D12

N79-17487

THE EFFECTS OF PARTICIPATORY MODE AND TASK WORKLOAD ON THE DETECTION OF DYNAMIC SYSTEM FAILURES[1]

Christopher D. Wickens
Colin Kessel

Department of Psychology
University of Illinois at Urbana-Champaign

ABSTRACT

The ability of operators to detect step changes in the dynamics of control systems is investigated as a joint function of, (a) participatory mode: whether subjects are actively controlling those dynamics or are monitoring an autopilot controlling them, and (b) concurrent task workload. A theoretical analysis of detection in the two modes identifies factors that will favor detection in either mode. Three subjects detected system failures in either an autopilot or manual controlling mode, under single-task conditions and concurrently with a "subcritical" tracking task. Latency and accuracy of detection were assessed and related through a speed-accuracy tradeoff representation. It was concluded that failure detection performance was better during manual control than during autopilot control, and that the extent of this superiority was enhanced as dual-task load increased. Ensemble averaging and multiple regression techniques were then employed to investigate the cues utilized by the subjects in making their detection decisions.

INTRODUCTION

Over the past decade, the aviation industry has witnessed a gradual change in the role of the pilot in the cockpit. Many traditional pilot functions have been replaced by on-board computers, and in some instances the pilot is no more than a supervisor (Sheridan, 1976) or monitor of automatically controlled functions. One task, however, that remains of critical importance to the operator of any aviation system, whether he is removed from the control loop or not, is that of monitoring all facets of aircraft performance for the occurrence of failures or malfunctions. The relatively low frequency of occurrence of such events does not diminish the importance of failure monitoring and detection, because the consequences of an undetected malfunction, or one that is detected after an unnecessary delay, can be disastrous, potentially resulting in the loss of the aircraft or of human life.

It can be argued in fact that one criterion that should be used in considering whether a pilot should remain in the control loop under particular conditions is his relative sensitivity to system malfunctions in the two modes of participation.

Young (1969) has argued strongly on the basis of his findings that the operator is more sensitive to system malfunctions as an active participant in the control loop, than as a passive monitor. In his experiment, subjects were required to detect various step changes in system order and gain. Conditions were compared in which the subject was an active controller and a *passive monitor* (who was observing the compensatory display produced by another active controller). Under these circumstances detection latencies varied from twice to five times greater for the monitor than the controller. A second study which also compared detection ability in the two modes, however, resulted in contradictory findings. Ephrath (1975) investigated failure detection performance in a two-dimensional simulated landing task as a joint function of participatory mode and workload. The "failures," which in this case were deviations introduced into the flight path rather than changes in system dynamics, could occur in either the pitch or yaw channel. Under different conditions subjects were either in control when a failure occurred or were monitoring a nonadaptive autopilot in control of that channel. The non-failed channel could also be either controlled or monitored. Ephrath's results indicated a clear superiority for detection on the monitored as opposed to the controlled dimension, both in terms of the smaller number of missed failures and of the shorter detection latency. This difference Ephrath attributed in large part to the increased level of workload involved in the controlled task.

Obviously, in many respects the studies of Young and Ephrath are not comparable. Young employed single-axis tracking with changes in system dynamics, while Ephrath employed dual-axis simulator control with "deviation" *failures*. *In addition, the monitoring conditions were different* in the two experiments, being influenced by adaptation in Young's study and not in Ephrath's. In this light, it is not surprising that the conclusions differed dramatically. Certainly, one of the most salient differences between these studies lies in the contrast between single- and dual-axis tracking and is inherent in the greater workload imposed in the latter condition.

While numerous other investigations of failure detection performance are present in the literature (Sheridan and Johansen, 1976; and Young, 1969), the studies of Young and Ephrath are the only two that have explicitly contrasted detection between the two modes, so that a direct comparison is possible. The present study was conducted with the intent of clarifying the nature of the superiority relation between the two modes. A question of specific interest was whether the difference in results between the results of Ephrath's and Young's study could be attributable to differences in concurrent task workload between the paradigms, and for this reason secondary task workload was manipulated orthogonally to participatory mode.

---

[1]Contractual support for this research was provided by the Life Sciences Program, Air Force Office of Scientific Research, Contract Number F44620-76-C-0009. Dr. Alfred Fregly was the Scientific Monitor of the contract. The authors acknowledge the invaluable programming assistance of Mr. Roger Hunt in the conduct of this research.

Theoretical Analysis of Failure Detection

The detection of a failure or change in the characteristics of a dynamic system requires that the detector have available two basic elements: (1) an internal representation of the state of the normally operating system - the expected value of state variables and their expected variability, (Veldhuyzen and Stassen, 1976; Pew, 1974) and; (2) a channel, or set of channels, of information concerning the current state of the system. Failures are detected when the information concerning the current system state is assessed to be sufficiently deviant from the representation of normal operation to warrant a decision. The decision process involved may be assumed to involve the application of some statistical decision rule (Curry and Gai, 1976). The following theoretical analysis, employing the conceptual framework described above and represented symbolically in Figure 1, will attempt to define the characteristics or attributes of each participatory mode that might be expected to enhance the sensitivity of failure detection in that mode.

Stability of internal model of dynamics. An evolving conception in control theory is that the operator maintains an internal representation or "model in the head" of the dynamic system that is being controlled. It is assumed here that this model provides the basis for predicting expected system outputs in response to known inputs - an internalized estimate of the transfer function of the system being controlled. This conception is consistent with that employed by Curry and Gai (1976), Miller and Elkind (1967), and others. With respect to failure detection, a critical characteristic of an internal model relates to its internal consistency or expected variability. For any given input to the system, the range or variability of possible outputs is a measure of this consistency.

It is proposed that when the operator is actively controlling, the stability of this internal model is considerably greater than when he is monitoring. This difference reflects the fact that, when controlling, the operator has a greater involvement with the system, and a direct knowledge of its input-output characteristics available by comparing his control inputs with the system response (Gould and Fu, 1966). This information is only available when monitoring if the monitored display is pursuit, and even then, knowledge of control inputs is not as precise, since a system response due to regulatory error correction, cannot easily be discriminated from one resulting from external disturbances. Thus provided with a smaller variance estimate of the normal state, detection in the control mode should show greater sensitivity to departures from this state induced by changing dynamics, than should detection in the monitoring mode.

Information channels. A second attribute of the control mode that predicts superiority of failure detection is the greater number of channels of information concerning the current state. When monitoring, information is provided to the operator exclusively via the visual channel (system error and its derivatives in the compensatory display or input and outputs

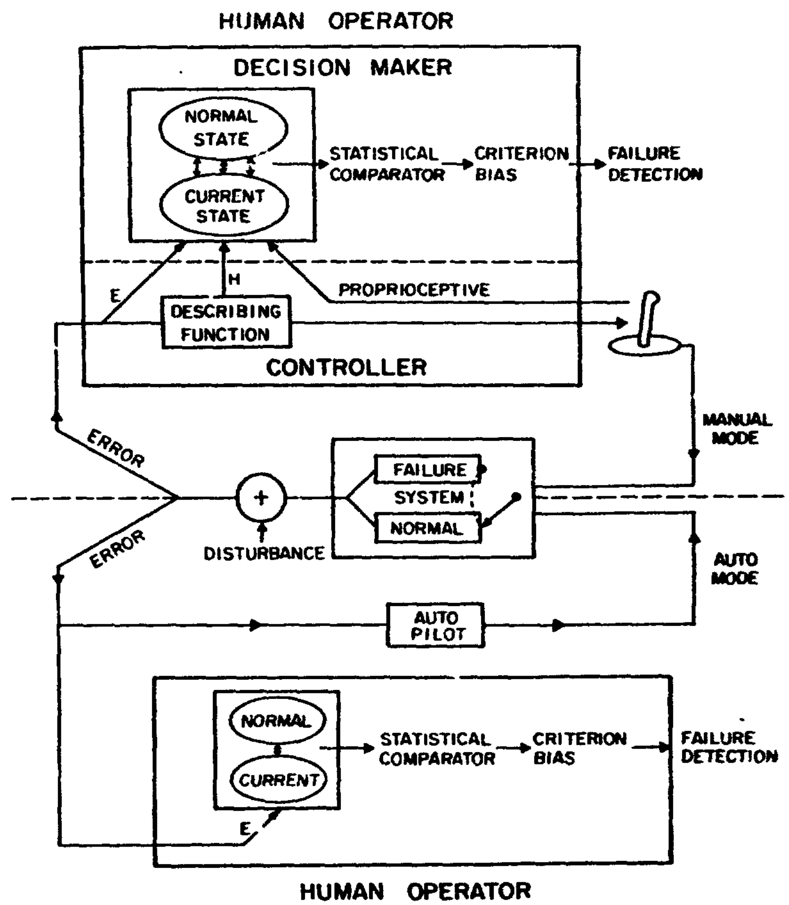

Figure 1. Schematic representation of human failure detection in manual mode (top) and autopilot mode (bottom).

plus derivatives in the pursuit display). On the other hand, in the control situation the operator also has available a proprioceptive channel of information concerning his own input to the control stick, independent of disturbances acting upon the system.

Although control input cannot directly reflect the occurrence of failures (except as failures initiate mechanical feedback from the control itself), it will do so indirectly to the extent that any compensatory adaptation that the operator initiates to a system change will be reflected in a change in his response characteristics (mean control position, velocity or acceleration) and/or the characteristics of the operator's open-loop transfer function. When controlling then, these proprioceptive channels will be available to the detection system to supplement the visual channels that are available in both monitoring and controlling (Figure 1).

While the controlling mode thus seemingly provides a distinct advantage over the monitoring mode by virtue of its added proprioceptive channel, it should be noted that this advantage is not invariably present for reasons related to the non-independence of control input and error. More specifically, if adaptation to the failure is rapid and complete, as may occur for example in response to shifts in system gain (Young, 1969), the obtained distribution of error following the change would show little or no alteration from that characterizing the normal operating state, while a change would be manifest in the characteristics of the control response and, therefore, the transfer function.

Failure to initiate any adaptive control, on the other hand, would leave unchanged the proprioceptive input, while changing both the nature of the error distribution and again, the resulting transfer function. In short, whether or not an adaptive response is implemented, the transfer function will change. If adaptation occurs, the response will change as well. If it does not, then the error distribution will be altered.

However, even provided with only two sources of information (transfer function plus error or control response) rather than three, a comparison of number of channels still favors the control mode over monitoring. Assuming that there is some degree of independence of information processing along the channels, the probability and/or speed of detecting change information along any one of two channels characterizing the control mode, should be greater than that of detecting change along the single visual channel available in the monitoring mode.

Differential sensitivity to visual vs proprioceptive information. Although a strict comparison of the number of channels of information available to a decision mechanism favors control over mointoring, an important caution should be noted. As described above, the operator is able to trade off the strength of the failure occurrence "signal" along the visual vs proprioceptive channel, to the extent that he engages in some degree of compensatory adaptation. As adaptation increases, proprioceptive "signal strength" increases at the expense of visual error "signal strength." Thus the prediction based upon the difference in number of information channels -- that control detection will be superior to monitoring detection -- is predicated upon the assumption that detection of change is equally efficient along all the channels (proprioceptive, transfer function, and visual). In other words this approach assumes that, whichever channels are employed in the control mode, their joint signal will be more easily detected than the single visual signal in the monitoring mode.

Mitigating against this conclusion, however, is a body of literature in psychology suggesting that the sensitivity to proprioceptive information is reduced relative to visual information particularly when the two sources are available at the same time and are conveying conflicting information (e.g. Jordan, 1972; Klein and Posner, 1974). Such a conflict, in fact, describes precisely the stiuation in which an operator has successfully adapted to a change in control dynamics. Under these circumstances, the visual error channel is providing information describing normal operation (since the appropriate gain, or lead-lag adjustment, has presumably been initiated to restore the original open-loop transfer characteristics), while the less sensitive kinesthetic channel conveys the information that a change has in fact been implemented. The predicted consequence of this conflict situation is that the operator will be less likely to detect the change than he would had no adaptation been achieved, the latter condition of course producing a visual signal equivalent to the monitoring mode. McDonnell (1966), in fact, has noted anecdotally such instances in which successful adaptation has been coupled with the failure to detect dynamic system changes.

Workload differences. A second characteristic of the manual control mode that predicts a reduced sensitivity to the occurrence of failures relates to the greater workload imposed by tracking than by monitoring. Numerous examples may be cited from behavioral literature that demonstrate the attention demands of purely perceptual tasks such as monitoring to be less than those of tasks such as tracking in which a requirement for the selection and execution of responses is also imposed (e.g. Keele, 1973; Kerr, 1973). This finding is verified as well in a direct comparison of controlling vs autopilot monitoring in the simulator (Johannsen, Pfender, and Stein, 1976). In the framework of the present analysis, if monitoring for and responding to failures is regarded as a "task" separate from tracking, then since the operator's attentional resources are limited, the greater workload demands imposed in the control mode than in the monitoring mode would predict poorer performance on the added "task" of failure detection in the former condition.

Whereas, workload differences make a clear prediction of detection differences in the single-task environment, this prediction is not as apparent when the performance of additional tasks is required. A common result emanating from much dual-task research is that tasks that are in themselves simpler or less loading are at the same time more vulnerable to performance decrements in a dual-task environment, as more demanding paired tasks "capture" a greater proportion of available attentional resources (Welford, 1968). In the current context, monitoring - the simpler task - should be more vulnerable to additional dual-task requirements than controlling.

Furthermore, to the extent that failure detection while tracking is dependent upon the processing of information integral to the tracking task, then the quality of this information -- and, therefore, the quality of detection itself -- will be preserved as tracking performance is guarded in the face of competing secondary task demands. In contrast, the quality of visual information available in monitoring will be predicted by this view to deteriorate; rendering detection while monitoring more vulnerable to loading than while controlling.

Summary

The implications of the preceding theoretical analysis are complex. In summarizing, two attributes of the controlling mode may be identified that would seemingly facilitate failure detection. A greater stability of the internal model of the system, and a greater number of channels available upon which to base failure detection decisions. At the same time, the latter advantage may be mitigated to the extent that: (a) adaptation takes place reducing the strength of a visual error signal and, (b) proprioceptive sensitivity is less than visual. In comparison the monitoring mode is also characterized by two attributes that could facilitate detections: a greater "strength" of the visual signal (if adaptation by an autopilot does not take place) and a lower level of workload.

Finally, it is argued that any advantage of monitoring over controlling attributable to workload differences might itself be dissipated as the competition for attentional resources is increased by imposing concurrent tasks. Clearly this interplay of factors is sufficiently complex to prohibit precise predictions concerning the superiority of one mode over the other. It does, however, facilitate a clearer identification of the nature of the failure detection task and allows predictions to be formulated concerning the differential effect of variables such as workload or control adaptation on detection performance.

In the following experiment, independent variables of participatory mode and task workload were manipulated to determine their effect on detection. Analysis techniques were then employed in an effort to identify further the nature of the processes operating in detection performance.

METHOD

Subjects

The subjects were three right-handed male university students enrolled in basic flight training courses at the Institute of Aviation. Subjects were paid at a rate of $2.50 per hour.

Apparatus

The basic experimental equipment included a 3 x 4 inch Hewlett Packard Model 1300 CRT display, a spring-centered, dual-axis tracking hand control (with an index-finger trigger) operated with the other hand, and a Raytheon 704 16-bit digital computer with 24k memory and A/D, D/A interfacing that was used both to generate inputs to the tracking display and to process responses of the subjects. The subject was seated on a chair with two arm rests, one for the tracking hand controller and one for the side-task finger controller. The subject's eyes were approximately 50 centimeters from the CRT display.

Tracking tasks. The primary pursuit-tracking task required the subject to match the position of a cursor with that of a target which followed a semi-predictable two-dimensional path across the display. The target's path was determined by the summation of two non-harmonically relate sinusoids along each axis. The frequencies were: X-axis, .08 and .05; Y-axis, .08 and .05. The position of the following cursor was controlled jointly by the subject's control response and by a band-limited forcing function with a cutoff frequency of .32 Hz for both axes. Thus the two inputs to the system were well differentiated in terms of predictability, bandwidth, and locus of effect (target vs cursor). The control dynamics of the tracking task were of the form $Y_c = \frac{1-\alpha}{s} + \frac{\alpha}{s^2}$ for each axis, where α was the variable parameter used to introduce changes in the system dynamics. These changes, or simulated failures, were introduced by step changes in the acceleration constant α from a normal value of .3, a mixed velocity and acceleration system with a high weighting on the velocity component, to α = .9, a system that approximates pure second order dynamics.

As the loading task, the Critical Task (Jex McDonnel and Phatac, 1967), was employed. This was displayed horizontally at the bottom of the screen and required the subject to apply force to the spring-loaded finger control in a left-right direction to keep the unstable error cursor centered on the display. The value of the instability constant λ in the dynamics $Y_c = \frac{k\lambda}{s-\lambda}$ was set at a constant subcritical value. Two values (λ = .05 and λ = 1.0) were employed on different dual task trials.

Experimental Task

Subjects participated in five experimental sessions of which the first two were devoted entirely to practice on the tracking and detection tasks, and the last three used to generate the experimental data. During the first practice day the subject performed only the two-dimensional pursuit tracking task. In the manual (MA) condition the subject performed the tracking manually while in the autopilot (AU) condition, his role in the control loop was replaced by a simulated autopilot control dynamics consisting of a pure gain and effective time-delay. The open loop gain was set at a constant value for all subjects, and the time delay value was adjusted for each subject to obtain an error measure in the AU condition equivalent to the operator's performance in the MA condition. This value of time delay was maintained throughout the rest of the experiment. Each trial, MA or AU, lasted 150 seconds.

To give the subjects some experience with the failed condition (i.e., the higher acceleration in the control dynamics), the subject received two trials (one AU and one MA) in which he tracked (or viewed the autopilot tracking) only the failed dynamics. Two demonstration trials were then presented in which the subject tracked in the regular condition, but the onset of each failure was cued by the presentation of a "F" on the screen. The subject was instructed to press the trigger to return the system to normal only upon the detection of the nature of the change. This training period was then followed by 8 regular detection trials (4 AU, 4 MA in alternating order). Each trial contained either 4 or 6 failures so that a total of 20 failures were presented in each mode.

The presentation of the failure was generated by an algorithm that assured random intervals between presentations and allowed the subject sufficient time to establish baseline tracking performance before the onset of the next change. Task logic also insured that changes would only be introduced when system error was below a criterion value. In the absence of this latter precaution, changes would sometimes introduce obvious "jumps" in cursor position.

During these detection trials, the detection decision was recorded by pressing the trigger on the control stick. This response presented a "T" on the screen and returned the system to normal operating conditions via a four-second ramp to the prefailure dynamics. If the subject failed to detect the change, the system returned to normal after six seconds. This was an interval within which it was assumed, on the basis of pretest data, that responses would correspond to detected failures and not to false alarms. The subjects were told to detect as many changes as possible as quickly as possible.

On the second day (dual-task) the subject performed the primary tracking task together with a side task, the Critical Task. After a refresher trial in the MA mode, the subject received a series of training trials to practice the side task, first in the AU and then in the MA mode. When acceptable criteria were achieved in the Critical Task and MA tracking individually, the subject then carried out these tasks together with the failure demonstrations, as described above.

Eight more experimental trials were then presented in which the subject performed all three tasks (tracking or monitoring, Critical Task, and failure detection). Two trials were presented in each mode at each level of Critical Task difficulty ($\lambda = 0.5, 1.0$). The subject was instructed to "do the side-stick task as efficiently and accurately as possible." The instructions, therefore, clearly defined the side task as the loading task while allowing performance on the tracking and detection tasks to fluctuate in response to covert changes in available attentional resources. In this manner, workload demands were experimentally manipulated, rather than being passively assessed.

During each of the final 3 days, following presentation of 4 warmup trials, the subjects received 2 replications of each of the six experimental conditions: AU and MA under single, and under the two dual task conditions. The order of presentation of the 12 experimental trials was counter-balanced across subjects and across days within a subject. The task logic, instructions, and experimental procedure was otherwise identical to that on days 1 and 2.

ANALYSIS

Assumptions from signal detection theory (Green and Swets, 1966) were employed to account for detection performance in terms not only of the proportion of failures detected (hit rate), but also the number of detection responses made in the absence of failures (false alarms). The signal detection-based sensitivity index reflects changes in both of these values. Some modification of classical signal detection analysis procedures was required because of the undefined nature of the response interval (Watson and Nichols, 1976). According to this procedure it is necessary initially to specify the interval following each failure signal to be designated as a "hit" interval. The data from a number of pretests, indicated that the distribution of subject responses, following signal occurrence, showed a peak at around three seconds and reached a relatively stable baseline by six seconds following a failure. Therefore, six-second intervals were defined as hit intervals, and the measure P(HIT) was simply the number of detection responses falling within the interval divided by the total number of intervals. The remaining duration of the trial (150 seconds - 10 x 4 or 10 x 6 depending on whether it was a 4-failure or 6-failure trial)* was similarly subdivided into six-second false alarm intervals. The measure P(FA) was computed as the number of false alarms divided by the number of false-alarm intervals.

Because of the relatively small number of signals presented, and the questionable applicability of the formal signal detection theory assumptions to the current data, the nonparametric measure of the area under the ROC curve, P(A), was employed as the bias-free measure of sensitivity (Green and Swets, 1966). Values to this measure were computed from the P(HIT) and P(FA) data by reference to tables in McNicol (1972).

This measure produces a score varying from 0 to 1.0 for which 0.5 represents chance performance and 1.0 represents perfect accuracy. Both the P(A) measure and the mean and standard deviation of detection latencies were computed at the end of each trial.

* The extra four seconds relates to the four-second ramp discussed on page

Tracking performance. Tracking measures of vector error, vector stick position, and Critical Task error were sampled every 60 msec and stored on digital tape for later data analysis. In addition, on a fourth channel, the occurrences of failures and responses were recorded. At the end of each trial, the RMS vector error and RMS error on the Critical Task (if performed) were computed.

RESULTS

Figure 2 presents the detection performance (sensitivity and accuracy data) averaged over subjects as a function of conditions, along with the data of individual subjects. The rationale for the joint speed-accuracy representation of Figure 2, is that these two variables represent different manifestations of an underlying performance metric. In any effort to compare "performance" across conditions then, the joint implications of speed and accuracy must be taken into account. For example, a condition that produces a high accuracy of responding might do so at such a prolonged latency that the utility of that decision in a real-world context is less than that of a more rapid decision with slightly lower expected accuracy. Thus the two dimensional representation can be conceptualized as containing a performance axis running from poor (slow and inaccurate) performance in the lower right to good (fast and accurate) performance in the upper left. Individual data points may be projected onto this axis. Furthermore, a speed vs. accuracy bias may be viewed as an axis orthoganal to the performance axis, and points also projected and compared along this axis. The units assigned to the performance index are clearly arbitrary but do require that an assumption be made with regard to the relative weighting of accuracy vs latency in detection. This weighting defines the scaling along the two axes or, equivalently, the slope of the performance axis. In the current data the weighting was made proportional to the intra subject variability of each measure. The elipsoids surrounding each point represent ±1 standard error confidence estimates along the latency and accuracy dimensions.

Figure 2 indicates that detection in the manual mode is apparently superior to autopilot detection. This superiority is manifest at all 3 levels of workload by a large reduction in response latency. This reduction more than compensates on the performance axis for the small loss in accuracy evident in the single task and easy dual task conditions ($\lambda = .5$). For the most difficult dual task condition, ($\lambda = 1.0$) manual detection is favored by both latency and accuracy.

A second aspect of the data of Figure 2 concerns the effect of concurrent task workload. The requirement to perform the critical task leads to a deterioration of detection performance in both modes. Increasing critical task difficulty also causes AU detection performance to decrease further, but counter intuitively appears to improve MA detection. Both trends are consistent across the data of all 3 subjects. The joint effect of dual task load and participatory mode thus takes on a form that sub-

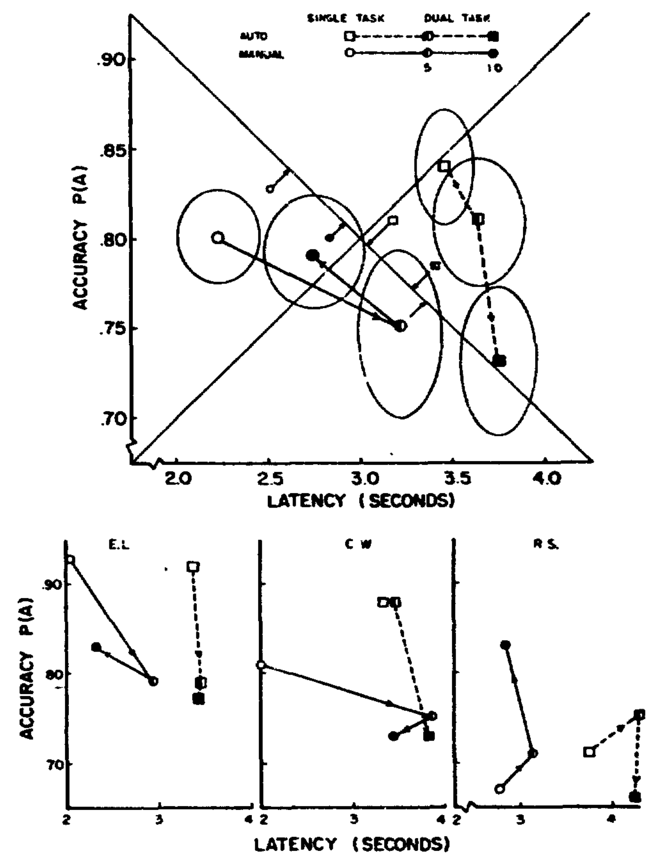

Figure 2. Detection latency and accuracy data of ~~subject means~~ (top) and individual subjects (bottom). One standard error confidence elipsoids surround subject data points.

stantiates data collected on 4 subjects in a shorter pilot experiment: namely that the derogatory effect of increasing concurrent task demands is greater in the AU mode than it is in the MA mode.

The data from the two tracking tasks are shown in Figure 3 presented again with the individual subject data. It is evident that under single task conditions, manual performance was closely matched by autopilot performance, and that an orderly increase in MA primary tracking error was produced by the increasing demand for processing resources imposed by the critical task requirement. This means that the improved detection in the difficult dual task MA condition cannot be attributed to improved tracking performance. Performance on the critical task itself was only slightly affected by its own difficulty, thus serving as a guarantee that this manipulation was an effective way of controlling the resources available for the primary tracking and detection tasks.

Figure 4 presents the ensemble averages of the tracking error and stick velocity following failures. Separate averages were computed for both detected and missed failures. The data are only presented for the single and the difficult dual task conditions. It may be argued that, to the extent that these profiles climb from their initial pre-failure level, there is information contained in the profiled signal that could be extracted by the subject as a cue that a failure had occured. Furthermore, to the extent that the hit and miss profies diverge, evidences is provided that this information was different on the two classes of trials and suggests that the information was used for detection. The converse however, does not necessarily follow.

Following this line of reasoning, the data from the auto condition (4a) indicate that failure information was present in the error signal, and was utilized in detection in the single task condition and to a lesser extent in the dual task condition. The lesser degree of separation here seemingly accounts for the greatly reduced detection accuracy in the dual task condition. An interesting reversal of this trend is evident in the manual condition (4b), where the error profile is greater on missed than on detected trials. This observation suggests that in response to the failures, subjects were making an adaptive adjustment of their control strategy in order to regulate tracking error (addition of lead to compensate for the higher order dynamics). It is possible that this adjustment was initiated prior to overt detection, and the resulting proprioceptive cues were employed as a basis for the detection decision. This hypothesis is neither supported nor refuted by the ensemble averages of response velocity in Figure 4c. It is apparent that some adaptation does take place, as the profiles rise from their pre-failure level. However, the only distinction between the hit and miss profiles suggests that there may be greater, rather than less response velocity associated with miss trials.

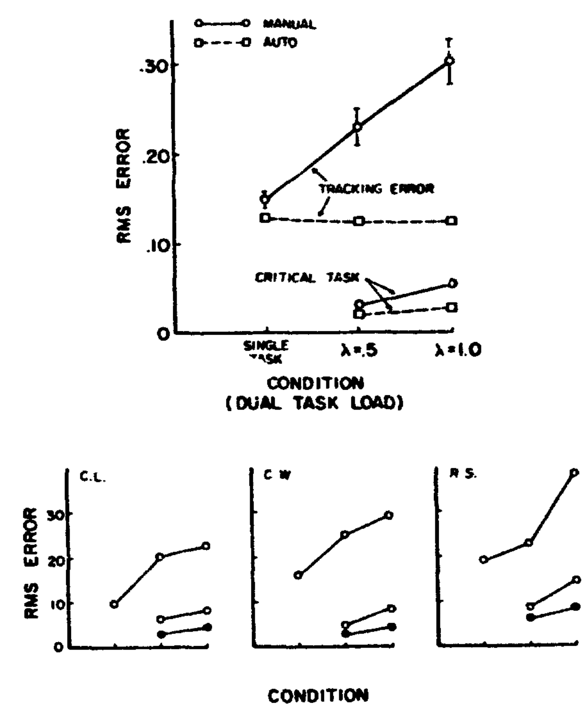

Figure 3. Primary tracking and critical task tracking error. Subject means (top) individual subjects (bottom).

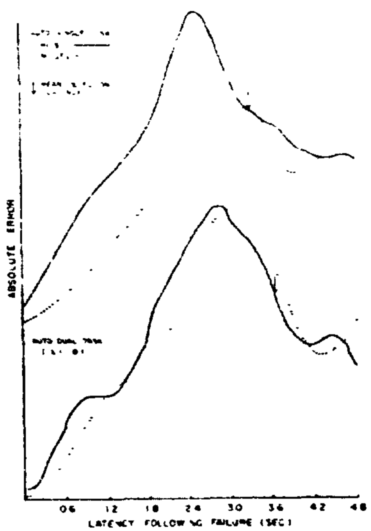

Figure 4a

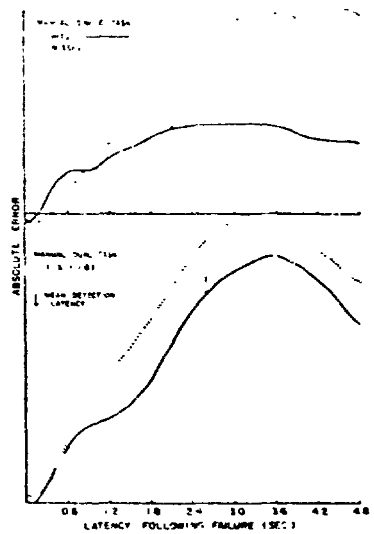

Figure 4b

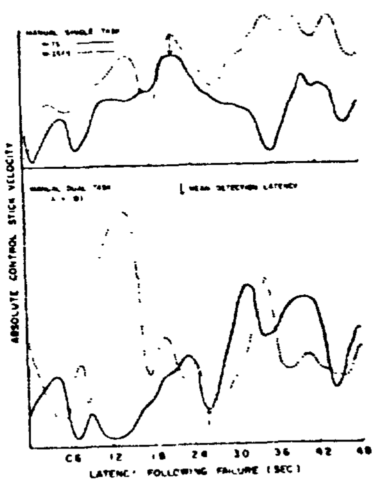

Figure 4c

Figure 4. Ensemble averages of tracking signals following failure occurrence.

In a related effort to discover the cues employed in reaching the detection decision, multiple regression analysis was used to predict detection latency, as a function of characteristics of the tracking error and response signals sampled at various time points following failure occurance. To the extent that a subject is relying upon a particular characteristic of the signal (e.g. error velocity) as an indication of failure, the value of this signal at some latency following the failure should correlate highly with detection latency. The results of this analysis failed to produce any strong trends however. In the AU condition characteristics of the error signal (error & error velocity) were low, but consistent predictors of response latency. The correlations with latency of the two error characteristics, sampled at .6, 1.2, 2.4 seconds following failure ranged from r = -.06 to -.39 (median r = -.12) for error and from r = -.10 to -.40 (median r = -.28) for error velocity. In the MA condition the hypothesis described above, that proprioceptive channels from control adaptation might serve as a detection cue, did not receive support. While this hypothesis predicts a negative correlation of response velocity (as an index of control adaptation) with latency, the obtained correlations between these variables were uniformly low and non-reliable.

DISCUSSION

The present data seemingly substantiate the conclusion drawn by Young (1969), that the sensitivity to failures is greater when the subject is an active participant in the control loop than when he is a passive monitor. While it is acknowledged that this superiority relation is undoubtably sensitive to particular task characteristics, for example the nature of the failure, the dynamics of the autopilot and its degree of adaptability, two characteristics of the current results argue for acceptance of their reliability. (1) The characteristics of the autopilot mode were clearly biased in favor of detection, in that the autopilot used in this experiment was incapable of any adaptive response which might attenuate the strength of the failure-induced error signal. This of course was not the case with the adaptive human operator in the MA mode, yet despite this advantage AU detection was found inferior. (2) One characteristic substantiates the data of all subjects that have run on earlier pretests in our laboratory.* Namely that AU detection, while not invariably less accurate, is always considerably slower, a point of note in considering that Young reported only latency data in describing his findings.

In speculating upon the reason for this shift along the speed-accuracy bias with participatory mode, it is possible that the greater motor involvement of the operator while engaged in active tracking responses (MA mode) serves to "prime" the response mechanism or lower the detection criterion, causing a detection response to be triggered on the basis of

*The present data represents the culmination of 4 experimental pretests. The third of these pretests, identical in format to days 1 & 2 of the current experiment generated that data that were presented at the June meeting in Cambridge.

less perceptual evidence. In the present data, while this evidence is less in the manual mode, it is presumably of equal or superior quality, thus maintaining nearly equal accuracy despite the more rapid responses.

A second noteworthy aspect of the data relates to the interaction of participatory mode and dual task load, again a consistant finding across all 3 subjects (and the previously run subjects in the pretest,. This finding is in accord with the argument made in the introduction relating to the greater vulnerability to the withdrawal of attentional resources of the less demanding monitoring task. It is also a finding which, if substantiated has certain practical implications. These concern the conditions under which the greatest advantage is gained in terms of failure sensitivity by keeping the operator in the control loop -- that is, conditions of higher workload. Unfortunately these conditions are at the same time those under which other arguments might dictate replacement of the human operator by the autopilot.

With regard to the cues utilized in detection and the sources of MA superiority outlined in the introduction the data were not greatly illuminating. The greater stability of the internal model remains a viable source of superiority that is in no way contradicted by the present results. Similarly the data also suggest in the MA mode the activation of all three "failure information channels" described in Figure 1, and thus a greater number of information channels than in monitoring. Error information is clearly present (Figure 4b), while control adaptation appears to take place as well, leading to the availability of proprioceptive information. This is indicated by the reduced error profiles of the manual condition - particularly for hit trials (Figure 4b), and by the modest increase in response velocity (Figure 4c). One puzzling aspect of the data is that neither the ensemble averages nor the multiple regression analysis provided any evidence that response velocity was actually <u>used</u> in reaching the detection decision. It is of course possible that this particular measure was not the appropriate one to tap the process of control adaptation, and further investigation is necessary to determine alternative measures.

It is apparent that the greater workload of the MA condition does not greatly interfere with the detection process (or if it does, this interference is more than compensated by the favoring factors of model stability and information channels). The task of detection is seemingly one that is sufficiently integral with that of controlling, that a kind of dual task facilitation results when the two are performed concurrently. This is a characteristic that preserves detection sensitivity in the face of increasing dual task demands.

Finally, some mention should be made concerning the presence of individual differences. To some extent these are inevitable, particularly in a task configuration as complex as the current one, requiring dual task performance in the AU mode and triple task performance in the MA. Given the subject's flexibility to allocate resources differentially to the two or three tasks, as well as his ability to adapt various criteria on the speed-accuracy detection bias, it is perhaps somewhat surprising that the individual subject data in Figures 2 and 3 are as consistent as they are. Nevertheless, the importance is acknowledged of acquiring more data to replicate and substantiate the trends reported here.

Curry, R. E., and Gai, E. G. Detection of random process failures by human monitors. In T. B. Sheridan and G. Johannsen (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Ephrath, A. R. Detection of system failures in multi-axes tasks. Proceedings of the 11th annual NASA-University conference on manual control. NASA TMX, 1975, 62, 454.

Gould, E. E., and Fu, K. S. Adaptive model of one human operator in a time-varying control task. Proceedings of the second annual NASA-University conference on manual control, NASA SP, 1966, 128, 65-96

Green, D. M., and Swets, J. A. Signal detection theory and psychophysics, New York. John Wiley and Sons, Inc., 1966.

Hays, W. L. Statistics. London: Holt, Rinehart, and Winston, 1969.

Jex, R., McDonnell, J. D., and Phatak, A. V. A "critical" tracking task for manual control research. IEEE Transactions on Human Factors in Electronics, 1966, 7(4).

Johannsen, G., Pfendler, C., and Stein, W. Human performance and workload in simulated landing-appraoches with autopilot-failures. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Jordan, T. C. Characteristics of visual and proprioceptive response times in the learning of a motor skill. Quarterly Journal of Experimental Psychology, 1972, 24, 536-543

Keele, S. W. Attention and human performance. California: Goodyear Publishing Company, 1973.

Kerr, B. Processing demands during mental operations. Memory and Cognition, 1973, 1, 401-412.

Klein, R. M., and Posner, M. I. Attention to visual and kinesthetic components of skills. Brain Research, 1974, 71, 401-411.

McNicol, D. Signal Detection Theory. London: George Allen & Unwin Ltd. 1972

McDonnell, J. D. A preliminary study of human operator behavior following step changes in the controlled element. IEEE Transactions on Human Factors in Electronics, HFE Tech. 1966, 125-129.

Miller, D. C., and Elkind, J. J. The adaptive response of the human controller to sudden changes in controlled process dynamics. IEEE Transactions Human Factors in Electronics, HFE, 1967, 8 (3), 218-223.

Pew, R. W. Human perceptual-motor performance. In B. H. Kantowitz, (Ed.), Human information processing tutorials in performance and cognition. New York: Wiley, 1974.

Sheridan, T. B. Preview of models of the human monitor/supervisor. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Sheridan, T. B., and Johannsen, G. Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Swets, J. A., and Kristofferson, A. B. Attention. Annual Review of Psychology, 1970, 21, 339-366.

Taylor, M. M. Detectability theory and the interpretation of vigilance data. Acta Psychologica, 1967, 27, 390-399.

Veldenhysen, W., and Stasen, H. G. The internal model: What does it mean in human control. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Watson, C. S., and Nichols, T. L. Detectability of auditory signals presented without defined observation intervals. Journal of Acoustic Society of America, 1976, 59, (3).

Welford, A. T. Fundamentals of Skill. London: Metheun, 1968.

Young, L. R. On adaptive manual control. IEEE Transactions on Man-Machine Systems, V. MMS-10, 1969, 4, 292-331.