Systems with Human Monitors:
A Signal Detection Analysis

Robert D. Sorkin

Purdue University

David D. Woods

Westinghouse Research and Development Center

ABSTRACT

Automated factories, the flightdecks of commercial aircraft, and the control
rooms of power plants are examples of decision-making environments in
which a human operator performs an alerted-monitor role. These human-
machine systems include automated monitor or alerting subsystems operating
in support of a human monitor. The automated monitor subsystem makes
preprogrammed decisions about the state of the underlying process based on
current inputs and expectations about normal/abnormal operating conditions.
When alerted by the automated monitor subsystem, the human monitor may
analyze input data, confirm or disconfirm the decision made by the automated
monitor, and take appropriate further action. In this paper, the combined au-
tomated monitor–human monitor system is modeled as a signal detection sys-
tem in which the human operator and the automated component monitor par-
tially correlated noisy channels. The signal detection analysis shows that
overall system performance is highly sensitive to the interaction between the
human's monitoring strategy and the decision parameter, C_a, of the automated
monitor subsystem. Usual design practice is to set C_a to a value that optimizes

Authors' present addresses: Robert D. Sorkin, Department of Psychological Sci-
ences, Purdue University, West Lafayette, IN 47907; David D. Woods, Westinghouse
Research and Development Center, 1310 Beulah Road, Pittsburgh, PA 15235. Re-
quests for reprints should be sent to R. D. Sorkin.

CONTENTS

the automated monitor's detection and false alarm rates. Our analysis shows that this setting will not yield optimal performance for the overall human-machine system. Furthermore, overall system performance may be limited to a narrow range of realizable detection and error rates. As a result, large gains in system performance can be achieved by manipulating the parameters of the automated monitor subsystem in light of the workload characteristics of the human operator.

1.   SYSTEMS WITH AUTOMATED AND HUMAN MONITORS

One component of human-machine systems in complex decision-making environments, such as power plant control rooms, air traffic control, or commercial aircraft flightdecks, is the alerted-monitor function. An alerted-monitor system consists of an automated-monitor subsystem operating in conjunction with a human monitor. The automated alerting subsystem makes preprogrammed decisions about potentially troublesome situations by comparing incoming data on process state with stored expectations of normal and abnormal operating conditions. When alerted by an output from the automated subsystem, the human monitor may analyze input data, confirm or disconfirm decisions/actions made by the automated subsystem, or take appropriate further action. For example, in the computerized air traffic control system (Goldmutz, Kefaliotis, Kleiman, Rucker, Schuchman, & Weathers,

1981), there are subsystems that will monitor incoming flight data, predict potential air traffic problems, and call the human monitor's attention to these problems. This paper uses the techniques of Signal Detection Theory (see Green & Swets, 1966) to model and explore the performance of the joint automated-monitor–human monitor system. The analysis shows that overall alerted-monitor performance is highly dependent on the interaction of the parameters of the automated alerting subsystem and the operator's workload and monitoring strategy.

Traditionally, automated monitors have consisted of messages about simple set-point (preset threshold value) violations. When the measured condition exceeds (or drops below) this level, a message is displayed to the operator. Experience has shown that operators frequently have problems identifying, prioritizing, and responding to abnormal conditions with this type of alarm system (Banks & Boone, 1981; Cooper, 1977; Kragt & Bonten, 1983). For example, one review of power plant control rooms found that:

> Operators at all plants complained about the high number of nuisance alarms. The reasons for their occurrence varied. At one plant there was a blank, supposedly nonfunctional, annunciator window that would occasionally alarm. The maintenance and operational people had been unable to determine its cause, but an acknowledgement, silence, and reset were required on each occasion. In many cases alarm set-points were known by operators to be too sensitive to normal transients. As a consequence slight deviations or transients, thought of as normal, would set the alarm off even though no further operational action was required. Maintenance or calibration operations often caused recurring alarms that were a nuisance. The net results of the many false alarms is a "cry-wolf" syndrome which leads to a lack of faith in the system and a causal attitude towards the constant presence of certain alarms. On many occasions operators were observed to casually silence and acknowledge alarms without further concern or surveillance of plant status; the alarms had become "old friends." (Seminara, Gonzalez, & Parsons, 1977, p. 12)

As a result, there is considerable interest in applying computer technology to develop new types of automated monitoring systems that can improve the detection and correction of abnormal occurrences. One approach focuses on improved organization and presentation of alarm data, cf., the alarm presentation techniques described in Berson, Po-Chedley, Boucek, Hanson, Leffler, and Wasson (1982), Thompson (1981), and Visuri (1983). Another avenue is the development of automated decision systems that monitor multiple parameters on a number of input channels and use more complex (algorithmic or potentially heuristic) information processing to identify potential abnormal conditions, for example, the alarm analysis techniques described in Bastl and

Felkel (1981) (cause-consequence models); Bray, Nelson, Blackman, and Fowler (1984) (rule-based systems); Gimmy and Nomm (1982) (decision tables); Lind (1981) (function-based systems). While there is considerable activity in the design of automated monitors, essentially no theoretical or empirical information is available to guide designers with respect to how human or machine subsystem characteristics affect the performance of the overall alerted-monitor system.

What subsystem factors might affect overall system performance? To explore this question, let us consider a simple example of an alerted-monitor system: the familiar smoke detector. Smoke detection systems vary in their ability to discriminate fire from nonfire conditions. In addition, different consequences are associated with the different types of possible outcomes — failure to detect and respond to an actual fire has much higher negative consequences than a false alarm. Therefore, the designer can vary the amounts of evidence required for a "fire" decision on the part of a smoke detector to take into account the relative consequences of the different possible outcomes. In this example, designers might reasonably require relatively little evidence of a fire before outputting an alarm. This strategy will maximize the smoke detector's hit rate (correct detection of actual fires), but it will also increase the detector's false alarm rate (and the overall number of alarms) for a fixed ability to discriminate fires from nonfires. But consider the consequences of a high false alarm rate on the performance of the subsequent human monitor. A busy human monitor may soon learn to ignore the smoke detector's alarm signal, considering it a false alarm and not worthy of a shift in attention from more pressing duties. The performance of the overall smoke detector–human monitor system would be worse than if the smoke detector were set to emit fewer alarms.

2.   SIGNAL DETECTION ANALYSIS OF ALERTED-MONITOR SYSTEMS

The smoke detector example points out the need to understand how subsystem characteristics determine overall system performance. To begin to provide guidance for the design of alerted-monitor systems, we explore some possible interactions between the automated- and human monitor subsystems. This analysis is based on the techniques of Signal Detection Theory; the results show that the automated monitor's response criterion and the human monitor's workload and time-sharing strategy are important determinants of overall alerted-monitor system performance.

2.1. Signal Detection Systems

Systems that monitor selected process data (a noisy input channel) for potentially abnormal conditions (signal events) can be modeled as signal detection devices. These devices must discriminate between the presence of signal-

plus-noise (abnormal condition) events and noise-alone (normal condition) events. Figure 1 illustrates such a signal detection system. This system first makes a measurement (an observation), which can be considered as a multidimensional input vector, X, and then computes a unidimensional statistic, Z, based on the observation vector and on stored information about the expected characteristics of signal and nonsignal events.

The statistic, Z, is then evaluated against some criterion value, C, which is the output threshold or set-point. If the statistic is greater than or equal to C, the output will be "Yes, there is a signal." If not, the response will be "No, there is no signal." The criterion, C, is usually chosen on the basis of the prior probability of signal (or noise) and the costs and benefits of the possible decision outcomes (correct detection, missed signal, false alarm, correct rejection).

How well the detection system can discriminate between signal-plus-noise and noise-alone events depends on the probability density distributions of the statistic, Z (Figure 2). The probability (density) of obtaining a particular value of Z given a signal-plus-noise event is f (Z/SN); the probability (density) of obtaining a particular value of Z given noise-alone is f (Z/N). If these distributions are widely spaced with low variance, there will be few errors in attributing a given value of Z to a signal-plus-noise or noise-alone cause. Computation of a good statistic requires prior knowledge of what input vectors will usually be generated by signal-plus-noise or noise-alone events. (An optimal statistic for this process is the likelihood ratio.) The assumption that the distributions on Z are normal (Gaussian) in form is a good approximation to the behavior of many detection systems.

Sensitivity (d') and Criterion (C) Parameters

Signal detection systems can be represented by two parameters: (1) a sensitivity parameter, d', which specifies how effectively the system can discriminate signal-plus-noise events from noise-alone events on the channel, and (2) a response criterion parameter, C, which specifies how much evidence is required for a Yes decision. The d' parameter is defined as the normalized separation between the distributions; that is, the difference between the means of the distributions, μ_{sn}-μ_n, divided by the standard deviation of the distributions (assuming that the distributions on Z have equal variance). Values of d' near zero will yield performance at chance levels; levels of d' above 4.0 will yield essentially errorless performance. It is convenient to consider the decision dimension Z as normalized, so that the standard deviation of the distributions is equal to one and the separation of the means is equal to d'. It follows that d' and C are expressed in standard deviate units.

The second parameter, C, is the criterion value used to partition Z into the response categories of Yes or No. The important aspect of C is its position relative to the means of the two distributions. The probability of the system correctly responding Yes given that a signal has occurred, $P(Y/SN)$, is called the hit probability. The probability that the system incorrectly responds Yes given

Figure 1. The basic signal detection system. (See text for discussion.)
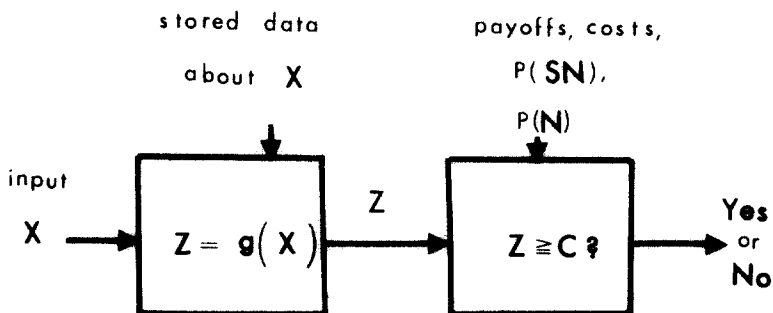
Figure 2. Probability distributions of Z, the decision statistic. Values of Z greater than C lead to a Yes response.
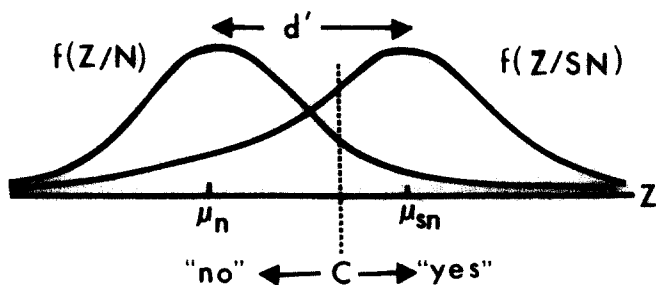
that no signal has occurred, $P(Y/N)$, is called the false alarm probability. These probabilities are computed from the density distributions on Z as follows:

$$P(Y/SN) = P[(Z \geq C)/SN] = {}_C\int^{\infty} f(Z/SN)\, dz \qquad (1)$$

$$P(Y/N) = P[(Z \geq C)/N] = {}_C\int^{\infty} f(Z/N)\, dz \qquad (2)$$

Suppose that $d' = 2.0$, that is, $\mu_n = 0$ and $\mu_{sn} = 2.0$. If C were equal to 1.0, it would be centered between the means of the two distributions. The resulting hit and false alarm probabilities would be, from a table of the normal distribution, .841 and .159, respectively. For a particular value of d', one can evaluate the hit and false alarm values resulting from any value of C. Figure 3 gives hit and false alarm rates for some representative values of d' and C.

Moving C along the Z dimension toward the noise distribution (smaller values of C) yields both higher hit and higher false alarm rates. Moving C toward the signal-plus-noise distribution (higher values) yields lower hit and lower false alarm rates. A small value for C implies that the system is liberal at

Figure 3.   Hit [$P(Y/SN)$] and False Alarm [$P(Y/N)$] probabilities for selected values of d' and C.

d'	C	$P(Y/SN)$	$P(Y/N)$
2	0	.977	.500
2	1	.841	.159
2	2	.500	.023
1	0	.841	.500
1	0.5	.691	.308
1	1	.500	.159

responding Yes; a high value of C implies a conservative criterion for responding Yes.[1]

Receiver Operating Characteristic

The performance of a signal detection system can be characterized by a Receiver Operating Characteristic (ROC), which is a plot of the hit probability, versus the false alarm probability, for all possible values of C and a fixed value of d'. The $d' = 1$ and $d' = 2$ curves of Figure 4 are thus complete plots (all values of C) of the hit and false alarm values shown in Figure 3. The curve labeled 3 is the plot of hit and false alarm values for all values of C for the more sensitive system with a d' equal to 3.0. It is evident that systems with higher d' values will have higher hit rates and lower false alarm rates. The higher the d', the farther the ROC will be from the positive slope diagonal (which is the ROC for $d' = 0$, or chance performance). Since each ROC curve is computed with a fixed d', each curve represents a constant ability to discriminate signal-plus-noise from noise-alone. However, the particular operating point on an ROC (a particular hit and false alarm pair) reflects the liberal or conservative aspect of the system's response criterion, C. As C shifts toward smaller values (move from point 0,0 toward 1,1 along an ROC curve in Figure 4), there is an increase in both hit and false alarm rates and in the overall (Yes) response rate. For single-stage detection systems, the best operating point on a given ROC (i.e., the optimal C value) depends on the costs and benefits of the various outcome categories as well as the prior probability of a sig-

---

[1] In the signal detection literature it is customary to define the response criterion parameter as β rather than C; β is the ratio of the ordinates of the decision distributions when Z equals C. That is, $\beta = [f(C/SN)/[f(C/N)]$. We use C in this paper for ease of exposition of the theory.

Figure 4.   Receiver Operating Characteristics for d' values of 1, 2, and 3.
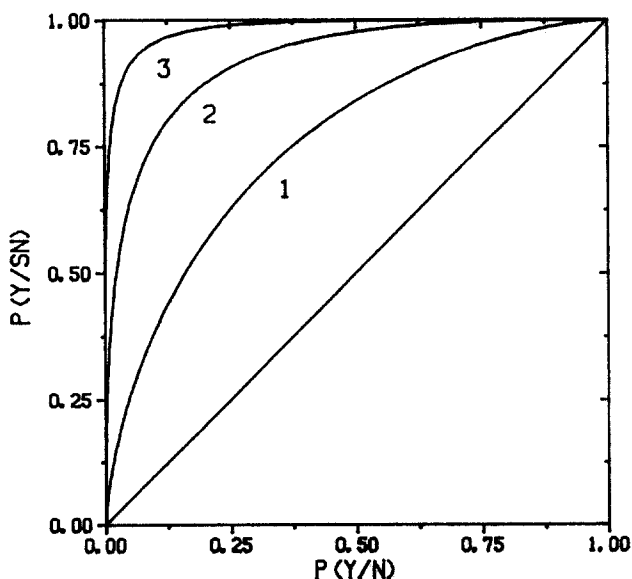
nal occurring (see Egan, 1975). For example, in some situations it will be desired to maximize total percentage correct, while in other applications it will be necessary to minimize the number of missed signals or false alarms. Different decision strategies prescribe different settings of the C parameter (Green & Swets, 1966; Sheridan & Ferrell, 1974; Swets, 1964).

## 2.2. General Model of the Alerted-Monitor System

An alerted-monitor system consists of an automated monitoring subsystem operating in conjunction with a human monitor. An automated subsystem monitors a noisy input channel (i.e., selected process data) on which occasional signal events occur (i.e., potential abnormal conditions). When alerted by subsystem output, the human monitor may analyze input data, confirm or disconfirm the decisions/actions made by the automated monitor, or take appropriate further action. The operator also may be busy with tasks not related to the monitor task.

In signal detection terms the alerted-monitor system can be represented as the two-stage detection system shown in Figure 5. Each stage is specified by its own sensitivity (d') and response criterion (C) parameters.[2] A statistic, Z_a,

---

[2] There is evidence supporting modeling of human monitors as a signal detection system in a variety of different tasks and work environments, including industrial inspection, supervisory control, sonar/radar, and medical image systems (e.g., Bisseret, 1981; Davies & Parasuraman, 1981; Kantowitz & Sorkin, 1983; Swets & Pickett, 1982).

Figure 5. Signal detection model of an alerted-monitor system. The inputs to the automated monitor (X_a) and the human monitor (X_h) are subsets of the input vector, X. An output from the automated monitor subsystem may cause the human operator to make an observation on the input channel.
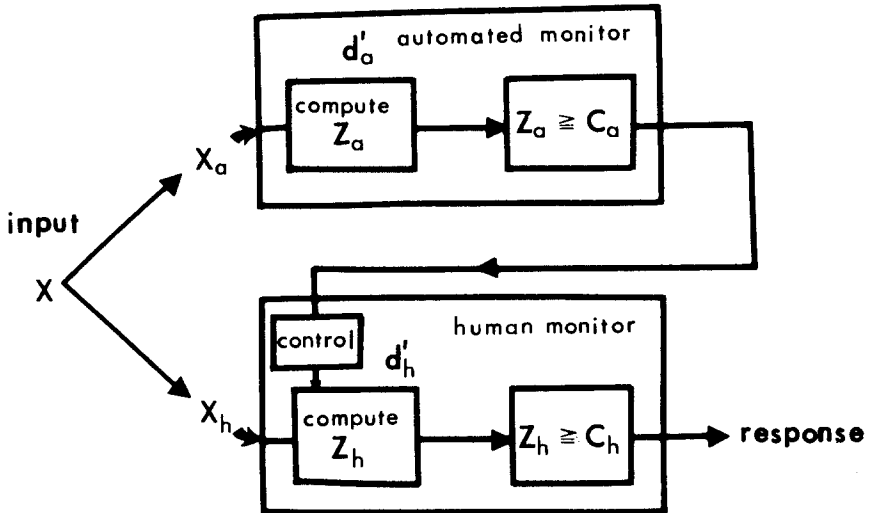
is computed by the automated monitor based on its input, X_a. The magnitude of Z_a relative to the criterion, C_a, determines whether or not an alarm is sent to the human monitor control stage. If an alarm is sent, the human monitor may compute a statistic, Z_h, based on its input, X_h. If Z_h is greater than or equal to C_h, a Yes (confirming) system response will be made. It thus takes a Yes response from both the automated and human subsystems for a Yes system response to any input, X.

The human and automated monitors do not necessarily observe precisely the same noisy input vector, X. The data input, X_h, to the operator and the input, X_a, to the automated subsystem are subsets of X, and each may include elements not available to the other system. For example, the operator may have historical or contextual data unavailable to the automated system, while the automated subsystem may include complex data-processing algorithms or heuristics (e.g., expert systems) that the human cannot duplicate within task constraints. Differences in sampling rates, precision of measurement or display, or offset in the time of observation by each subsystem also decrease the overlap between the channels monitored by the human and automatic systems. In signal detection terms, the two subsystems can be described as monitoring channels that possess some statistical correlation or, equivalently, some degree of common noise. The magnitude of the common noise source, compared to the total noise in each channel, defines the correlation between the channels (Jeffress & Robinson, 1962).

In actual systems one would expect to find the correlation between the automated and human channels to range from 0 to 1, depending on the particular application. For example, a smoke detector may continuously test the conductivity of the air near the ceiling of a room, while a human's observation may be based on a visual check for smoke in lower areas. The correlation between these types of observations may be fairly low. In an automated chemical plant, the same unprocessed sensor data might be available to both the human and the automated monitor; in that case the correlation between those channels could be very close to one.

While the literature includes descriptions of many situations where responses to alerting signals were delayed or omitted (e.g., Kantowitz, 1982; Reason & Mycielska, 1982; Seminara, Gonzales, & Parsons, 1977), little data are available describing the effect of variations in alerting subsystem parameters on the behavior of the operator who must receive and process the output. Reviews of problems with actual alarm systems suggest that operator behavior depends on assumptions about the alerting system's sensitivity, its response criterion (false alarm rate), and the relative importance and likelihood of the condition being signaled. For example, airline pilot ratings of the acceptability of flightdeck alerting signals depend on the perceived false alarm rates as well as the perceived urgency or importance of the alerted condition (Williams & Simpson, 1976).

The effect of alerting system characteristics on the human monitor's performance may also depend on the operator's workload (monitor task load plus other task loads). Prediction of the accuracy and speed of human performance on one task in a multitask environment has been a major concern (Enstrom & Rouse, 1977; Whitaker, 1979; Wickens, 1979). The alerted-monitor task can be considered a special case of the multitask or time-sharing situation. While there is some controversy about the nature of limitations in human information processing in different types of multitask situations (Allport, 1980; Kantowitz & Sorkin, 1983; Lane, 1982; Norman & Bobrow, 1975), we would expect workload factors to be important determinants of the interaction between the automated and human monitor subsystems. In the next sections we analyze the performance of the alerted-monitor system as a function of the key subsystem parameters of alerting subsystem response rate and operator time-sharing strategy.

2.3. Alerted Monitor Operating Characteristics

The performance of a two-stage detection system, such as shown in Figure 5, can be analyzed by determining the effect of various assumptions about the parameters of each of the detection subsystems on the overall system operating characteristic. To get an idea of the possible range of system performance, consider a hypothetical two-stage detection system in which the two stages possess

equal detectability and the observations made by each stage are optimally combined for a system decision. The overall detectability of such a system is given by $\sqrt{2}\ d'$, where d' is the detectability of each stage.[3] In the alerted-monitor system the observations are not combined optimally because the alarm is a binary signal to the operator, and because the operator usually observes the input channel only when alerted by the automated monitor stage.[4] Hence, overall performance in the alerted-monitor system generally will be less than the optimum value of $1.414\ d'$.

The performance of a two-stage detection system, in which binary responses from the two stages are combined for a system decision, was examined by Pollack and Madans (1964). The alerted-monitor system is a special case of the Pollack and Madans two-stage system, in which the second stage (the human operator) observes the noisy channel only when alerted by a signal from the first stage (the automated monitor). Pollack and Madans found that the maximum system d' occurred for such a system when the two stages had equal detectabilities and equal response criteria. They found that over an intermediate range of response criteria, maximum system performance was approximately equal to $1.2\ d'$, a possible 20% performance advantage for the two-stage system.

Figure 6 illustrates the effects that different response criteria have on system performance for an alerted-monitor system with equal subsystem d's, $d'_a = d'_h = 2.0$. Each curve is the overall system operating characteristic for all possible operator criteria and a fixed setting of the automated monitor's criterion, C_a. For example, a stringent C_a value of 1.5 results in few false alarms (and few correct detections). Each curve starts at the origin (0,0), moves toward (1,1), and suddenly terminates.[5] Recall that these system operating characteristics were computed by fixing the parameters d'_a, d'_h, C_a and by sweeping through all values of C_h. When C_h is sufficiently small, the human operator is so liberal that he/she says Yes whenever any observation is made. At that point, the human simply mimics the behavior of the

---

[3] If there were n stages, the overall detectability would be $\sqrt{n}\ d'$. This relation follows from the assumption that the noise or variability of each stage is equal and statistically independent. The standard deviation of a detection statistic based on n independent observations is thus $1/\sqrt{n}$ times that based on one observation.

[4] We recognize that in actual work environments the human operator may also monitor the input channel independent of outputs from alerting subsystems. For example, in one nuclear power plant critical incident (U.S. Nuclear Regulatory Commission, 1983), the operators correctly detected abnormal plant conditions requiring operator action, independent of automatic monitoring subsystem output.

[5] The curves are not proper receiver operating characteristics in the decision theory sense, since $P\ (Y/SN)$ does not increase from (0,0) to (1,1) with monotonically decreasing slope (see Tanner & Sorkin, 1972).

Figure 6. System operating characteristics for $d'_a = d'_h = 2$. Each curve is the system operating characteristic for a fixed value of C_a, evaluated over all possible values of C_h. The highest curve is for $C_a = -0.5$; C_a increases in steps of 0.5 for the lower curves. (C_a is in d' or standard deviate units.) The outer envelope of the curves approximates the receiver operating characteristic for a detection system with a $d' = 2.4$.
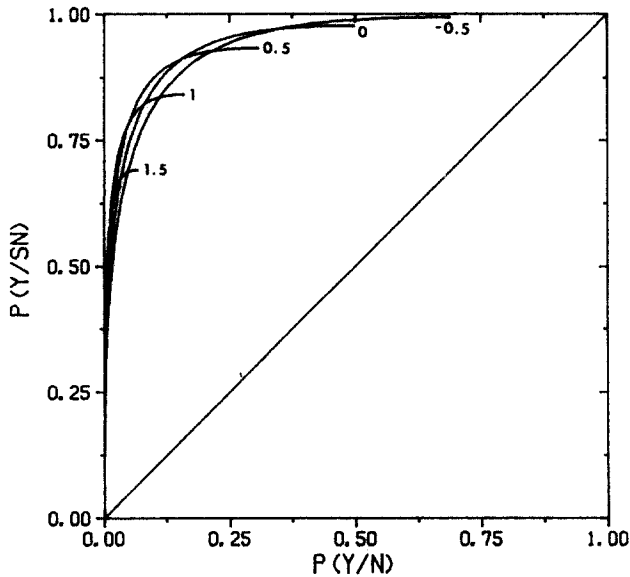
automated stage; that is, given no alarm the human says No, and given an alarm, the human automatically says Yes. This condition exists in the region where each operating characteristic terminates. Hit and false alarm rates higher than these termination values are not achievable. Because the system is mimicking the automated stage at these points, the locus of all the termination points is in fact the ROC for the automated stage alone. It follows from examination of the operating curves that the lowest system d' will be equal to the d' of the automated stage alone.

The outer envelope of the system operating curves in Figure 6 represents the highest d' available with the alerted-monitor system. This envelope is approximately equivalent to an ROC for a system with a d' slightly more than 1.2 times the individual subsystem d's, as shown by Pollack and Madans. In cases where the between-channel correlation is greater than zero, the same general relationships hold, but there is a smaller gain in the combined system d'. Unequal subsystem d's also result in lower system d's (see also Swennson, 1980, for a related analysis).

Pollack and Madans also collected data on the performance of combined person and machine (simulated) monitors using an auditory signal detection

task. Their study employed a set of six observers whose response ratings were pooled on each trial. They found that the performance of the combined observers/machine system was generally superior to that of the observers alone, but less than that predicted by the optimal combination rule. This discrepancy was especially evident at high signal-to-noise ratios.

2.4. Automated Monitor — Human Monitor Interactions

In this section we investigate the effects on system performance of different forms of interaction between the human operator's parameters (d'_h, C_h) and the automated monitor's criterion (C_a). The higher is d'_a (and d'_h), the better will be overall system performance. Obviously, the system designer will want to maximize d'_a. This parameter will generally be limited by automated detector hardware considerations and by lack of knowledge of the to-be-detected events, such as the particular patterns of sensor data that correspond with abnormal events. However, the automated monitor's response criterion, C_a, is normally *not* limited by similar considerations; it can be set to any desired level. The setting of C_a will determine the relative proportion of the various outcome categories (e.g, the number of false alarms relative to misses) and the overall alarm rate of the automated monitor. Since the C_a parameter is not limited by hardware and software factors, we confine our analysis of system interactions to those between C_a and the human subsystem parameters, C_h and d'_h.

Dependence of C_h on C_a: Dependent Criterion Strategy

One interesting form of interaction between the human and automated monitor subsystems would result if C_h were some function of C_a. This could occur if, as the number of false alarms from the automated monitor increases, the human monitor required more evidence to confirm the alerted condition (that is, if the human became more conservative as the automated monitor became more liberal). We call this possibility the *dependent criterion strategy*. To illustrate this strategy, we have assumed that the operator criterion increases with the probability of an output from the automated subsystem, $P(\text{alarm})$:

$$C_h = 2\ P(\text{alarm}) \tag{3}$$

where

$$P(\text{alarm}) = P_a(Y/SN)\,P(SN) + P_a(Y/N)\,P(N) \tag{4}$$

where $P_a(\ \ )$ refers to the hit or false alarm rate of the automated subsystem and $P(SN)$ and $P(N)$ are the prior probabilities of signal and noise. In these calculations we have assumed $P(SN) = P(N) = 0.5$.

Figure 7 shows the system operating characteristic that results from these assumptions. The curve has a prominent peak along its lower portion; it increases from (0,0) to a peak hit rate of about 0.74 and then falls rapidly. This

Figure 7. The consequences of subsystem interaction if the human monitor's criterion depends on the output (alarm) rate of the automated monitor. Each subsystem $d' = 2$. Performance is evaluated for all values of C_a, under the assumption that $C_h = 2 P(\text{alarm})$.
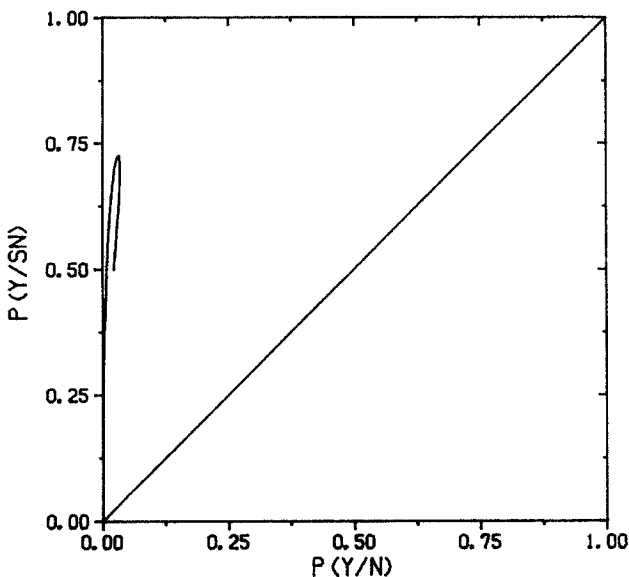
means that effective system performance will be possible only over a very narrow range of low output rates from the automated monitor. At intermediate or high rates, system performance quickly drops off to a level determined by the sensitivity and criterion parameters of the automated system alone. High system hit (and false alarm) rates are impossible to achieve with the human in the system.

Dependence of d'_h on C_a: Operator Sampling and d'-Allocation Strategies

In addition to the possibility that the operator criterion may depend on the alerting signal rate, high rates may lead to situations where the operator either ignores some alerting signals or makes observations with a reduced detectability. Suppose that the operator is sufficiently "busy" that there is a limit to the operator's ability to process information in all of the tasks (cf. Allport, 1980; Lane, 1982). That is, there is a set of primary tasks that demands attentional capacity, and only the remaining attentional resources are allocated to the input channel. Given this situation, at least two types of strategies would result in a dependence of operator sensitivity on the automated system's parameters. We call the first possibility the *operator sampling strategy.* Under this strategy, the operator makes observations only on a subset of the alerted

events; the rest are ignored. The second type of strategy is termed the d'-allocation strategy. In this case we assume that the operator will make an observation on the input channel whenever alerted, but will observe with a reduced d'.

Figure 8 shows combined system performance under the operator sampling strategy. Given an alerting signal from the automated subsystem, the operator will observe the input channel with some probability, P(observe). To show the general effect of an operator sampling strategy, we have assumed that P(observe) decreases linearly from a value of 1.0 when the alerting rate, P(alarm), is zero, to a value of 0.5 when P(alarm) is 1.0. That is, when alarms occur infrequently, the operator will observe the input channel on every occurrence of the alarm. When alarms are very frequent, the operator will observe the input channel only half of those times. This is not an extreme assumption, since the human monitor observes the input channel on at least half of the alerted events. On those occasions when the operator does not make an observation, the alarm is ignored and the system makes a No response.

Each of the curves shown in Figure 8 correspond to a fixed value of C_a (shown) and a corresponding value of P(observe), evaluated over all values of C_h. High values of P(alarm), e.g., negative values of C_a, result in operating characteristics close to the chance line (e.g., the $C_a = -0.5$ curve). In general, under operator sampling strategies the system operating characteristics trace curves in the lower-left quadrant of the ROC space. If the sampling probability decreases to zero, system performance will decrease all the way to chance. One example of the sampling strategy may be seen in the reaction of busy air crew personnel to flightdeck warning signals. The flight crew may be set to automatically cancel a warning signal as soon as it occurs, rather than investigate the potential problem signaled by the warning.[6]

The second type of dependence between operator sensitivity and alerting subsystem parameters follows from the assumption that there are limits to the amount of detection capacity that can be assigned to the alarm channel. Operator observations are made on all alerted events, but high alerting rates lead to a decrease in the d' that the operator may use. To examine this idea, it is necessary to postulate how detection capacity might be allocated across different numbers of observations. We have postulated two different functions relating d'_h to the alerting rate, P(alarm), as illustrated in Figure 9. The solid curve of the figure results from the assumption that $(d'_h)^2$ decreases linearly with the number of observed events: That is,

$$d'_h = d'_a \sqrt{1 - P(\text{alarm})} \tag{5}$$

where $d'_a = 2.0$.

---

[6] The motivation for this behavior may not be only because a false alarm is expected; the air crew may quickly cancel warning signals because of the signal's aversive and distracting nature and their potential to disrupt vital communications (Patterson, 1982).

Figure 8.   The consequences of subsystem interaction under the sampling assumption. Each curve is the system operating characteristic for a fixed value of C_a (shown in d' or standard deviate units) and all possible values of C_h; $d'_a = d'_h = 2$. The Human Monitor is assumed to make observations on the input channel with a probability that decreases linearly with the alarm rate of the automated monitor: $P(\text{observe}) = 1 - [P(\text{alarm})/2]$.
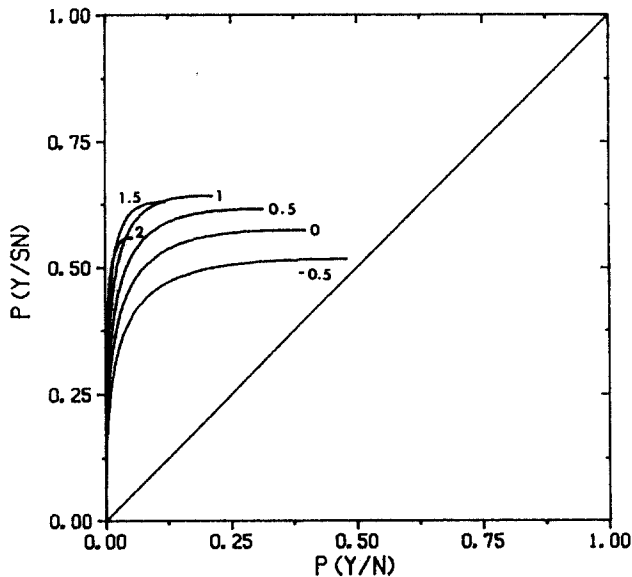

This function was chosen to provide detectabilities between 1.0 and 2.0 over most of the range of $P(\text{alarm})$. These values generally constrain d'_h between modest above-chance performance and the moderately good performance of the automated stage ($d' = 2.0$). The effect of this strategy on system performance is shown in Figure 10. Each operating characteristic corresponds to a fixed value of C_a (shown) and a corresponding value of d'_h, evaluated over all levels of C_h. The system operating characteristics approach chance performance as C_a becomes increasingly negative and $P(\text{alarm})$ goes to 1.0.

An alternative relationship between d'_h and $P(\text{alarm})$ is shown as the dotted curve of Figure 9. In this case, d'_h is equal to 4.0 when very few alerting events occur. When $P(\text{alarm})$ is high, however, d'_h decreases to a limiting value of 0.5. These values were chosen to illustrate the case when increases in low alarm rates can produce significant effects on operator detectability. The equation describing this relationship is:

$$d'_h = 2d'_a[7P(\text{alarm}) + 1]^{-1} \tag{6}$$

Figure 9.   Two forms of the assumed dependence of operator detectability on the alarm rate of the automated monitor (see text).
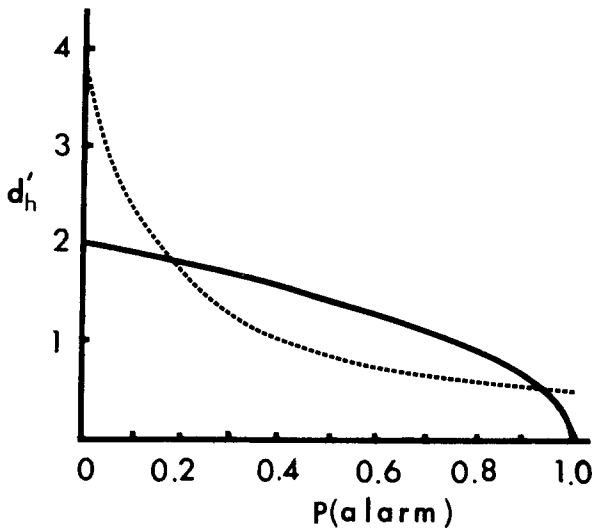

where $d'_a = 2.0$. Figure 11 illustrates the effect of this strategy on the system operating characteristics. The curves are similar to those of Figure 10 but reach a minimum level of system performance equal to $d' = 1.0$ for increasingly negative values of C_a.

This analysis of automated-human subsystem interaction indicates that the dependent criterion, operator sampling, and d'-allocation strategies each result in characteristically different effects on system performance. Figures 8, 10 and 11 also illustrate that the particular interaction function assumed will determine (1) how rapidly the family of operating curves will shift toward the chance line, and (2) whether they will reach a limiting value of d' that is greater than zero.

The particular d'-allocation strategy specified by equation (6) (the dotted curve of Figure 9) has the interesting property that high values of $P(\text{alarm})$ do not produce additional decreases in d'_h. That is, as the number of alerting events rises and the operator's workload increases, performance on the alerted-monitor task suffers smaller and smaller decrements. This implies that the operator's performance is not limited by a fixed amount of processing capacity that must be shared among the alerted-monitor and the other operator tasks. Rather, the operator's information-processing capacity is "elastic" and may increase somewhat as the total task demands increase. This idea is consistent with the reports of some investigators (cf. Allport, 1980; Kahnemann, 1973; Wickens, 1979). Whether operator performance exhibits fixed or elastic

Figure 10. The consequences of system interaction under a d' allocation strategy (the solid curve of Figure 9). Each curve is the system operating curve for a fixed value of C_a (shown) and all possible values of C_h; $d'_a = 2$.
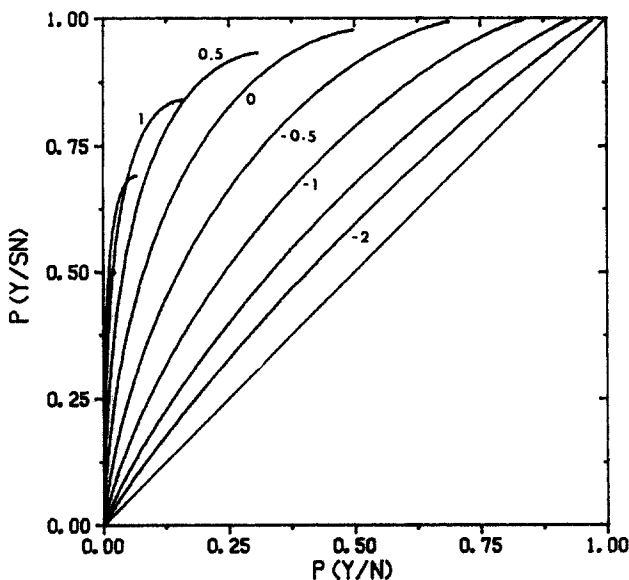

capacity properties depends on the specific characteristics of the operator tasks and the operator training. Lane (1982) and Wickens, Sandry, and Vidulich (1983) have reviewed many of the issues involved in this question. In general, we would expect that operator strategy in an alerted-monitor task would depend on the precise nature of such task and training factors. If the alerted-monitor task were a salient absorber of operator processing resources compared to other tasks, then small changes in the automated subsystem's criterion might result in large changes in operator resource allocation.

One interesting consequence of task interaction may be a combined sampling and d'-allocation strategy. Figure 12 illustrates one version of such a strategy. The operator sampling probability (dotted line) is fixed at 1.0 for values of P(alarm) less than 0.5, then decreases linearly to a minimum value of 0.2; that is, the operator observes the input channel at least 20% of the time he or she is alerted by the alarm subsystem. The relationship between d'_h and P(alarm) is assumed to follow a function similar to equation (6), but decreases to a minimum d' of 1.0 rather than 0.5. These particular values were chosen to illustrate the general effects of combining the two operator strategies. These assumptions are given by the following equations:

for P(alarm) < 0.5,

Figure 11. The consequences of system interaction under a d' allocation strategy (the dotted curve of Figure 9). Each curve is the system operating curve for a fixed value of C_a (shown) and all possible values of C_h; $d'_a = 2$. The curves converge on the ROC for a single-stage system having a $d' = 1.0$. The last three curves plotted are for C_a values of -1.5, -2.0, and -2.5.
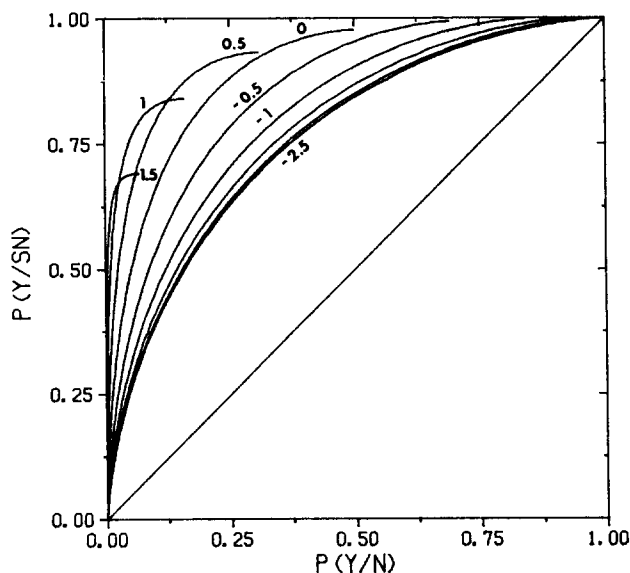
$$P(\text{observe}) = 1.0 \text{ and } d'_h = 1.33 \, [0.33 + 1.67 \, P(\text{alarm})]^{-1} \qquad (7)$$

for $P(\text{alarm}) \geq 0.5$,

$$P(\text{observe}) = 1.8 - 1.6 \, P(\text{alarm}) \text{ and } d'_h = 1.0 \qquad (8)$$

Figure 13 illustrates the effects of these assumptions on the system operating characteristics. The parameter of the curves is the value of the alerting subsystem criterion. This figure has both the sampling and d' allocation aspects shown in the previous figures. As $P(\text{alarm})$ increases, the system performance curves drop toward the chance line and toward the origin.

We can make the additional assumption that the human operator will attempt to match his or her criterion to that of the automated subsystem, for example, $C_h = C_a$. This strategy maximizes the system d' when the automated and operator subsystems are independent and have equal d's. The result is the operating characteristic shown in Figure 14. This curve has one prominent peak at a low criterion level and a second apparent peak at an intermediate criterion value. Thus, for this system, efficient performance is possible only for certain values of C_h and certain system hit and false alarm rates.

Figure 12.   An example of a combined sampling and d' allocation strategy. Neither the operator's d' nor the operator's probability of observing decreases to zero at high alarm rates.
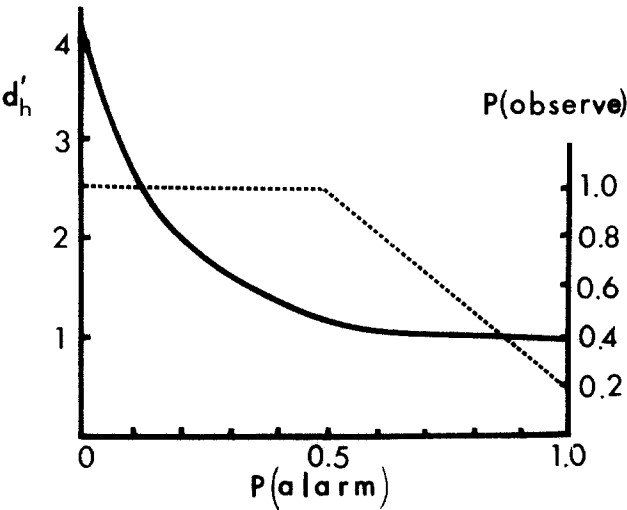

Figure 13.   The consequences on system performance of the combined strategy shown in Figure 12. Each curve is the system operating curve for a fixed value of C_a (shown) and all possible values of C_h; $d'_a = 2$.
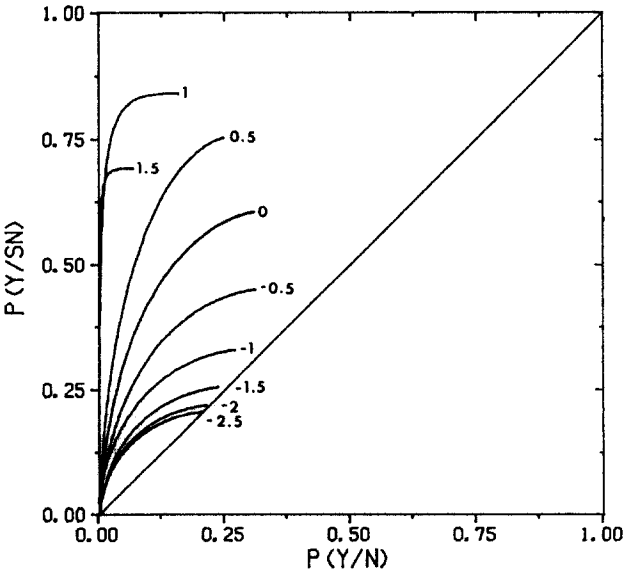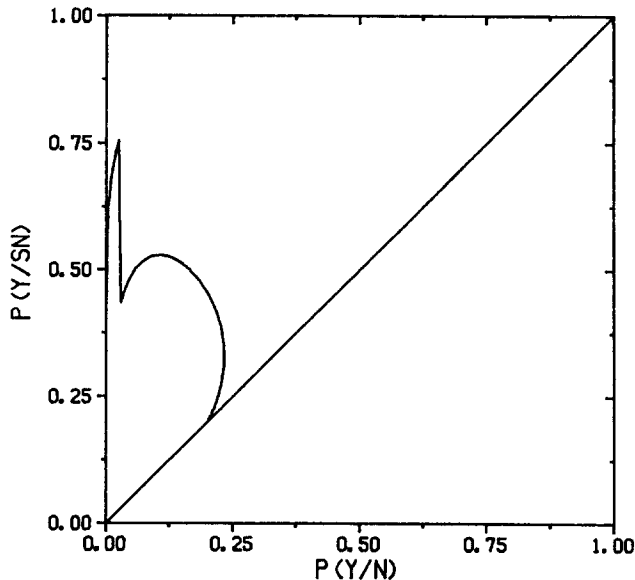
Figure 14. The system operating characteristic for the system shown in Figures 12 and 13 under the additional constraint that $C_h = C_a$. Performance is evaluated over all possible values of C_a.

These values may not correspond to those required for a given application and probability of signal. Furthermore, the complexity of the operating characteristic may be unknown to the operator and the system designer; system performance at desired hit and false alarm rates may never be achieved because of improper specification of C_a. Although the distinctive operating characteristic of Figure 14 results from arbitrary assumptions about operator monitoring strategy, we have included it to illustrate the potential impact on system performance that may result from interactions between the human and automated subsystems.

Summary of the Interaction Analysis Results

This analysis of representative types of interaction between the human and automated subsystems has pointed up the following relationships between the subsystem parameters and their effect on overall system performance:

1. The operator's criterion may depend on the parameters of the automated subsystem, especially C_a. This may be a rational operator strategy in high information-processing load circumstances, although overall alerted-monitor system performance can be impaired. Effective system performance may be limited to very narrow ranges of (low) system hit and false alarm rates (Figure 7).

2. Operator sampling or sensitivity (d') allocation strategies may result if the total task loading (alerted-monitor task plus other tasks) is sufficiently great. As a consequence, good system performance may be limited to certain hit and false alarm levels. For example, under the sampling strategy, performance may be restricted to the lower-left quadrant of the ROC space. A sampling or d'-allocation strategy will cause system performance to drop rapidly either to an asymptotic level or to a chance level, depending on the parameters of the interaction function (Figures 8, 10, and 11).

3. The types of interaction between the operator and automated subsystems will depend on the characteristics of the particular alerted-monitor task situation, including the operator's taskload, the expected probability of signals and non-signals, and the desired system hit and false alarm rate. Appropriate system design and operator training could (a) reduce the task loadings and interactions and increase the operator's effective processing capacity; (b) provide accurate specification to the operator of the desired system operating point and of the automated system's criterion, detectability, and relative importance; and (c) improve the accuracy of operator monitoring performance (and possibly reduce the correlation between the operator and automatic subsystems).

3.   DISCUSSION AND IMPLICATIONS FOR SYSTEM DESIGN

3.1. C_a in the Design of Alerted-Monitor Systems

The preceding analysis demonstrates the importance of designing automated monitoring systems based on the performance of the total human-machine alerted-monitor system, rather than on the performance of the automated subsystem alone. A critical parameter of automated monitors with respect to overall system performance is the criterion C_a. This parameter can be set to any desired value because, unlike d'_a, the criterion is usually not constrained by considerations of cost or computational capacity. The designer of an automated subsystem is likely to set C so as to minimize the number of signals missed by the subsystem. This is because, from the point of view of subsystem performance alone, a greater cost is usually associated with a miss than with a false alarm. Nonetheless, our analysis of the combined human-machine monitoring system, shows that this approach can lead to the impaired performance of the total system. In other words, optimal criterion placement for one stage of a two-stage detection system may not be identical to optimal placement in the single-stage case.

Analysis of the two-stage signal detection model also demonstrates that quantitative predictions can be made about alerted-monitor system performance via system operating characteristics, if some data are available about how operator detectability and criterion depend on the characteristics of the auto-

mated monitor and on other task conditions. Our analysis provides a conceptual framework to guide empirical studies of automated-monitor–human monitor performance. The important independent variables in such studies are the processing load imposed by non-monitor tasks and the response criterion of the alerting subsystem C_a. Relevant dependent variables include operator criterion, C_h, operator observing probability, P(observe), and operator sensitivity, d'_h. These variables should be assessed in conditions that include a range of operator and automated subsystem criteria. We are presently working toward obtaining this empirical specification of alerted-monitor system performance.

3.2. Complex Alerted-Monitor Systems

The signal detection analysis also suggests possible design features for advanced alerted-monitor systems. One possibility involves the use of multiple output criteria in the automated subsystem. A simple example is an alerting signal explicitly coded to indicate the conservatism of the criterion (C_a) that has been exceeded. The operator could then shift resource allocation strategies based on the level of alarm. Designers of alerted-monitor systems might also use multiple subsystem criteria in more complex ways, such as having different modes of human-machine interaction triggered by different criterion values (cf. Sheridan, 1982, for one categorization of levels of automated-decision system–human decision-maker interaction). For example, a monitoring system might combine a relatively liberal subsystem criterion that signals an advisory or decision support mode of human-machine interaction. A more conservative criterion could trigger a second mode of interaction, where the automated subsystem directly activates control systems subject to subsequent operator review and intervention.

A second possible design feature is to present directly to the human monitor the likelihood estimate, for example, the value of the decision statistic, Z, computed by the automated subsystem. From a display point of view, this suggestion could imply an analog form of data presentation (i.e., presentation of automated subsystem output as a continuous variable, "degree of alarm," with response criteria highlighted, rather than only as alarm/no alarm categories). A report by Woods, Wise, and Hanes (1982) contains an example of such a display. Likelihood presentation is one mechanism that could allow the human monitor to inspect (although not necessarily duplicate) the process by which the automated subsystem reaches its conclusion in cases of uncertainty or conflict. Finally, interface features that reduce the correlation between the automated and human monitor's input channels can increase the performance capability of the combined system. This is a subtle aspect of system design, because in many situations there may be an advantage in redundancy between these channels.

In some task environments there is the possibility of considerable interaction among the alerting signals themselves, and alerting signals may be missed or confused (Cooper, 1977; Patterson & Milroy, 1980). Patterson and Milroy discussed the confusability of auditory alarm signals on aircraft flightdecks. The present analysis does not address this problem, since we have assumed that the detectability and discriminability of the actual alerting signals are infinite. Extending the signal detection analysis to include these factors introduces two kinds of complications. First, the presence of noise in the task environment can impair the detectability of the alarm signal itself. The general problem of an alerting subsystem whose output signal is communicated over a noisy channel is part of a problem described by Dr. James P. Egan as the "telegraph relay" problem. This assumption increases the complexity of the alerted-monitor model; the system becomes a three-stage detection system.

The second complication results from the existence of multiple alarm channels. How would an operator establish response priorities to a set of four automated monitors, each of which has unique pay-off/penalty contingencies, prior probabilities of signal, and sensitivity and criterion parameters? It is clear that the allocation of processing resources would be a significant aspect of such a task.

3.3. Conclusion

As technological changes shift the human role to primarily one of supervisory control (National Research Council, in press; Woods, 1982), monitoring an array of multiple automated monitors could well be the major function of the operator. Describing and predicting the performance of such a system would be critically dependent on describing the system behavior on each channel (and the channel interactions), in terms of the models proposed in our analysis. Indeed, the alerted-monitor system could provide a general paradigm for describing the behavior of operators in a variety of multiple channel and time-shared tasks.

REFERENCES

Allport, D. A. (1980). Attention and performance. In G. Clayton (Ed.), *Cognitive psychology*. London: Routledge.

Banks, W. W., & Boone, M. P. (1981). *Nuclear control room annunciators: Problems and rec-*

ommendations (NUREG/CR-2147). Springfield, VA: National Technical Information Service.

Bastl, W., & Felkel, L. (1981). Disturbance analysis systems. In J. Rasmussen & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures.* New York: Plenum.

Bisseret, A. (1981). Application of signal detection theory to decision making in supervisory control. *Ergonomics, 24,* 81–94.

Berson, B. L., Po-Chedley, D. A., Boucek, G. P., Hanson, D. C., Leffler, M. F., & Wasson, R. L. (1981, January). *Aircraft alerting systems standardization study, II. Aircraft alerting system design guidelines* (DOT/FAA/RD-81-38-II). Springfield, VA: National Technical Information Service.

Bray, M. A., Nelson, W. R., Blackman, H. S., & Fowler, R. D. (1984). *Response tree evaluation: Implications for the use of artificial intelligence in process control rooms* (NUREG/CR-3461). Springfield, VA: National Technical Information Service.

Cooper, G. E. (1977, June). *A survey of the status and philosophies relating to cockpit warning systems* (NASA-CR-152071). NASA Ames Research Center, Moffett Field, California.

Davies, D. R., & Parasuraman, R. (1981). *The psychology of vigilance.* New York: Academic.

Egan, J. P. (1975). *Signal detection theory and ROC analysis.* New York: Academic.

Enstrom, K. D., & Rouse, W. B. (1977). Real-time determination of how a human has allocated his attention between control and monitoring tasks. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-7,* 153–161.

Gimmy, K. L., & Nomm, E. (1982). Automatic diagnosis of multiple alarms for reactor control rooms. *Transactions of the American Nuclear Society, 41,* 520.

Goldmutz, L., Kefaliotis, J. T., Kleiman, L. A., Rucker, R. A., Schuchman, L., & Weathers, D. (1981, March). *Automated en-route air traffic control (AERA) concept* (Report No. FAA-EM-8-3). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration, Office of Systems Engineering Management.

Green, D. E., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Jeffress, L. A., & Robinson, D. E. (1962). Formulas for the coefficient of interaural correlation for noise. *Journal of the Acoustical Society of America, 34,* 1658–1659.

Kahnemann, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kantowitz, B. K. (1982). Interfacing human information processing and engineering psychology. In W. C. Howell & E. A. Fleishman (Eds.), *Human performance and productivity: Information processing and decision making.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kantowitz, B. K., & Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships.* New York: Wiley.

Kragt, H., & Bonten, J. (1983). Evaluation of a conventional process-alarm system in a fertilizer plant. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13,* 586–600.

Lane, D. M. (1982). Limited capacity, attention allocation, and productivity. In W. C. Howell & E. A. Fleishman (Eds.), *Human performance and productivity: Information processing and decision making.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Lind, M. (1981). The use of flow models for automated plant diagnosis. In J. Rasmussen & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures.* New York: Plenum.

National Research Council. (in press). *Research and modeling of supervisory control behavior.* Washington, DC: National Academy Press.

Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft.* CAA Paper 82017. Civil Aviation Authority, London, November.

Patterson, R. D., & Milroy, R. (1980, February). *Auditory warnings on civil aircraft: The learning and retention of warnings* (CAA-7D-5-0141). MRC Applied Psychology Unit, Cambridge, England.

Pollack, I., & Madans, A. B. (1964). On the performance of a combination of detectors. *Human Factors, 6,* 523–531.

Reason, J., & Mycielska, K. (1982). *Absent-minded? The psychology of mental lapses and everyday errors.* Englewood Cliffs, NJ: Prentice-Hall.

Seminara, J. L., Gonzalez, W. R., & Parsons, S. O. (1977). *Human factors review of nuclear power plant control room design* (NP-309). Palo Alto, CA: Electric Power Research Institute.

Sheridan, T. B. (1982). *Theory of man-machine interaction as related to automation.* (MIT Industrial Liaison Report, 12-72-82). Cambridge, MA: MIT Press.

Sheridan, T. B., & Ferrell, W. R. (1974). *Man-machine systems: Information, control, and decision models of human performance.* Cambridge, MA: MIT Press.

Swensson, R. G. (1980). A two-stage detection model applied to skilled visual search by radiologists. *Perception & Psychophysics, 27,* 11–16.

Swets, J. A. (Ed.) (1964). *Signal detection and recognition by human observers.* New York: Wiley.

Swets, J. A., & Pickett, R. M. (1982). *Evaluation of diagnostic systems: Methods from signal detection theory.* New York: Academic.

Tanner, W. P., Jr., & Sorkin, R. D. (1972). The theory of signal detectability. In J. B. Tobias (Ed.), *Foundations of modern auditory theory,* New York: Academic.

Thompson, D. A. (1981). Commercial aircrew detection of system failures: State of the art and future trends. In J. Rasmussen & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures.* New York: Plenum.

U.S. Nuclear Regulatory Commission. (1983, March). *NRC fact-finding task force report on the ATWS event at Salem nuclear generating station, Unit 1, on February 25, 1983.* (NUREG-0977). Springfield, VA: National Technical Information Service.

Visuri, P. J. Multi-variate alarm handling and display. (1983). *Proceedings of the International Meeting on Thermal Nuclear Reactor Safety* (NUREG/CP-0027). Springfield, VA: National Technical Information Service.

Whitaker, L. A. (1979, April). *Dual-task Interference as a Function of Cognitive Processing Load* (Report Number 840) Naval Submarine Medical Research Laboratory, Submarine Base, Groton, CT. Naval Medical Research and Development Command, Research Work Unit MR000.01.01-5.

Wickens, C. D. (1979). Position Paper — NATO Conference on Workload. In N. Moray, (Ed.), *Mental workload: Theory and measurement.* New York: Plenum.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors, 25,* 227–248.

Williams, D. H., & Simpson, C. A. (1976, October). *A systematic approach to advanced cockpit warning systems for air transport operations: Line pilot preferences* (NASA SP416). NASA Aircraft Safety and Operating Problems Conference, NASA Langley Research Center, Hampton, VA, pp. 617–644.

Woods, D. D. The cognitive coupling of man and machine. (1982, October). *Second International Workshop on Safety and Reliability of Industrial Computer Systems,* West Lafayette, IN.

Woods, D. D., Wise, J. A., & Hanes, L. F. (1982). *Evaluation of safety parameter display concepts* (NP-2239). Palo Alto, CA: Electric Power Research Institute.