of a set of features, the value of a linear discriminant function

$$Y_i = v_{i1}x_{i1} + \cdots + v_{im_i}x_{im_i} \quad (1)$$

is calculated. Based on previous experience with the task, estimates are made of the discriminant function coefficients v_{ij}, $j=1,2, \ldots,m_i$, with which to combine the feature values x_{ij} to obtain the discriminant function score Y_i that best differentiates observations of events from the rest of the task related observations. Estimates of the mean and variance of the discriminant function over observations of events and over the rest of the observations are also formed. The a posteriori probability that an event has occurred is generated using the value of the discriminant function score, the estimates of the means and variances of this score over events and "non-events", and an estimate of the a priori probability of the event.

If the human operator is forced to make a yes/no response on the presence of an event, we might assume that he chooses the response which maximizes his expected reward. We can then express his decision in a signal detection manner and state that he should respond "yes, an event related to task i has occurred" if the following inequality holds:

$$\frac{P(e_i/Y_i)}{1 - P(e_i/Y_i)} > \frac{V_{CR_i} + C_{FA_i}}{V_{H_i} + C_{M_i}} \quad (2)$$

$P(e_i/Y_i)$ is the a posteriori probability that an event related to task i has occurred. The value of this probability is generated by the event detection model. V_{CR} is the value of correctly responding "no event" (a correct rejection), C_{FA} is the cost of incorrectly responding "event" (a false alarm), V_H is the value of correctly responding "event" (a hit), and C_M is the cost of incorrectly responding "no event" (a miss).

It is predicted, then, that if the operator is forced to make a yes/no decision on the presence of a task related event, he calculates the likelihood ratio of the event (the left hand side of Eq. (2)). He compares the magnitude of the likelihood ratio with a threshold determined by the values of correct responses and the costs of incorrect responses (the right hand side of Eq. (2)). He responds "event" if the likelihood ratio exceeds the threshold.

THE EVENT DETECTION EXPERIMENT

An experiment has been run employing a situation in which subjects simultaneously monitor several dynamic processes for the occurrence of events and make yes/no decisions on the presence of events in each process. Figure 1 illustrates the display observed by the subjects in the experiment. The static display was generated on a Tektronix 4010 by a time-shared DEC-System 10 and depicts the measured values of the outputs of nine processes over 100 sampling intervals (i.e., 101 points). The processes had identical second order system dynamics with a natural frequency of 0.75 rad/sec and a damping

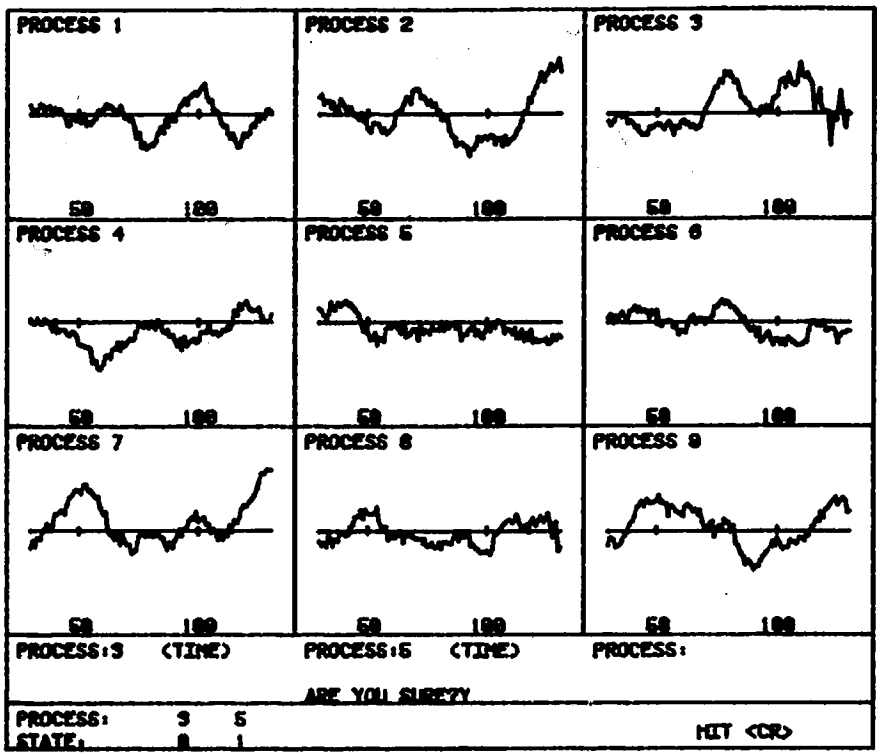

Figure 1. The Multiple Process Monitoring Situation

ratio of 0.5. Samples were taken at 0.2 second intervals. The inputs to the processes were zero-mean Gaussian white noise sequences of identical variance. The displayed measurements were obtained by corrupting the process outputs with additive zero-mean Gaussian white noise sequences which normally had identical variance. The measurement noise variance was normally selected to yield measurements with signal-to-noise ratios of 25.0. An abnormal event in a process was defined by an increase in the measurement noise variance such that the signal-to-noise ratio following an event occurrence was decreased to 95% of the signal-to-noise ratio of the preceding measurement. Thus, abnormal events became more pronounced with each measurement following their occurrence.

After scanning the nine process histories, the subject was given an opportunity to key in the numbers of processes in which he had decided an abnormal event had occurred. He was then given feedback regarding the actual states of the processes he had keyed in ("1" indicating the normal state, "0" indicating the abnormal state). An iteration in a trial was completed by erasing the display, scoring the subject's performance, and returning all abnormal processes detected by the subject to the normal state. Another iteration was then begun by generating a new display depicting the process histories advanced 10 sampling intervals in time as illustrated by Figure 2. (The dashed vertical lines indicated to the subject the point at which he last responded to each process.)

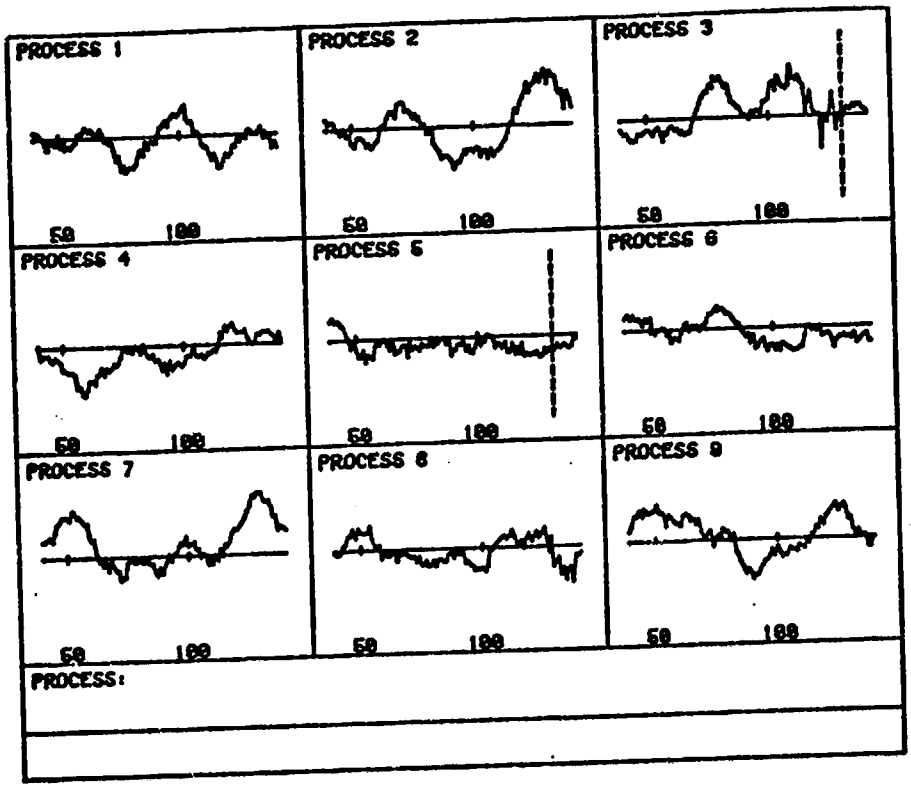

Figure 2. An Updated Display

The subject was allowed to respond to as many events as he thought had occurred. He was awarded points for his hits, receiving high scores for responding to events soon after their occurrence and lower scores for tardier responses. A fixed number of points was deducted for each false alarm. The subject was allowed to study the displays as long as he wished, but any time taken beyond the first minute on each iteration reduced the score awarded for hits made on that iteration.

Eight subjects were given three trials spaced over several days. Each trial was 20 iterations long. The first and third trials given half the subjects were identical, with one event scheduled to occur per iteration. Their second trial scheduled the same events as the first and third trials, but also scheduled an additional event occurrence each iteration. The rest of the subjects were given the same trials in different order so that two events were scheduled to occur per iteration in the first and third trials while one occurrence per iteration was scheduled in the second trial. (Not all scheduled events actually occurred. If an event was scheduled to occur in a process in which a previous event had not yet been detected by the subject, the scheduled event was deleted from the trial.) Events were scheduled to occur uniformly over the nine processes and over the 10 new points displayed for each process on each iteration (the last 10 points on the first iteration, in which all 101 points were new) within the constraint that no two events could occur in a process within 30 sampling intervals of

each other.

Before each trial, the subject was told the average number of new events he could expect to occur per iteration. He was not given any information regarding the dynamics of the processes, but was told that he could expect the processes to exhibit similar characteristics when operating normally. He was also not told what parameter changes defined events, but was told that all events would generally exhibit similar characteristics, and all would become more pronounced as time passed. The subject was given several iterations of training before each trial during which solid vertical lines were included on the process histories to mark exactly when and where events had occurred.

During each trial, the subject was asked to keep a log in which he described his strategies for event detection and noted characteristics of the process measurements he used in his attempts to detect events. After each trial, he was asked to order these characteristics in terms of their usefulness in event detection.

APPLICATION OF THE MODEL TO THE EXPERIMENTAL SITUATION

The event detection model suggests that the human operator in the experimental situation just described extracts various features from his observations of the process measurements. He attempts to select features which characterize the presence or absence of task related events. Through his experience with the processes, the operator has formed estimates of the discriminant function coefficients with which to combine the features to obtain a discriminant function score. He has also formed estimates of the means and variances of this score over observations of events and over the rest of his observations. The operator generates the likelihood ratio that an event has occurred based on the value of the discriminant function score. his estimates of the means and variances of the score, and his estimate of the a priori probability of an event occurrence. He compares the likelihood ratio with a threshold that is based solely on the values of correct responses and the costs of incorrect responses and responds "event" if the likelihood ratio exceeds the threshold.

Four features of the process measurements were selected for use with the event detection model. Selection of these features was guided by the comments of the experimental subjects regarding the characteristics of the process measurements they found useful in event detection. The first feature involves the magnitude changes between successive measurements in a sequence of the most recent measurements. The second feature involves the presence of reversals in direction in this sequence (changes from positive slope to negative, or vice versa, of the line segments connecting the measurements of the sequence). The third feature tests for the simultaneous occurence of large magnitude changes and reversals. The fourth feature, like the first, is a measure of magnitude changes, but it is much more local in that it

involves only the four most recent measurements of the process output.

In extracting these features from the process measurements, the values of the features over recent measurements are weighted more heavily than the values over earlier measurements. The weight decreases exponentially with the age of the measurement and the rate of this decrease is a free parameter. The value of the first feature, for example, a measure of the magnitude changes between successive measurements in a sequence of the n most recent measurements of a process' output, is given by

$$x_1 = \{ \sum_{k=1}^{n-1} |z(k+1)-z(k)| \cdot \exp[-\beta(n-1-k)] \} / \sum_{k=1}^{n-1} \exp[-\beta(n-1-k)] \qquad (3)$$

where z(k) is the k th measurement in the sequence, z(n) is the most recent measurement, and β is the free parameter governing the relative weighting of the feature's value over recent and earlier measurements in the sequence.

In the generation of the likelihood ratio of an event in a process at a given iteration of an experimental trial, the sequence of process measurements over which the features are calculated ends with the last measurement displayed for the process on that iteration. The cutoff length used in extracting the features from the process is a free parameter. Values of the features over process measurements taken earlier than the cutoff are not calculated (or, effectively, are assigned zero weight). If the subject responded "event" to the process at some point following the cutoff, then features are calculated over only those measurements occurring after this response. The information on the state of the process that the subject gains when he responds to the process motivates this constraint. If the process is in the normal state, then on succeeding iterations the subject knows that if an event has occurred, it must have occurred following his last response. If the process is in the abnormal state, then the process is reset to normal when the subject keys in his response. On succeeding iterations the subject knows that if another event has occurred in the process, it must have occurred following his last response. In either case, the subject (and the model) should calculate features only over measurements occurring after the subject's last response.

The estimation of discriminant function coefficients requires a representation of normal and abnormal process measurements. This representation was formed using the process histories displayed to the subject on his third experimental trial. The process histories are separated into two groups of sequences - normal and abnormal. Sequences of measurements beginning when a process was returned to the normal state and ending when an event occurred are defined to be normal. Sequences of measurements beginning when an event occurred and ending when the process was returned to the normal state are defined to be abnormal. The values of the four features were calculated over the entire length of each of the sequences in the two groups. A discriminant analysis was then performed on the resulting two groups of feature values to determine the discriminant function coefficients v_j, j=1,2,...,m, with which to combine the features to best differentiate between the two groups. The mean value and the variance of the

resulting discriminant function scores for the sequences in each of the two groups was also calculated.

The final requirements for application of the event detection model to the experimental situation are estimates of the a priori probabilities of event occurrences and the selection of a threshold against which likelihood ratios of events can be compared. For experimental trials in which one event was scheduled to occur per update of the display over the nine processes monitored by the subject the a priori probability of an event occurrence in each process was fixed at 1/9. For trials in which two events were scheduled to occur per display update, the a priori probability was fixed at 2/9. The threshold against which the likelihood ratios of events are compared is assumed to remain constant through an experimental trial. The magnitude of this constant is a free parameter.

RESULTS

Figure 3 compares the event detection performance of the model with the actual performance of each of the eight subjects in the third trial of the experiment. In this trial, 20 events were scheduled to occur in the trials given subjects A,B,C, and D, while 40 events were scheduled to occur in the trials given subjects E,F,G, and H. In applying the model to each of these trials, the number of measurements over which features were extracted (cutoff

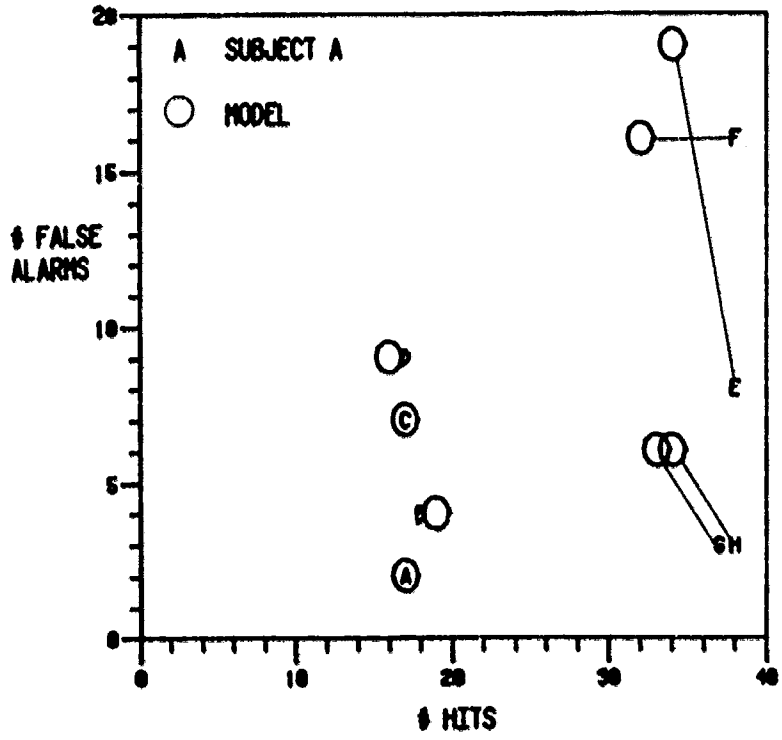

Figure 3. Comparison of Model with Subjects on Third Trial

length) and the relative weighting of recent and older points (β) were adjusted to improve the fit of the model's performance to each subject's performance. The value of the threshold against which likelihood ratios of events were compared was also adjusted to improve the fit. Figure 3 reveals a high degree of correspondence between the model's performance and the performance of most subjects.

Figure 4 compares the event detection performance of the model with the actual performance of the eight subjects in the second trial of the experiment. In this trial, 40 events were scheduled to occur in the trials given subjects A,B,C, and D, while 20 events were scheduled to occur in the trials given subjects E,F,G, and H. In applying the model to each of these trials, none of the parameters of the model were changed from the settings used to obtain the results presented in Figure 3. Despite the fact that the numbers of events scheduled in these trials differ from those in the trials used to assign the values of the parameters, the correspondence between the model's performance and the performance of most subjects is reasonable.

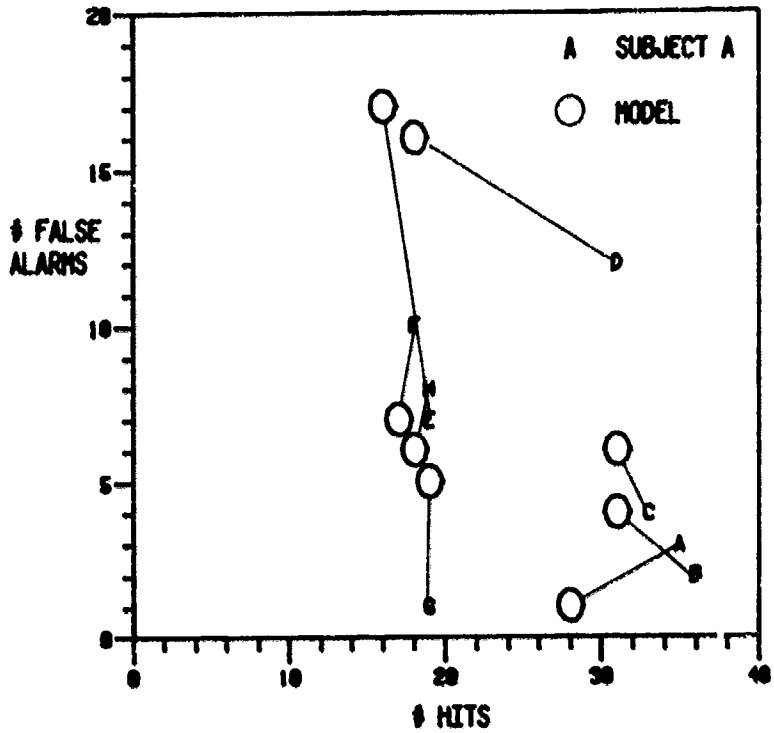

Figure 4. Comparison of Model with Subjects on Second Trial

Table 1 compares the mean detection times (in terms of the number of sampling intervals which elapsed from the occurrence of an event to the time of its detection) for hits common to both subject and model in the trials presented in Figures 3 and 4. It should be noted that the fact that the mean detection times of the model are consistently smaller than those of the subjects is an artifact of the manner in which the model's performance was investigated. The model was tested on the process histories displayed to a subject in his experimental trial. In these trials, a process was returned to the normal state at the point at which the subject detected an event in

the process. Thus, in going over the process histories the model can never respond to an event later than the subject responded to it. If the model fails to respond to an event by the time of the subject's response, the model is scored as having missed that event.

Table 1. - Comparison of Mean Detection Times

	Trial 1		Trial 2	
Subject Code	Subject	Model	Subject	Model
A	24	19	24	20
B	26	21	25	19
C	28	17	28	17
D	32	21	30	16
E	18	13	22	17
F	17	14	27	16
G	20	17	27	20
H	20	17	22	18

We plan to evaluate the model in the near future using a somewhat different approach. Rather than running the model over the process histories displayed to a subject on an earlier experimental trial (and constraining the timing of the model's responses by the timing of the subject's responses in that trial), we will use the model in place of the subject in the event detection experiment. Processes in which events occur will then remain in the abnormal state until the model responds to the process The only constraint on the model's detection times will be the end of the experimental trial. Because the model's detection time for each event need no longer be less than or equal to the subject's detection time for that event, we expect that, for a given number of hits, the model's threshold can be raised to achieve the longer mean detection times and smaller numbers of false alarms characteristic of the subjects in the experiment.

CONCLUDING REMARKS

In applying the event detection model to the experimental situation described in this paper, we studied a situation in which the subject was forced to respond yes or no to the possibility of an event related to each of nine processes. In general, the human operator is not forced to make such yes/no decisions with respect to each of his tasks. Instead he uses his estimates of the probabilities of task-related events (which the event detection model generates) in deciding how to allocate his attention among his tasks. We plan to run an experiment investigating the human's attention allocation performance in a multiple process monitoring situation similar to

the one employed in the event detection experiment discussed here. Data from this experiment will be used to develop and validate a model of attention allocation performance in multi-task situations. (The modeling of human attention allocation performance in multi-task situations is considered in [11].) This model might be used in conjunction with the event detection model as a part of the design process for, and the implementation of, automated decision making systems.

REFERENCES

1. Senders, J. W.: "The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems," IEEE Trans. Human Factors in Electronics, vol. HFE-5, no.1 pp.2-5, Sept. 1964.

2. Smallwood, R. D.: "Internal Models and the Human Instrument Monitor," IEEE Trans. Human Factors in Electronics, vol. HFE-8, no.3, pp. 181-187, Sept. 1967.

3. Carbonell, J. R : "A Queueing Model of Many-Instrument Visual Sampling," IEEE Trans. Human Factors in Electronics, vol. HFE-7, no.4, pp. 157-164, Dec. 1966.

4. Carbonell, J. R.; Ward, J L.; and Senders, J. W : "A Queueing Model of Visual Sampling: Experimental Validation," IEEE Trans. Man-Machine Systems. vol. MMS-9, no.3, pp.82-87, Sept. 1968.

5. Senders, J. W ; and Posner M. J. M.: "A Queueing Model of Monitoring and Supervisory Behavior," in T. B. Sheridan and G. Johannsen, Eds., Monitoring Behavior and Supervisory Control, New York: Plenum 1976, pp.245-259.

6. Gai E. G.; and Curry, R. E.: " A Model of the Human Observer in Failure Detection Tasks," IEEE Trans. Systems, Man. and Cybernetics, vol. SMC-6, no.2, pp.85-94, Feb. 1976.

7. Sheridan, T. B.; and Tulga, M. K.: "A Model for Dynamic Allocation of Human Attention Among Multiple Tasks," Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 1978.

8. Rouse, W. B.: "human-Computer Interaction in Multitask Situations," IEEE Trans. Systems, Man, and Cybernetics, vol. SMC-7, no.5, pp. 384-392, May 1977.

9. Tatsuoka, M. M.: Multivariate Analysis. New York: Wiley, 1971.

10. Afifi, A. A.; and Azen, S. P.: Statistical Analysis. New York: Academic Press, 1972.

11. Rouse, W. B.; and Greenstein, J. S.: "A Model of Human Decision Making in Multi-Task Situations: Implications for Computer Aiding," Proceedings of the 1976 International Conference on Cybernetics and Society, Washington, Nov. 1976.