NASA CONTRACTOR
REPORT

NASA CR-125
C. 1

NASA CR-1258

HUMAN VISUAL SAMPLING PROCESSES: A SIMULATION VALIDATION STUDY

by John W. Senders, Jaime R. Carbonell,
and Jane L. Ward

Prepared by
BOLT BERANEK AND NEWMAN, INC.
Cambridge, Mass.
for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JANUARY 1969

NASA CR-1258

HUMAN VISUAL SAMPLING PROCESSES:
A SIMULATION VALIDATION STUDY

By John W. Senders, Jaime R. Carbonell,
and Jane L. Ward

Issued by Originator as Report No. 1681

Prepared under Contract No. NAS 1-5059 by
BOLT BERANEK AND NEWMAN, INC.
Cambridge, Mass.

for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the persistent critical attention of Dr. Rayford Saucer of NASA-Langley; and the high order of technical skill of Dr. Christine Kris who was in charge of the Electro-oculograph recording. Our pilots: Lt. Col. Dante Memmolo, Capt. James Fisher, and Lt. Col. Phillip MacNamara were, of course, the keystones of the project, the only truly indispensable participants in the program. Mr. Eugene Kwasniewski maintained the equipment in running order from day to day, no mean task.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

The results obtained in some fundamental investigations (BBN Report No. 1246) strongly support the notion that the behavior of the pilot or controller/monitor is largely determined by, and calculable on, the basis of physical characteristics of the system. Based on these results a validation study more nearly approaching operational situations was undertaken. This study used an instrumented flight simulator, operational pilots, and complete measuring, recording and analysis techniques necessary to validate (and improve) the methods offered by the theory. The applied goal of our program was to make possible an a priori evaluation of the work load which any set of information sources to be monitored and controlled will place on the man. We obtained the following results: a measure of the goodness of prediction of the Nyquist models developed in the basic study, and the development and a measure of goodness of prediction of a new queueing/cost-effectiveness model.

2. DISCUSSION OF BACKGROUND INFORMATION

There have been two rather distinct sources of thinking and experimentation which have gone into this research project. One of these stemmed from the classic observations made at Wright-Patterson Air Force Base in the period between 1949 and 1954. The other rests on some of the early attempts to utilize information theoretic notions in quantifying some human "information processing channel capacity." The WPAFB Studies of pilot eye-movements (Refs. 1-9) were directed at finding out the patterns of eye-movements actually used by pilots and the interpretation of these as indicators of the relative importance of the various instruments. To quote from Reference (1):

> Aircraft instruments are designed to provide pilots with information needed to control aircraft in flight—information such as the attitude of an aircraft with respect to the ground, its location in three-dimensional space, and the rate of change of its attitude and location vectors. Knowledge of how pilots use their eyes when they are flying on instruments, i.e., how they obtain data from separate instruments in order to combine bits of discrete information into a total picture of "what the aircraft is doing," is fundamental to a basic understanding of the function served by aircraft instruments. Such knowledge should guide the engineer in designing functional instruments. It can form a scientific basis for improving the

design of aircraft instruments, increasing the efficiency of pilots, and simplifying the task of instrument flying.

The present series of studies of eye-movements of pilots was undertaken to answer such questions as the following: How often is each instrument checked during particular maneuvers? How much time is required to check each instrument? What percentage of the total time available during critical maneuvers is spent in obtaining information from each of the different instruments? How are the frequency and duration of eye fixations influenced by factors such as instrument design, instrument arrangement, instrument lighting, pilot experience, and the particular maneuver being flown at the time?

And -

It is reasonable to assume that frequency of eye fixations on any given instrument is an indication of the relative importance of that instrument. The length of fixations, on the contrary, may more properly be considered as an indication of the relative difficulty of checking and interpreting particular instruments. The pattern of eye movements -- i.e., the link values between instruments -- is a direct indication of the goodness of different panel arrangements.

The information theoretic notions were discussed by Senders (10, 11, 12) in 1955-1958 in a series of papers directed toward the problem of the division of attention among a number of instruments. We quote from one of this series (12):

> Every time a new instrument is added to the instrument panel, it is feared that the visual system will be overloaded. Until we have some way of estimating (1) the visual load capacity of the human operator, and (2) the load presented by any particular system, such a statement cannot be contradicted or confirmed. It would be of considerable value to the aircraft designer, as well as of such general interest, to have such estimates.
>
> It is difficult, if not impossible, to estimate the information presented by a continuously varying instrument if consideration is given only to the instrument itself, apart from its use. It is still more difficult to estimate the total information flow from a display consisting of a multiplicity of instruments differing from one another in a variety of ways.

For a complex of M instruments it should be possible, therefore, to determine the total time (to be) spent by measuring or calculating W_i (signal bandwidth) and Q_i/N_i (signal-to-error ratio) for each instrument, multiplying the calculated frequency and duration of fixation, and summing over M. If one follows this approach the minimum utilization time for M instruments, Min T_M, is given by:

$$\text{Min } T_M = K \sum_i W_i \log_2 4^C \frac{Q_i}{N_i} \qquad \text{(Ref. 12)}$$

We quote from Reference 12:

> Let T_o be the duration of the duty cycle during which all instruments must be observed. Then, if $T_o > \text{Min } T_M$, one can try to add another instrument on the basis of known parameters of the flight condition, the sensing element, and the frequency response of the instrument, Q_i from the flight condition, and N_i from the pilot's instructions or set, such an overall evaluation of a panel is not impossible.

The data upon which such an information analysis could be performed would consist of graphic or numerical recordings of instrument readings during real or simulated flight conditions, and of the accompanying pilot eye-movements. From the instrument readings the power and maximum frequency of the instrument "noise" could be computed, and the above hypotheses could then be tested. Varying W, Q, and N (by instruction or determined by post-experimental analysis) and measuring performance as affected by flicker rate and light-time fraction should provide us with functional relationships among these many variables. These should, in turn, indicate the relation between the theoretical sampling values and those which actually can be used, and between the computed information per presentation and the time required for its assimilation.

The sampling notions themselves are related to other works in which periodic sampling (intermittence) was imposed on the human tracker as an experimental attack on the problems of voluntary observing behavior. The comments made in these earlier papers are still valid. To quote from reference 10:

> Most research involving visual-motor tasks has been done with continuously presented stimuli. However, the actual use of virtually all visual displays involves either the intermittent presentation of information or its intermittent receipt by the operator. Radar systems, for example, scan areas (or volumes) sequentially and generate visual displays which may have repetition rates ranging from (as few as ten) per minute to as high as 10 per second. Even when a display presents information continuously, the man who monitors it usually has more than one thing to attend to and therefore divides his attention between tasks. He thus may impose on the displays his own intermittency or sampling behavior ...
>
> In this experiment the subjects were required to track two pointers, each in a separate instrument. They were scored on the amount of time that both pointers were held simultaneously within their marked target areas. The instruments were illuminated intermittently but simultaneously by means of a rotating sector disc in front of a light source. Thus the S could see only

> periodically, but this did not force him to sample the instruments periodically... Thus it is of interest to study the combined effects of imposed intermittency of a periodic nature and the intermittency (probably aperiodic) produced by the operator's sampling. One might hypothesize that there would be an interaction between the intermittency inherent in the S's sampling of the component instruments and that imposed by the periodic illumination. Such an interaction would perhaps lead to a local maximum of performance at some frequency below fusion or the frequency at which performance is as good as with fusion... In the case of the apparatus used we are concerned not only with shifts of attention but also with concomitant shifts of the eyes necessitated by the physical separation of the two aspects of the display. As a result, we might expect the frequencies at which shifting could occur to be much lower than in the case of the PSMT*, and consequently a different relationship between performance and rate of interruption. If the operator shifts to the left display, for example, and finds it dark he is forced to wait until the next flash occurs.

And further (11):

> The implications of the results...are of interest. Since the peak values obtained are less than those obtained for continuous exposure of the

---

*Pedestal Sight Manipulation Test

stimuli, it may be deduced that the normal sampling behavior of the human tracker is aperiodic. If this had not been the case the peak value should have risen to that of the continuous case for some value of presentation frequency. It may be further stated that periodic sampling at a rate sufficient to transmit all the information in the continuous time functions which are being tracked is not a suitable mode of operation for the human operator. It is probable that the human operator takes account not only of the instantaneous position of the indicator but also of its first derivative and adjusts his sampling to accord with some combined function of the two aspects of the display... If a man(S), tracking or observing time functions on two or more separated displays (as in the operation of an aircraft or automobile), had mounted in his eyes lights which shone on that which he fixated, and no other lighting were present, an outside observer and the man himself would have entirely different opinions of the process being carried out. To S the situation would be very similar to the usual one in which the stimuli which he is tracking or observing are continuously illuminated. The external observer, on the other hand, would see a situation in which, with only a casual observation, he would perceive no order or regularity, and he would be astonished to find that S performed as well as under conditions of continuous illumination. In the experiment described above, we have reversed the procedure and imposed on S

> a regular and easily describable set of conditions in which the stimuli are exposed equally often and for equal durations in all cases of a particular trial.
>
> When we permit S to sample freely, and record and analyze the sequence of fixations which he makes, we certainly expect that there will be a relationship between the rate at which information is presented on a particular dial, or, in other terms, the bandwidth of the signal presented on the dial, and the frequency with which S looks at that dial. It is this functional relationship which is being sought.

Unfortunately, the lack of information about the nature of the displayed signals in the WPAFB studies prevented the use of those results to test the primitive theory proposed in 1955 (12). As a result a series of laboratory studies was done and the adequacy of the various models to predict visual behavior in laboratory situations was established (13). Since actual flight vehicles are much more complex than laboratory devices, a validation effort was undertaken using a modified Link C-11B Flight Simulator. The goal, then, is the calculation of the workload imposed by a system on its operator on the basis of the signals which are displayed to him. This is, of course, tantamount to relating the operator's visual behavior (as an index of workload) to the dynamics of the vehicle and to the mission which the system is performing.

The term "workload" subsumes a number of related concepts. From a strictly operational point of view it means some measure of the performance decrement on an alternative or secondary task when the task whose workload is being measured is being performed to a criterion level. Alternatively workload can be conceived of as a direct measure of the "effort" being put forth by a human operator while performing a task independent of the performance of the task itself. Or, as still another alternative, one can calculate the theoretical demand of a task assuming the human operator to be a particular kind of ideal operator defined in some analytical or heuristic way. In a general way we would like to know "how hard a man has to work" in order to perform a particular task. It is an important thing to know or to be able to estimate since the probability of success of a man-machine system depends on the reliability of the man as well as on that of the machine, and to a large extent, the reliability of the man is a function of the load that is placed upon him.

Knowles (14) summarizes the notion of operator workload as follows:

> ...in the design of equipment and the development of operator and training procedures it is important to be able to answer questions such as: How easy is this equipment to operate? How much attention is required? How much learning is involved? How well will the operator be able to perform additional tasks? All these questions deal with some aspect of what has come to be called operator workload. Essentially, how busy is the operator?

The unanswered questions are:

How does one measure the workload which is imposed on the operator by a system in being?

How does one predict the workload which will be imposed on the operator by a system in prospect?

The former question may be answered to some degree by the techniques described by Knowles (14) which involve operator loading tasks. Briefly, an operator loading task is one which is to be performed at the same time that the operator is engaged in performing the task under investigation to some established criterion level. The reduction of performance in the loading task as compared with the level achieved in the absence of the primary task is an index of the loading placed on the operator by that primary task. The basic assumptions underlying this methodology are:

1. The operator is a single channel system.

2. The channel has a fixed capacity.

3. The capacity has a single metric by which any task can be measured.

4. The components of workload are additive linearly no matter what the sources of the load.

For such a system to work, there must also be an assumption that the operator can so schedule his working that the primary tasks are always adequately dealt with, and are mutually non-interfering.

Such a "workload" measurement technique depends on the availability of some kind of hardware or system simulation, and is therefore not applicable to the earliest design stages. Furthermore, it does not attempt to distinguish the load placed on the operator by poor interface design from the inherent load which would exist in a perfectly human engineered machine. The only way in which one could use an external loading task to accomplish this latter goal would be to use it in a large number of simulators possessing different interfaces.

An alternative method of estimating operator workload is that used by Siegel and Wolf (15). This method depends on "source data" about the task which are derived from a prior "task analysis." The source data are concerned with such factors as: the average time required for the operator to perform the task; the standard deviation of that time; probability of performing the subtask successfully; the necessary waiting time before which the subtask cannot be begun; and the subtasks which must be performed next in the events of success or failure of the subtask. The technique introduces stress and its effects on performance into its calculations and generates distributions of trials on which the operator either completes the tasks successfully or runs out of time. Then the distributions for alternative system forms can be compared. Such a technique is useful for the examination of existing systems but could be used only with difficulty for systems which are hypothetical or conceptual.

Ekstrom (16) applied both a loading task and an analytical approach to the estimation of the relative workloads associated with two modes of operation of an automatic flight control system and found that the loading task confirmed the results of the analysis to a marked degree. The technique used in her analysis was essentially that described by Lindquist and Gross (17). It involves a Second-by-Second Operational Analysis which provides instrument use-time estimates as a basis for laying out panels. It also provides estimates of the loading as a function of time placed on the pilot. This is used as a means of locating critical periods in the mission. The means of arriving at many of the numerical estimates is not too different from that described by Siegel and Wolfe (15).

The analysis of Lindquist and Gross (17) is largely based on the application of the theoretical analysis proposed by Senders (12); and this in turn had led to the more elaborate notions expressed in the present paper.

The assessment of the loading to be placed on an operator should come at the earliest moment in the design of a man-machine system. Ideally it should come before prototype design and construction and should be based on an analytic and theoretical foundation. Nor should its application be limited to the human engineering design of the man-machine interface but should rather lead to estimates of the reliability of the human component of the total system. The loading task approach depends on the existence of a simulator or a prototype of the actual system hardware; the analysis of Siegel and Wolfe (15) depends on opinions of operators of similar systems. What we propose to do here is to go from abstract specifications of the system to estimates of the loading which will be placed on an operator.

3. SEVERAL THEORETICAL APPROACHES TO VISUAL SAMPLING

The various analyses which will be presented rest on the following assumptions:

1. Visual attention is the major component of operator workload.
2. The various signals which must be monitored demand attention in a way which is dependent on the characteristics of the signal and the required precision of readout of the signal by the human operator.
3. The human operator is effectively a single channel device capable of attending to only one signal at any time.
4. The probability of failure at any time is equal to the probability that two or more signals will demand simultaneous attention.

A discussion of various approaches to modeling human visual sampling behavior follows. Details of these models may be found in the literature (20, 22, 23).

3.1 Periodic Sampling

In a previous report (12) we have considered the question of a human monitor of some large number of informational displays. A theoretical model was suggested based on the assumption that the task of the monitor was to reconstruct the time functions presented on each of the displays. As a result it was possible to present equations based on the assumption of periodic sampling, which predicted quite well the average behavior of experimental subjects in the highly constrained laboratory situation (18). In fact, however, in that laboratory situation the task of the observer was specified to be the detection of a value of each of the signals which exceeded some preset value, not that of signal reconstruction. If one takes as the goal of a sampling system this latter task, it is possible to generate a completely different sampling strategy from that of the sampling theorem.

It is the case that, when the latter task is the one presented and there are equal signal powers and equal significant deviations, the periodic sampling model provides adequate estimates of the distribution of attention of the visual sampling behavior of subjects. However, the estimates are only of the means of distributions of intervals between observations. There is nothing in this model which permits an estimate of the distributions themselves or suggests that such distributions will in fact exist.

That there are such distributions is apparent from the data (13) and from purely logical considerations (19). Aperiodicity in the sampling of any one instrument would result from almost any configuration of instruments and bandwidths except for cases where all the signals had identical bandwidths and identical significant deviations, or in certain other equally unlikely cases where a totally periodic scanning process would be possible.

Faced with the observed data about the distributions of durations both of observations and of interobservation intervals, one is led to attempt to account for these distributions in a rational rather than in a descriptive way. The basis of this approach is that the interval between observations is a function of the value previously read, and the duration of an observation is a function of the value then being read by the observer. A discussion of this model follows.

3.2 Conditional Aperiodic Sampling

Let us consider the problem faced by a monitor of limited channel capacity confronted with many (more than two) signals to attend to, but concerned with detection of extreme readings rather than with reconstruction of the signal. Such a monitor may serve not as a channel for the transmission of a complete time function but rather as a channel for the transmission of discrete pulses in time. For any function one might assume that there is a threshold value of the function which calls for the transmission of a message, and all values of the function below this limit call for no transmission. In other words, a monitor observes the time functions and does nothing so long as they remain with a "safe" interval. When a function exceeds the limits of safe operation the monitor emits a signal which might be either the present value of the function or some other signal. (Of course, this includes the exercise of control functions as well.) We may now ask what the appropriate sampling strategy will be for this monitor. It is easy to see that if the permissible error, between the function as presented and the function as read, is equal to the amplitude of the function, no observation is needed. Similarly, if the permissible error approaches zero then the information to be absorbed per sample increases and a longer time will be required for the monitor to accept and transmit the information. If the function at the moment of observation has a zero value (i.e., its mean), then the next sample may be deferred until a time T when the probability that the function will exceed the limits of safe operation exceeds some subjective probability threshold or until the moment when a new reading will be statistically independent of the old, whichever comes first.

The discussion above is valid for a monitoring (i.e., open-loop) task. Serious complications might arise when the HO is engaged in a control (i.e., closed-loop) task, in which the closer he controls, the smaller the observed deviations are, but the more often he may have to look at the instrument. See Section 6.2 for a complete discussion.

If the limit of safe operation is some L standard deviations away from the mean, then, as time since the last reading increases, the variance of the distribution of signal amplitudes around the expected value also increases. Thus, there will come a time when the probability of a significantly deviant reading will exceed the probability threshold. At that point a sample might be taken. If the function when observed is greater than zero, i.e., is some fraction of the way toward the limit L, then the point at which the probability reaches or exceeds the threshold probability will in general come sooner, and the sample must be taken after a shorter interval. As the observed value of the function approaches L, the inter-sample interval approaches zero.

The above description assumes one possible strategy: the probability threshold criterion (20). Other sampling strategies might involve the sampling of the function at the moment when the probability of exceeding the limit is a maximum, or sampling when that probability exceeds a certain threshold, or sampling according to a "variable signal bandwidth" rule. Reference 20 presents analyses of such strategies, and shows how one might calculate the interval between a present observation of a signal and the next observation, granted that the signal possesses certain well-defined characteristics.

The various strategies are not necessarily mutually exclusive. The actual process of "condition sampling" is probably a combination of two or more strategies. The mathematical model which is appropriate depends on the momentary goal of the observer.

For any of these models, the maximum interval between observations would be that time for which the next sample is statistically independent of the former. In general, the intervals will be shorter than that. Thus, samples will be taken by the observer more often than would be calculated on the basis of the sampling theorem. If the observer, human or inhuman, can also detect velocities, then it has been shown that the maximum permissible interval between samples may be doubled (21).

Since some low velocities will be below threshold, we would expect that low-frequency signals would be sampled more often than would be predicted by any simple theory. This is observed to be the case in earlier laboratory studies (20).

Carbonell (22) proposed a model of visual sampling, which can also be considered as a conditional sampling model. Its other features, however, make it distinctive from other models of that class and therefore will be discussed separately (see Section 3.4).

Smallwood (23) proposed a related model for visual sampling in a monitoring task. His conception is based on the probability of exceeding the threshold, estimated on the basis of an internal model that the HO is postulated to formulate about each instrument.

3.3 Transition Probabilities (Link Values)

We have considered various models for sampling strategies which will permit an observer to achieve some specified goal. Either by applying the models to hypothetical or known signals, or by direct measurement of the relative times spent observing the various instruments or display devices in a man-machine system, we can arrive at estimates of the probability of fixation of each of the signal sources or instruments. The utility of such estimates is apparent: the greater the probability that a signal will be fixated, the more centrally should it be displayed in the visual field. The original series of pilot eye-movement studies was aimed at determining by direct measurement the various fixation probabilities and using them to determine the locations of the instruments then used in a variety of aircraft.

In addition to measurement of fixation probabilities, these studies also determined the successive pairs of instruments fixated. The goal was to establish "links" between instruments which could act as a guide to the placement of instruments relative to one another. Here again the utility of the estimated transition probabilities is apparent: the greater the probability of transition between two signals, the closer together they should be displayed.

One model for the transition process is a Markov chain that treats the observer as if he drew at random from the set of displayed signals with probabilities equal to the fixation probabilities each time a transition is to be made. Such an observer would make transitions between instruments without regard for any real or imagined relation between signals displayed. Although it is not contended that pilots in fact behave this way in aircraft, it is nonetheless the case that the predictions of the model are in close enough accord with the actual link values measured in flight to have served as a basis for instrument panel layout decisions. The predictions of the model and the results of laboratory studies (17,18) are even closer and suggest that in the laboratory situation the model can account for nearly all of the observed behaviors. The laboratory data were gathered on a set of random, unrelated signals. Thus, there would be no basis for selection other than the probability of fixation itself. A deviation on one instrument was not indicative of the signal which might be observed on any other. In the aircraft, on the other hand, there would be two processes which would determine the next item to be selected by the observer. If the prior observation found the signal inside the acceptable region, then the selection of a signal for the succeeding observation could be made on an appropriately weighted random basis. If the prior observation were of a significant deviation, then the coupling which existed between the displayed signals would lead to a rational selection of an instrument which might be presumed to be related to the observed deviation. Our expectation therefore should be that the demands of the simple model will constrain partially the sampling behavior of the pilot and that this will be particularly true of those signals which have the highest probabilities of fixation. The same instrument can be fixated successively. As a result, the observed frequencies and durations of observation will not be those that would have been calculated on the basis of signal bandwidths and accuracy requirements, but the distributions will be modified as a result of the possibility of these successive fixations (18).

References 18 and 19 present data which show the extent of agreement between model and behavior.

3.4 An Economic Queueing Model of Many-Instrument Visual Sampling

From readings, direct observations, and discussions with pilots of high performance aircraft, it is possible to assess some important characteristics of their task. There is a wide variety of instruments, though they can generally be lumped into two groups: "desired-value" instruments (altitude, attitude, speed, etc.) and "threshold" instruments (indicator lights, passing over a point on the ground, etc.). Some instruments may be high frequency (roll), but under certain conditions involve a not too high cost. Others may be of rapid fluctuations and involve a high cost (altimeter on a landing maneuver). Others may be very slow in changing, but the cost of some unexpected variation may be extremely high (fuel gauge). Obviously, the cost structure depends on the situation, i.e., on the type of maneuver the pilot is executing.

In relaxed moments, the pilot may space his readings far apart. Uncertainty grows with time and there is always the possibility of, for example, a sustained drift in altitude. In high performance planes, if the pilot does not control, the plane will not come back to its "zeroed" condition. So, in the absence of control, the instrument indications will not be zero-mean.

A model was developed under the present contract which attempts to explain and match the behavior of pilots under actual flight conditions. As such, and in contrast to other models, it places the pilot in a closed loop which includes his visual, mental, and manual tasks.

The model is based on the concept of the different instruments competing for the attention of the pilot. Some may be unimportant under a given flight condition, but many should be looked at, the urgency of doing so being measured by the risk incurred if the corresponding value is beyond a certain threshold. Costs are assigned to each instrument; at each sampling instant the decision as to what instrument to look at is based on comparing for the different instruments the combined effect of both the probability of exceeding the threshold and a cost of exceeding that threshold. Effectively, the instruments queue for the pilot's attention; the instrument with the highest priority at each instant is then served (looked at). Though this model can attempt to reproduce both relaxed and loaded (high-demand) flight situations, is especially addressed to the latter which is obviously the critical one for all practical applications.

Contrary to all other models, which are addressed to monitoring tasks, the present model does not assume that the signal of each instrument is zero-mean Gaussian. On the contrary, it postulates a continuous random walk with predicted distributions that, in the absence of control, are Gaussian with a mean equal to the last observation. Only the action of the pilot brings under control the signals which otherwise may tend to drift; probabilities of exceeding thresholds monotonically increase with time in the absence of control actions.

The central idea of this model is, however, its use of what may be called an economic or cost-effectiveness approach. It assumes that a pilot is capable of making a rational conditional decision (instrument selection) on that basis before looking at an instrument each time. The model postulates that he is trying to minimize the risk involved in not observing the other instruments.

A formal discussion of this model, together with implementation details of a simulation following it and a very encouraging statistical validation, appear in Section 6.3 of this report.

3.5 General Discussion of Models

The foregoing sections have discussed in a general way the mechanisms proposed to explain the distribution of visual attention in a multidimensional dynamic environment. The experiments which were reported in References 18, 19, and 20 provide validation for various parts of the theory, in particular the relationship between monitoring laboratory data and the model proposed in 1955 and in 1963. The results of this series of experiments provide strong support for the hypothesis that human observers behave in a way which is influenced by the bandwidths and accuracy requirements of the signals which they are to monitor. However, there were indeed failures of agreement between the models and the data and these in turn have led us to more sophisticated analyses of the determiners of the distribution of visual attention. Among these are the direct perception of visual rate, the assignment of differential costs to extreme deviations of different signals (22), the effect of nonveridical perception of the signal characteristics by the observer (23), and the effects of coupling and correlation between the various displayed signals (20). The models of References 22 and 23 have been tested by computer simulation against the earlier data obtained in 1954 and 1955 and reported in 1958 (13). Here it is possible to choose parameters such that the behaviors exhibited by the models are well within the range of those behaviors exhibited by human subjects. Thus, the first model of 1955 appears to describe quite well the kinds of behavior which result when observers are required to monitor some relatively small number of completely uncorrelated, bandlimited, normally distributed time functions and to report whenever any of these time functions exceeds some symmetrically disposed, arbitrarily selected limit value; Smallwood's model (the subjective one of Reference 23) has resulted in a close approximation to the actual behavior displayed by these subjects in the monitoring laboratory situation. Carbonell's model (the queueing model of Section 3.4 and Reference 22) has given excellent approximation to visual sampling behavior in actual flight simulations, as will be discussed in Section 6.3 of this report. The ultimate utility of such models, of course, may be two-fold. On the one hand, they provide a basis for examination of attentional mechanisms and for identifying the characteristics of the world of stimuli which actually serve to elicit observing responses on the part of observers; and, on the other hand, they provide a potential means of estimating a number of parameters relating to the way in which people will behave when confronted with real tasks in real systems. For the moment it is this latter application of the theory which concerns us. If these models (or any one of them) hold true for controllers and monitors of real systems, then we can calculate a number of things about such systems. Among these would be the frequency with which the signal would be observed. This, of course, means that we can estimate in a statistical way the interval between observations, their duration and the sequences in which they would be made. We could, for example, calculate the probability that the monitor would in fact observe a transient signal of any duration which appeared on an instrument. Thus some index of the "reliability" of the human monitor in detecting signals could be obtained. In order to accomplish all or even part of these ambitious goals our task must be to show that the behavior of human beings in real systems does in fact follow the rules laid down by the models. Of course, most real-systems are markedly different from the systems that have been traditionally used in the laboratory. The differences are of the following kinds:

1. The signals presented on one instrument are very much related to the signals presented on the other instruments whereas in the classic laboratory situation the signals were specified as being unrelated.

2. The nature of the control process used by the system controller modifies the short term statistical characteristics of the signals whereas in the monitoring laboratory situation the response activity of the monitor had no effect whatsoever on the behavior of the signal.

3. The distribution of signal values is non-Gaussian whereas in the traditional laboratory situation the signal is constructed to be Gaussian.

4. The signal does not show a consistent tendency to regress to the mean whereas in the traditional laboratory situation the signal is assumed to have a mean value equal to zero.

5. There are no fixed arbitrary limits on each instrument but, rather, these vary as a function of time depending upon the immediate requirements of the machine and pilot, and on the mission to which the combination is dedicated. On the other hand, in the classic laboratory situation, the limits were fixed, and equal for all instruments.

For these reasons, it was felt necessary to perform more realistic experiments to avoid the oversimplifications of the classic monitoring laboratory experiments. These experiments consisted in the simultaneous recording of both eye movements and panel instrument signals when actual pilots were flying an instrumented landing approach flight in a LINK trainer simulator. The rest of this report is devoted to a description of these experiments (Section 4) and subsequent data analyses (Section 5), and to the use of the data in validation studies of some models (Section 6). Special emphasis is placed on the queueing model which was designed in this contract with the actual flight situation in mind rather than the traditional monitoring one.

4. EXPERIMENTS

The instrumentation requirements of the ground-based simulator to be used in the validation studies were set forth as part of Contract No. NAS1-5959. The relevant portion of the contract statement of work reads as follows:

"The Contractor will formulate a logical and/or mathematical model of a sampling process and test this model for validity against data gathered from operational flight status, licensed commercial pilots or equivalent operating the C-11 trainer under the following conditions:

(a) constant course

(b) constant altitude

(c) constant speed,

provided that the recorded data shall include as a minimum:

(a) heading

(b) speed

(c) altitude

(d) pitch, roll and yaw attitudes

(e) pilot control deflections

(f) continuous recordings of pilot eye motions."

In addition, Part 3.3.1 calls for I.L.S. data records.

4.1 Instrumentation

The C-11B Trainer located at Bolt Beranek and Newman's Cambridge facility was placed in operational status according to the maintenance and adjustment manual furnished by the original manufacturer. In addition, the general flight characteristics of the machine were altered in the direction of increased performance so as to bring the machine more nearly in accord with modern fighter aircraft. The machine was flown by two rated instrument pilots. In their opinion, it was in general a "pretty good airplane" and suitable for the maintenance of skill.

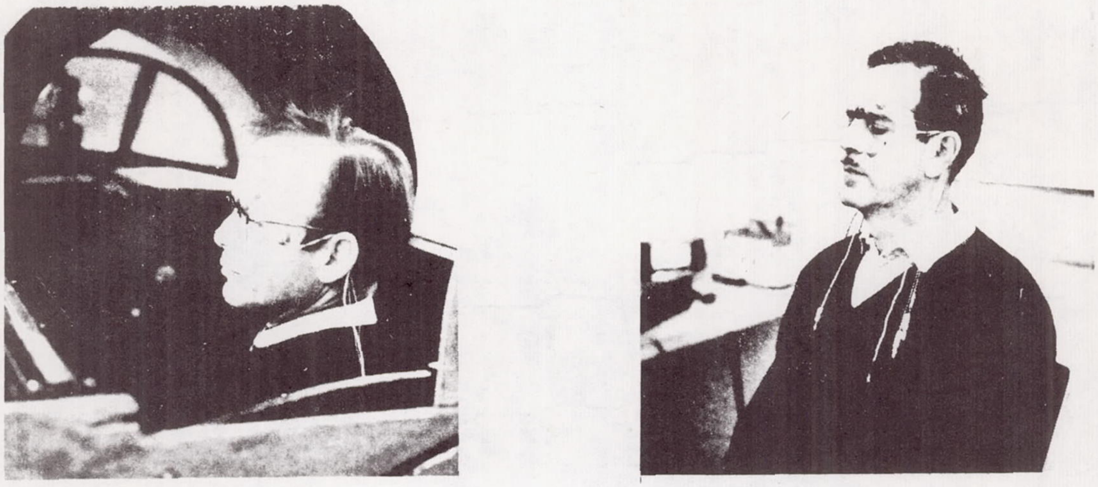
The canopy of the original C-11B trainer was lined with a translucent screen behind which fluorescent lamps were mounted to provide for simulated lightning. This liner and the lamps were removed, and suitable array of incandescent lamps was installed to provide light. Thus, the instruments and the controls in the cockpit were flood lit by white light. The canopy itself does not offer appreciable glare, nor does the light interfere with ordinary, safe operation of the aircraft simulator.

On each of the panel instruments an additional recording potentiometer or demodulator (as appropriate) was installed so that direct indications of all displayed functions could be obtained. In addition, potentiometers were installed on stick, rudder, flaps, and throttle to permit recording of the pilot's control activity.

In summary, then, to each of the servos which drives a flight indicator in the simulator there was attached an additional potentiometer to allow a DC signal to be obtained. Similar signals were obtained from controls. The signals were conducted to a Goodyear L-3 analog computer which served as a buffer between the simulator and the digital computer. From the analog computer the signals were conducted to a set of filters which eliminated line and potentiometer noise and then into the 20-channel analog-to-digital conversion system. The digitized signals were then fed into the PDP-1 computer which stored them on the drum in real time. The analog-to-digital conversion system was driven by an oscillator at the rate of ten cycles per second, scanned the twenty lines of data coming from the analog computer. Thus, all the visual information which entered the cockpit, and all the control information which left the cockpit was recorded.

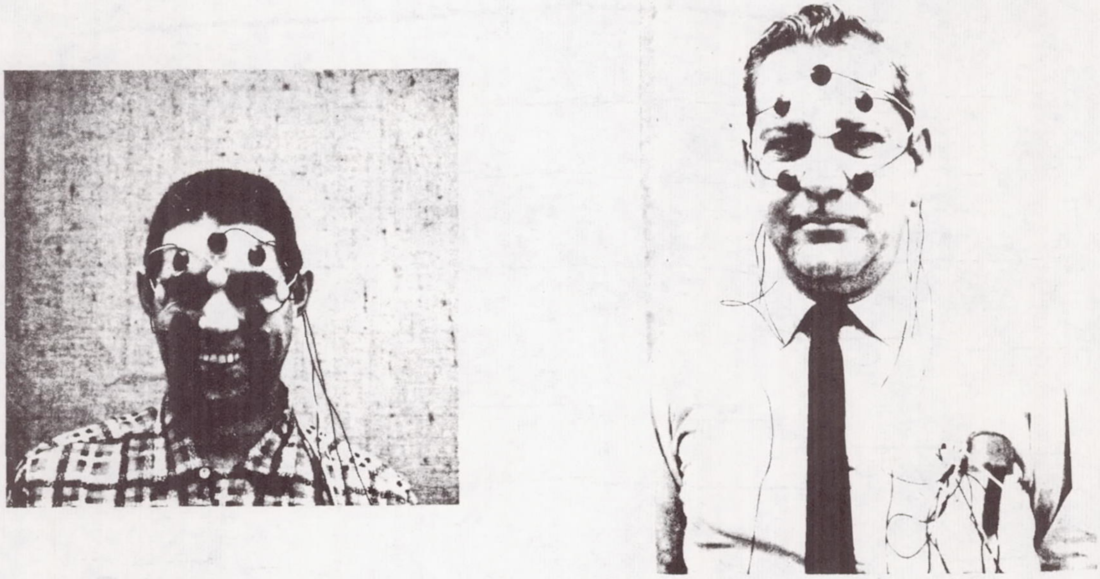
In addition, eye position had to be recorded (Ref. 24) if we were to test the predictions of the various models of the theory against actual practice. An electro-oculographic eye position recording system was used to record eye position (24). The means of obtaining such records are shown in Figure 1(a) through 1(d). The pilots opined that the process of wiring, and the wearing of the electrodes constituted less of an interference with ordinary flight behavior than that produced by the wearing of oxygen mask and helmet.

The procedure of application of electrodes was as follows: a small puncture was made in the skin using a 22-gauge disposable sterile hypodermic needle; Beckman electrode paste was then rubbed gently into the skin immediately around and over the location of the puncture; Beckman permanently chlorided silver electrodes with pressure-adhesive attachment rings were then applied and the wires led back over the ears of the subject pilot. The whole procedure required approximately five minutes. After the electrodes were applied a waiting period of 15 minutes allowed the reactions of the skin-electrode interface to stabilize and the subject was ready for the simulator. Prior to the actual take-off the subject would provide the electro-oculographer with a calibration run. Samples of such runs are attached as Figure 2. Such a run is produced by successively fixating the instruments in each of the three horizontal rows beginning at the upper-left-hand corner moving to the upper-right-hand corner moving back to the center left and so on. A brief pause would be made at each of the instruments in order to provide a step record of the sort shown.



(a) MR. J.W. SENDERS

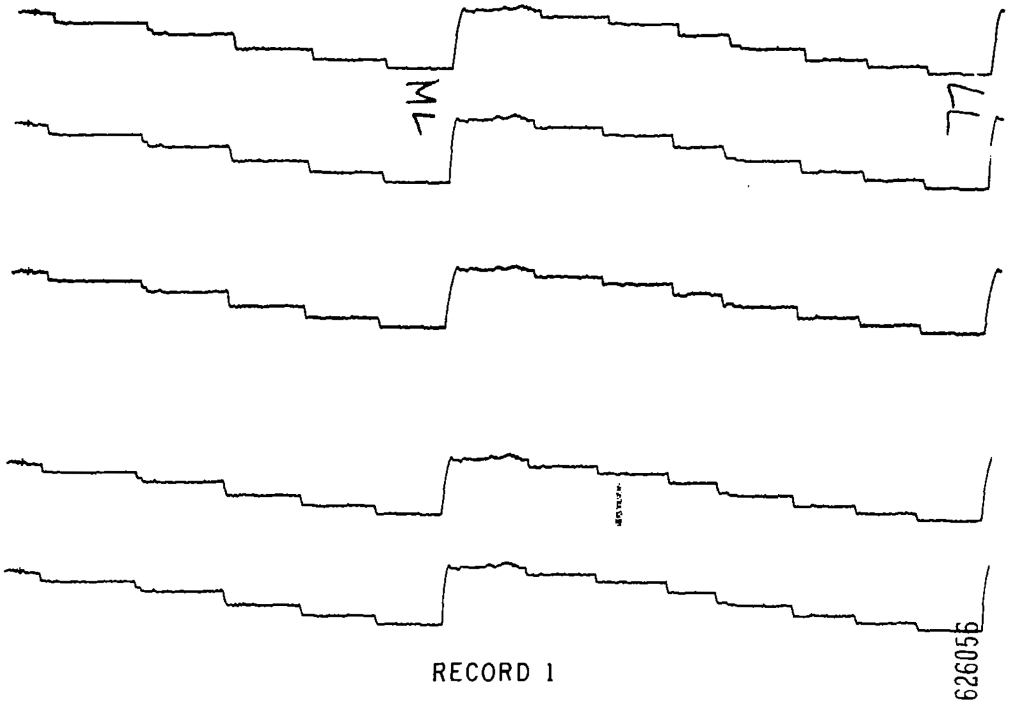
(b) COL. D.M. MEMMOLO



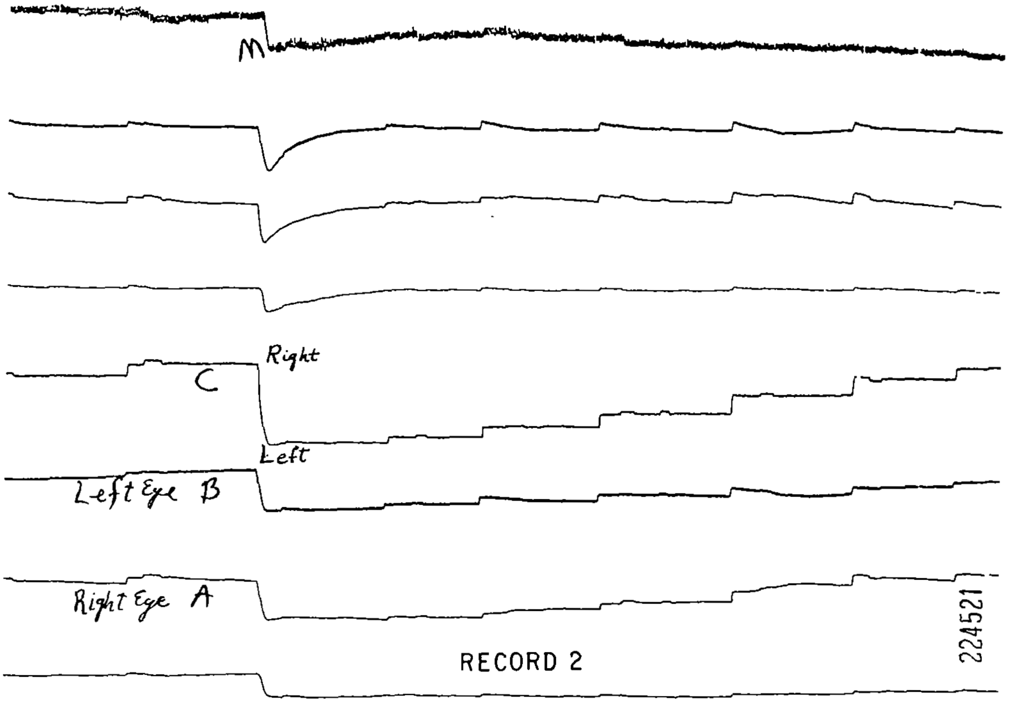
(c) LT. COL. P. J. MC NAMARA

(d) MAJOR J. F. FISHER

FIG. 1 ELECTRODE PLACEMENT

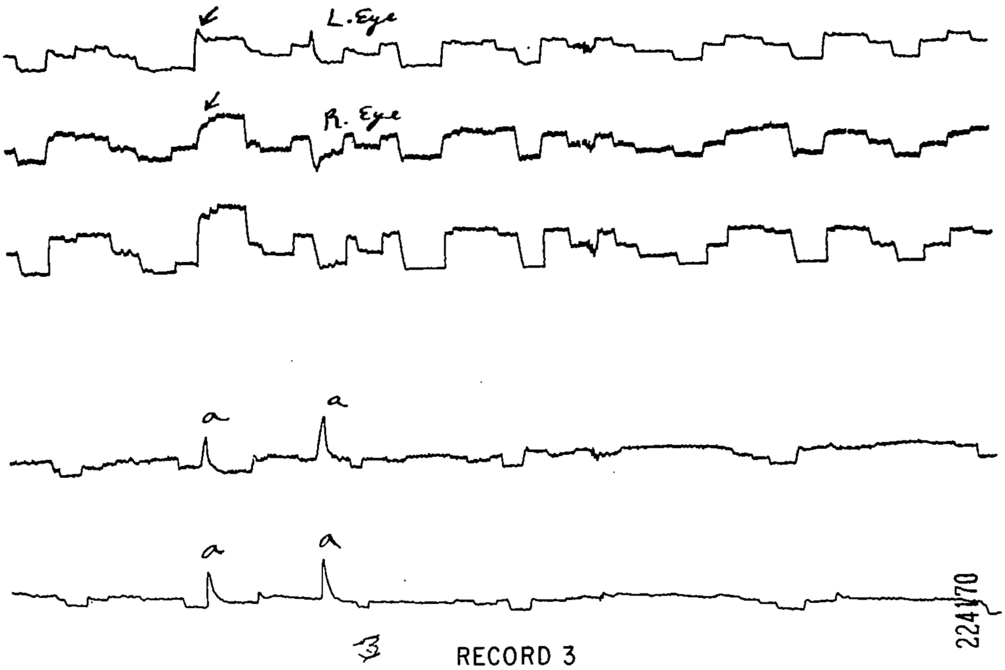


RECORD 1

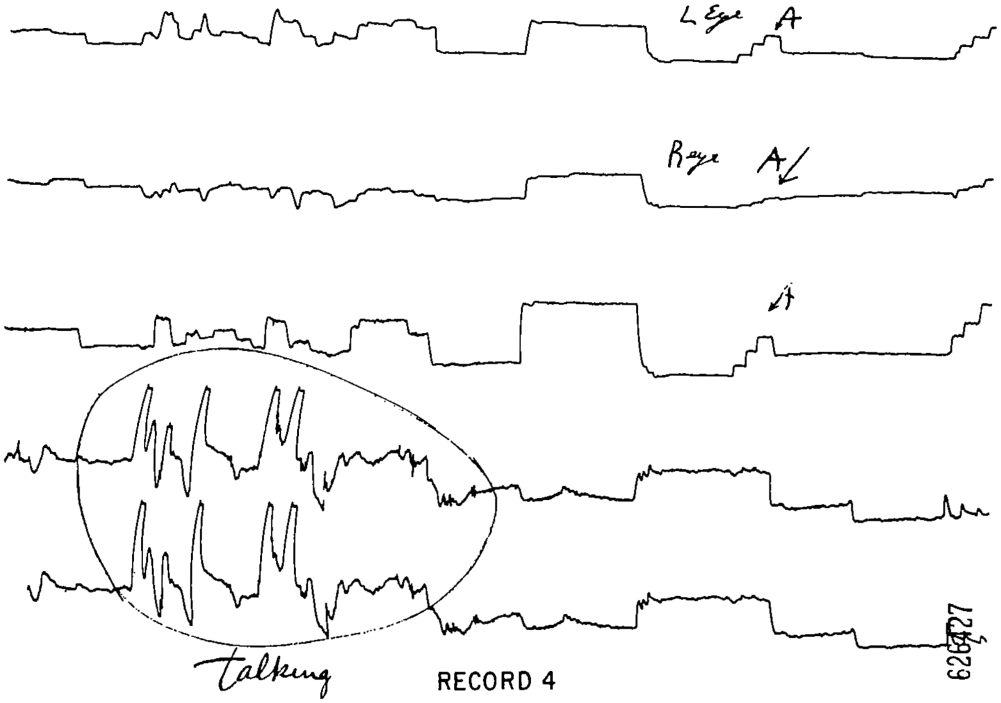


RECORD 2

FIG. 2 ELECTRO-OCULAR RECORDINGS



RECORD 3



RECORD 4

FIG. 2 (continued)

4.2 Discussion of Eye-Movement Records

Record No. 1 was taken on Mr. J. W. Senders and shows a number of readings derived from electrodes placed across the two eyes as shown in the photographs. The two upper traces and the two lower traces were made with temporarily chlorided silver electrodes, the center trace was that made with the Beckman permanently chlorided electrodes which were ultimately used. It can be seen that the signals derived from the two different kinds of electrodes are not significantly different in the short term. However, the decision to use the permanently chlorided electrodes was based on observations of drift over the long term which occurs with the temporarily chlorided electrodes.

Record No. 2 shows results for one subject, Col. Memmolo. The two lines marked A and B are recorded respectively from the right eye and the left eye individually. The curve marked C is recorded across the two eyes. It can be seen that in the series of fixations made by the eyes individually there are nonlinearities in the record which cancel when recorded across the two eyes. In addition, the total voltage derived, being equal to the sum of the two voltages, is more easily measured and discriminations more accurately made.

The four records at the top of the page were recorded respectively from horizontal and vertical electrodes and represent experimental attempts to improve the quality and discriminability of the signals. For our present discussion they serve no purpose.

The record marked 3 shows eye movements taken in flight and it is worthwhile to consider some of the characteristics of this method of recording which are demonstrated here. The points marked A represent blinks which appear strongly in the vertical electrooculographic potentials but are barely discernible in the horizontal left eye tract at the top of the record. As can be seen, the record for the right eye shows, if anything, a diminution of output with the blink rather than an increase, as indicated by the arrow. However, the various drifts, spikes and other nonlinearities in the eyes individually in this record taken during flight are, in general, suppressed by the technique of simultaneously recording across the two eyes. The vertical signals of this record are readable but obviously not as clear as those obtained from the horizontal movements. This is partly due to the eyelid and partly due to the fact that no convenient way was found at the time to add the vertical outputs of the two eyes. It is also apparent from looking at these records that there are times when the eyes do not appear to move together.

Record No. 4 shows the result of speech upon the eye movement recording. The circled segment represents voltages generated by movement of the skin under the electrodes and by muscular potentials during conversation. It is clear that these are largely suppressed in the bilateral recording but seriously interfere with the adequacy of the vertical signal recording. If one looks at the time marked "A", it is clear that a signal was obtained from the left eye which was barely represented in the right eye. One thing which became clear from examining these records was that the eyes of pilots do not necessarily move in concert, and that in some cases one eye dominates the other, which wanders in a relatively free way, apparently not seeing or not being used. It can be seen in this record that it is the left eye which is dominant and provides the major linear part of the signal.

4.3 Subject and Task Information

Earlier laboratory studies have shown that even as much as 30 hours at the rate of one hour per day of continuous observation of four simple unrelated signals did not permit subjects to reach true asympototic levels of performance. For this reason it was decided to use already trained and qualified jet pilots as subjects, and, further, to provide five hours of familiarization with the actual trainer and flight profiles which they were expected to fly. Four pilots were obtained from the Massachusetts Air National Guard, 103rd Fighter Wing, based at Logan Airport in Boston. From this group, three were chosen for the actual data taking. Table 1 provides relevant information about the pilots. The pilots were already familiar with the Logan Airport area and the radio aids of this region. It was decided to use an approach and final descent into the Logan instrument runway area as the task. The flight plan is that shown in Figure 3. The entire profile took somewhere between 8½ and 9½ minutes to fly. Table 2 shows the way in which time was distributed among the various legs of the flight plan. Following the five hours of familiarization, actual data runs were made using the well-practiced flight plan, and recording data with the techniques earlier described.

4.4 Data Taking

Each of the four pilots ran a series of three flights during which electro-oculograms were recorded on paper as well as being transmitted to the digital computer using the analog buffer as earlier described. Each run took approximately ten minutes which filled the available storage in the digital computer. It was discovered on the playback that the two ILS channels had not recorded during a part of the run and consequently the signals from the crosspointer instrument were not recorded. We therefore repeated a series of three runs with two pilots to provide data for the third phase of flight: the final approach to the threshold of the runway. The earlier set of data were adequate for the cross-country, letdown, and holding maneuvers since these maneuvers were not dependent upon crosspointer information.

TABLE 1

Brief History of Pilots

Name:	Dante M. Memmolo	Rank:	Lt. Col.
Age:	42		
Flying Experience:	7000 hours on Jets, Trainers, Bombers, and Transports		
Present Position:	Wing Staff Officer, Director of Flying Safety and Operations and Training Officer for Two Tactical Fighter Groups.		
Name:	Philip J. McNamara	Rank:	Lt. Col.
Age:	41		
Flying Experience:	5200 hours on Fighters, Trainers, Bombers, and Transports		
Name:	James F. Fisher	Rank:	Major
Age:	41		
Flying Experience:	5600 hours on Jets, Single and Multi-Engine Land Instrument Douglas DC-3		

TABLE 2

FLIGHT TIMETABLE

	8MI	MARKER	START 360°	COMPLETE 360°	OM	MM(½MI)
CLOCK TIME	53:00	53:30	56:15	58:07	59:35	62:15
TIME DIFFERENCE	0:00	0:30	2:45	1:08	1:28	2:40
CUMULATIVE TIME	0:00	0:30	3:15	4:23	5:51	8:31

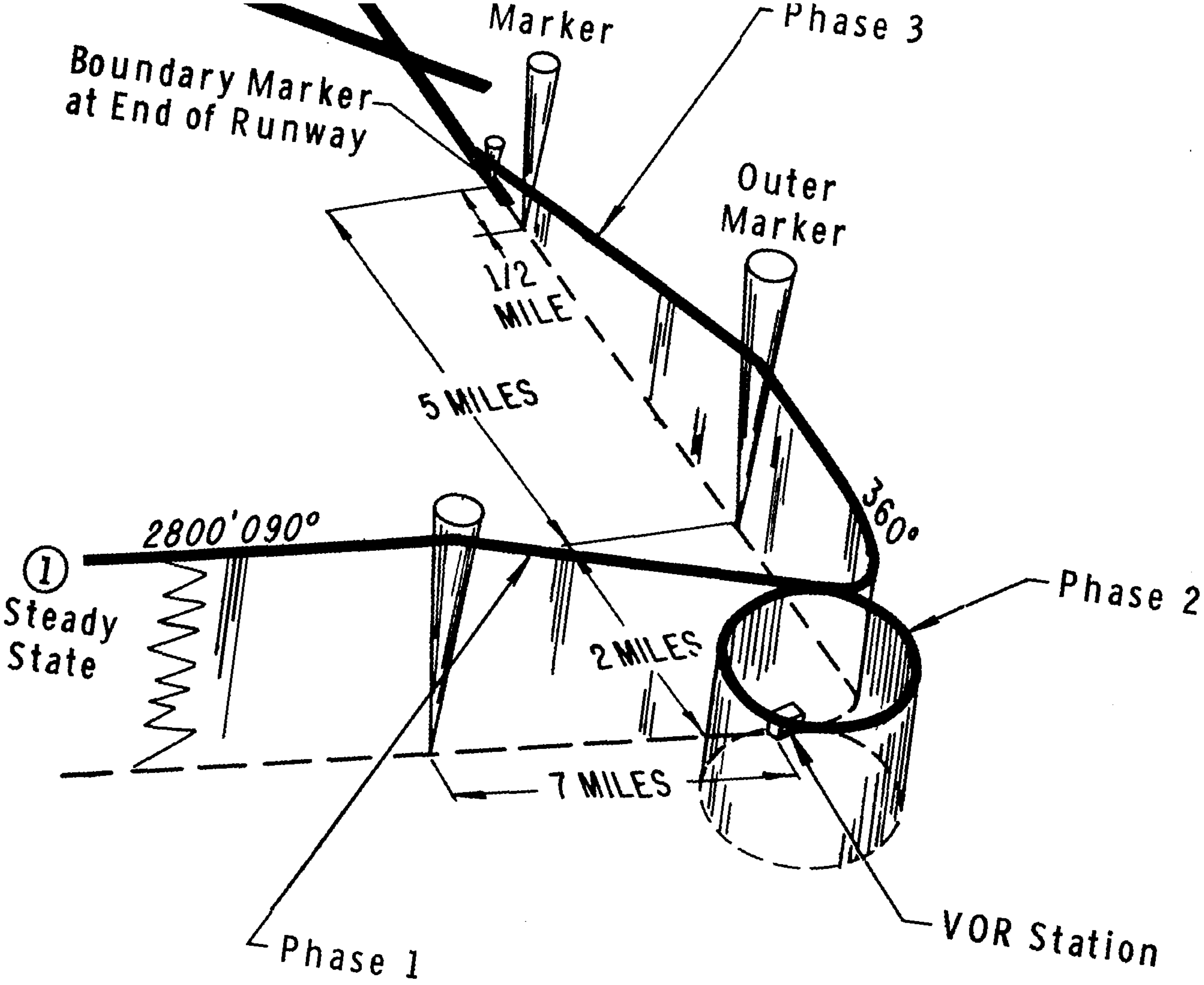


FIG. 3 FLIGHT PLAN

Data were taken at the rate of ten (10) samples per second. We possessed a total of 18 runs of 10 minutes each making a total of 108,000 samples of each of 18 channels of digital information. Our total data base then consists of 1.944×10^6 numbers recorded in digital form. This data base allows us to test any model of stimulus-determined behavior against the actual behavior exhibited by the pilots in this semi-operational situation.

4.5 The Pilot Questionnaire

A questionnaire was administered to the pilots involved in this study after the experimental part of the program was complete. The object of the questionnaire was twofold. Firstly it was an attempt to view this problem from the point of view of the pilot. Secondly, we wished to get subjective estimates of some of the parameters for the queueing model. For example, the pilots were asked what relative importance they placed on different instruments in different parts of the flight; they were asked for each instrument: (a) what minimum deviation is detectable; (b) what size deviation would cause them to initiate a corrective control action and (c) what would be a "panic" deviation. In addition, they were asked to describe as a function of time and the particular maneuver they were performing what their reaction would be from "calm" to "panic" if a particular instrument were unavailable to them. The results of a questionnaire of this type are not easily discussed in absolute terms. The application of the results of these questionnaires will be discussed in greater detail in the section of this report covering the queueing model.

5. DATA ANALYSIS

5.1 Data Storage

Digital analysis of the time sequences of each of the various instrument readings were required for the testing of the models; for example, processing the data corresponding to some ten minutes of altitude recording stored in the form of 6,000 samples, ten per second for six hundred seconds. The three phases of flight were each approximately 2,000 samples long. Segments of 1000 samples were transferred from digital tape into core memory and analyzed by a specially modified version of the BBN "Signal Analyzer" program (25). This process was repeated for each segment for each instrument for each pilot and for each run. The eye-movement records exist in two forms: 1) digital signals derived from electro-ocular potentials taken at the rate of 10 samples per second and stored in the computer on magnetic tape; 2) pen-tracings derived from the same electro-ocular potentials on paper as continuous time functions. Each of these two former methods were derived from a set of seven electrodes located around the eyes of the subject to give individual and combined lateral displacement and individual vertical displacements of the two eyes of the subject.

5.2 Eye-Movement Analysis

The raw records are similar to those shown in Figure 2 showing the calibration runs. It was necessary to pass each of the records through an editing procedure. The editing requirement arose largely as a result of drift in the electro-ocular potential records. The drift presumably was the result of minor changes in impedance through the electrode skin interface as well as through possible polarization of the junction at that place. The more serious problems were encountered in dealing with the vertical movement traces. It was not always possible to determine the absolute vertical position. However, the data possessed sufficient internal consistency so that a succession of two vertical movements in the upward direction implied that the beginning locus of fixation was on the lowest of the three possible vertical positions. In this way it was possible from time to time to anchor the judgments of the reader in determining the vertical position of the eye. In the horizontal case the magnitude of the deviations was larger and more uniformly related to eye position. Even here, however, some drifting did occur with the result that calibration points had to be found at various places on the run. Thus, if a measuring rule prepared on the initial calibration runs showed that the total pen excursion from the left most to the right most instrument was (let us say) one and one-half inches then if a place existed on the record where two horizontal locations one and one-half inches apart were in close proximity it would be assumed that the right most of these was position No. 8 and the left most position No. 1. Such cases occurred sufficiently often such that the minor drift which arose in the horizontal record could be easily taken care of. Certain portions of the electro-oculargraphic data were unreadable due to talking, muscle potentials, or instrumental artifacts. These were not edited and were coded as instrument No. 99. Where time marks failed constant paper speed was assumed. The paper speed was accurate to about 1%.

In general, as the readers gained experience on reading the electro-oculographic records, a great deal of internal consistency in the records became apparent and the speed of reading markedly increased. Certain patterns of behavior also became apparent which suggests that standard scanning patterns did exist even though these might be very different from one pilot to another. After the editing and identification procedure (which was done by hand) the paper strips were processed on a machine which provided a digital readout for excursions of the line in the vertical dimension, the paper being oriented horizontally in the machine. The reading point would be advanced to the beginning of an eye movement and a card punched. It would then be moved to the termination of that fixation and the new card punched. Also punched would be the identification of the horizontal and vertical locus of fixation. The new location would then serve as the beginning of a new fixation and so on until the records were complete. Thus, eye-movement time was not taken out explicitly but subsumed into fixation time. The records appear in the form of a series of IBM cards identifying for each fixation, the moment in time when it began, the moment in time when it ended and the point to which it was directed. This sequence of cards provided us with the means of determining the frequency with which each instrument was fixated, the average duration of fixation and the statistics relating to both frequencies and duration and also to calculate transition matrices of pairs of fixations from one instrument to another.

5.3 Signal Analysis

The signal data stored in the digital computer (mentioned earlier in this report) were available as an oscillograph readout. Figure 4 shows a complete set of waveform records taken during a flight of pilot MacNamara starting at point 700 and ending at point 4796, covering a span of 4,096 points, or 409.6 seconds. The various instruments are identified. (Figure 5 is a photograph of the instrument panel of the Link trainer used in these experiments.) The sensitivity and resolution of the system were adequate for our analysis tasks. It is of particular interest to examine the two channels which record eye-movement data. Although no attempt was made to analyze the data on these channels, it is clear that a program could be composed which would enable us to use the digital computer to obtain point estimates of where the pilot's eyes were fixated and, from what is stored on other channels, what he saw at the point of fixation.

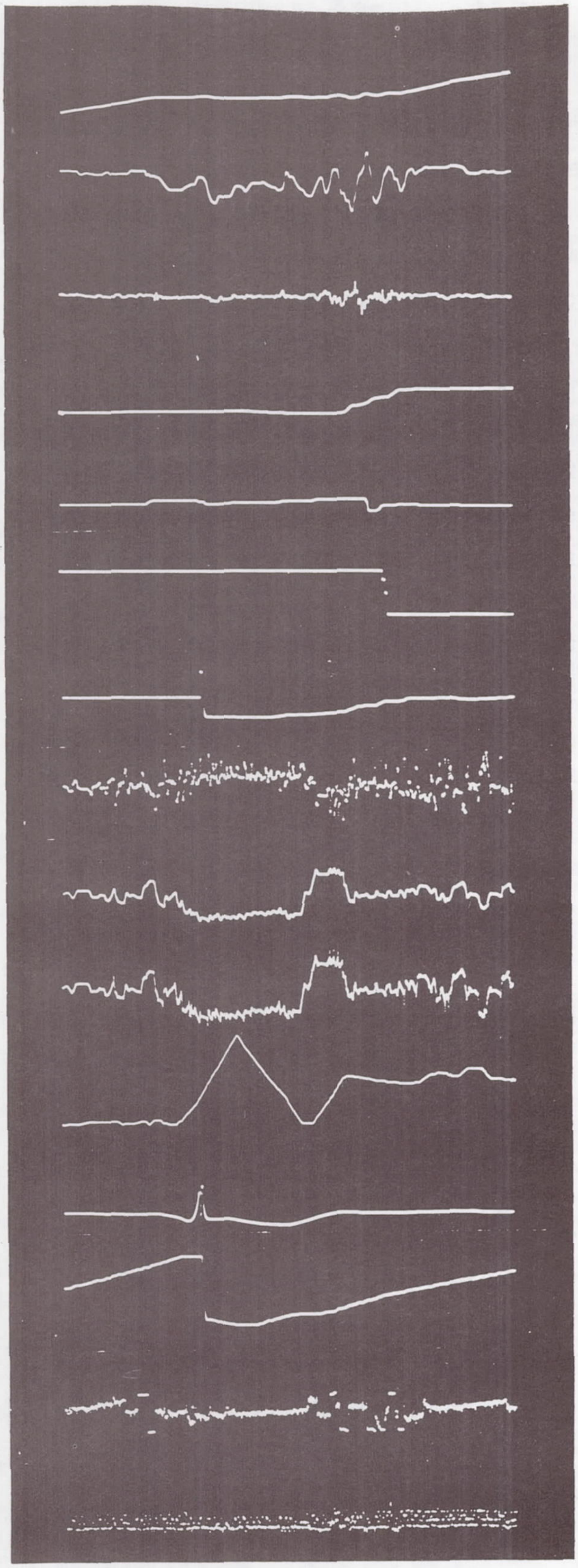


FIG. 4
SIGNAL DATA FROM PILOT MAC NAMARA, APPROXIMATELY 7 MINUTES OF FLIGHT.

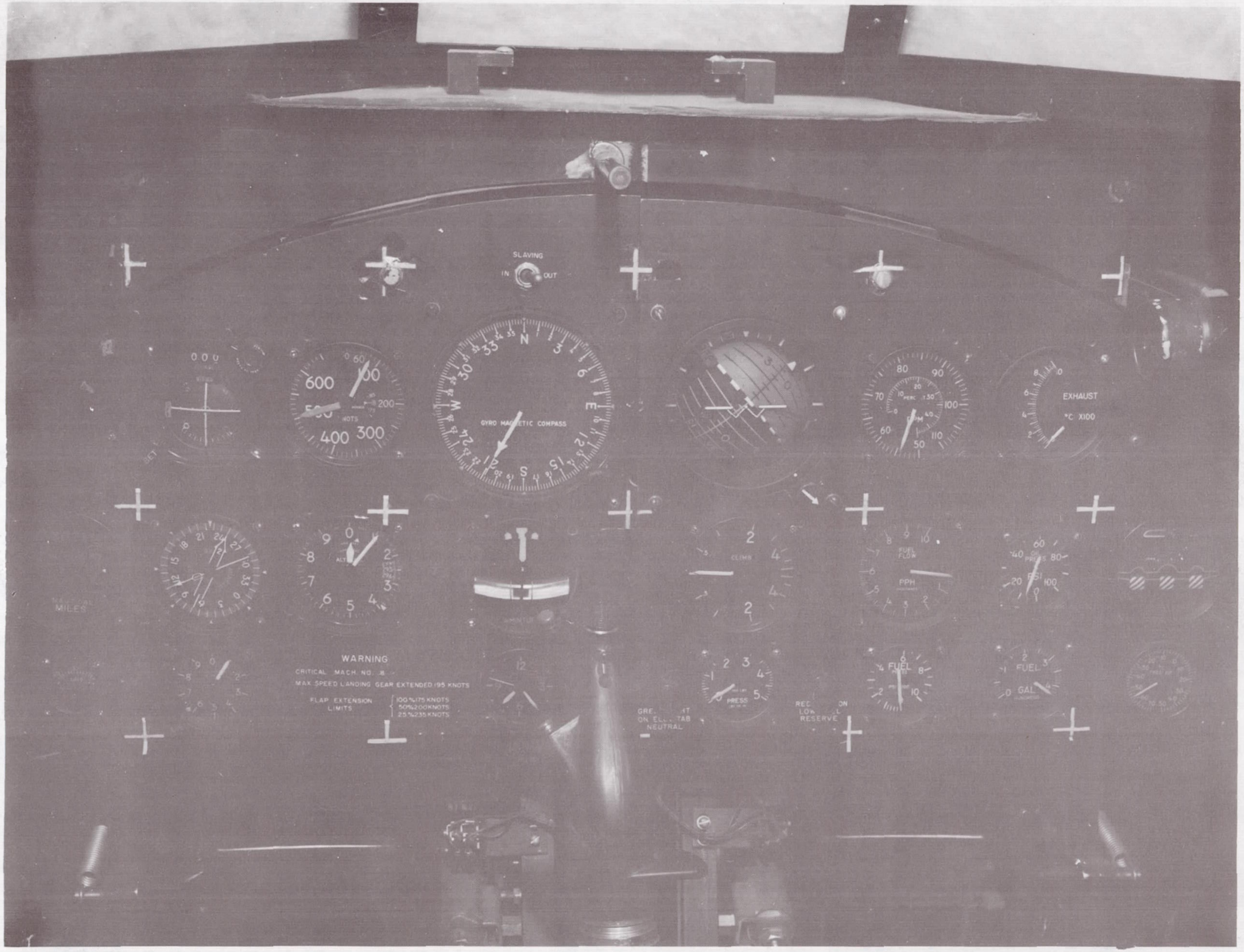
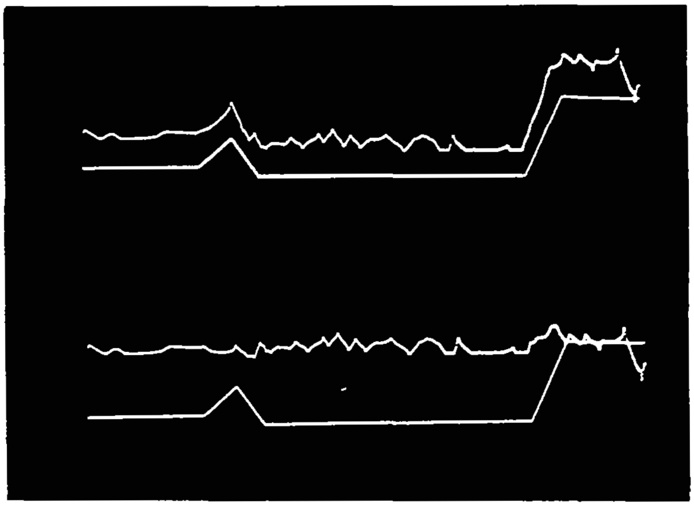


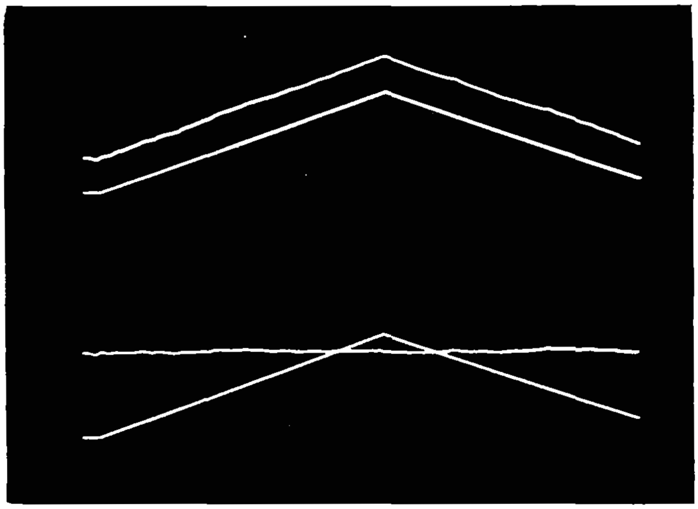
FIG. 5 LINK TRAINER INSTRUMENT PANEL

Since the models being tested assume that the pilot is responding to perturbations of signals around their intended values, estimates had to be made by the analyst of the goals of the pilots during each maneuver. The Signal Analyzer Program was modified to permit the calculations of spectral density functions on the perturbation data rather than on the signal as a whole. These perturbation data were obtained by subtracting what was estimated to be the intended time function (entered on-line during the analysis phase as a polygonal signal) from the actual data. Figure 6 illustrates this. The original program considered only the spectra of the signals around a zero mean. Figure 7 shows the spectra based on the original and on the now modified signals derived from the airspeed indicator. It is clear that there are significant differences in the relative power of the high- and the low-frequency components of the signal. The spectrum of the uncorrected signal shows a simple declining power as a function of frequency. When the signal is corrected the low-frequency power is reduced and the break frequency shifted from an estimated .09 radian per second to a value of .5 radian per second. The models being tested would assume that observing frequency would be related to the higher value of the two since it is indicative of the frequency with which perturbations occur around the desired time function of, in this case, airspeed.

Spectra of the corrected data were computed for three pilots for various portions of the flight path. The output photograph was scaled by means of a transparent overlay permitting a visual fit of a line at a constant power level and a line on the descending points which indicate diminished power as a function of increasing frequencies. From these two lines it was possible to obtain an estimate of the cut-off frequency and the slope with which the power falls off. Table 3 shows typical results of this analysis.

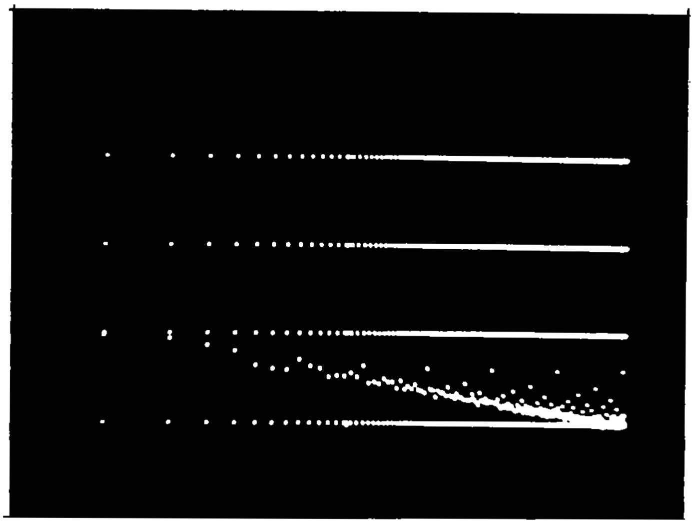


RATE OF CLIMB, PHASE II

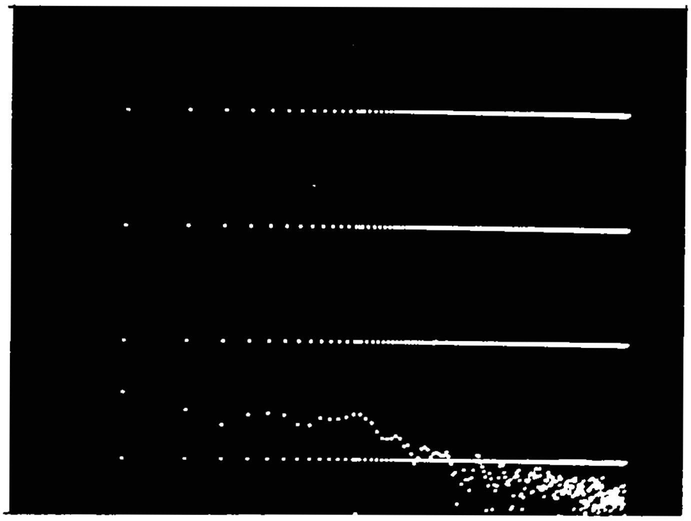


HEADING, PHASE II

FIG. 6 POLYGONAL CORRECTIONS



(a) UNCORRECTED



(b) CORRECTED

FIG. 7 POWER SPECTRA OF AIRSPEED INDICATOR

TABLE 3

POWER SPECTRUM CHARACTERISTICS OF DISPLAYED SYSTEMS

Phase I: Level Flight and Descent

Channel No.	Instrument	Memmolo		MacNamara	
		Slope db/oct	Cut-off rad/sec	Slope db/oct	Cut-off rad/sec
1	Heading	6	.25	6	.12
3	i.a.s.	7	.5	3	<.04
4	Altimeter	5	.5	5	.8
5	Pitch	12	1.0	7	.3
6	Roll	13	.95	6	.4
13	Rate-of-Climb	7	.15	12	.5

Phase II: Turn

Channel No.	Instrument	Memmolo		MacNamara		Fisher	
		Slope db/oct	Cut-off rad/sec	Slope db/oct	Cut-off rad/sec	Slope db/oct	Cut-off rad/sec
1	Heading	10	.95	7	.07	8	1.0
3	i.a.s.	6	.13	6	.2	5	.12
4	Altimeter	6	<.04	6	.3	6	<.04
5	Pitch	7	.5	9	.4	7	.6
6	Roll	12	.7	4	1.0	12	1.0
13	Rate-of-Climb	13	.25	7	.5	10	.3

TABLE 3 (Cont.)

		Phase III: Approach			
		MacNamara		Fisher	
Channel No.	Instrument	Slope db/oct	Cut-off rad/sec	Slope db/oct	Cut-off rad/sec
1	Heading	8	.5	12	.5
3	i.a.s.	3	.5	5	.95
4	Altimeter	6	<.04	6	.2
5	Pitch	11	.6	8	.7
6	Roll	12	.5	12	.7
13	Rate-of-Climb	6	.5	7	.5
14	Localizer	3	<.04	6	<.04
15	Glide Path	7	.5	6	<.04

The pilots were MacNamara, Fisher, and Memmolo. The phases of the flight are identified as: level flight and descent (Phase I), turn (Phase II), and final approach (Phase III). The reader may refer to Figure 3 which is a pictorial representation of the flight profile of the task. The various phases of flight mentioned above are identified in this figure. Each instrument was identified by a channel number (on which it was recorded on the 20-channel A-D conversion system) and by name (as it appears in the above table). The power of the signal which appears on that instrument, the slope in dB per octave, and the cut-off frequency in radians per second were taken from the spectral density photos.

6. MODEL VALIDATION STUDIES

Let us consider the application of the various models to the data obtained in the Link Simulator. It will be recalled that the various models may be hierarchically arranged in order of complexity. The simplest of these is the Senders model of 1955 which proposed that the sampling behavior of the subject will be in strict linear proportion to the effective bandwidths of the signals. A somewhat more sophisticated but still simple model proposes that this effective bandwidth which determines the sampling frequency will itself vary as a function of the pilot and the flight maneuver or condition in which he finds himself. Still further complexity is introduced by the queueing model of Carbonell which proposes the existence of a complex decision process. This process evaluates the relative uncertainty associated with each of the instruments, affects them by unitary costs, and selects for examination that instrument with the maximum cost if not looked at at that time. Various other models have also been proposed in previous reports.

6.1 Fixed Nyquist Model

Let us consider the problems which arise in considering the application of the data to the simple models. In particular, the signals differ so widely in their effective power that it may sometimes be the case that some signal will have a relatively high frequency content but a total power far less than that of some other signal. The computer, in calculating the power density spectrum of the signal, does not distinguish between various random noises and vibrations which may be present on any of the instruments, and real perturbations which can occur and which are of interest to the pilot. The pilot, however, can and does ignore the noise and for the most part concerns himself with those perturbations which reflect the actual behavior of the system which he is controlling and the inputs to that system from the outside. Other factors may also enter. The simple models make no assumptions about control behavior. In general they assume that the mean value on each signal is what (in this report) we call the desired time course of the various signals. They further assume that the random perturbations around these desired time courses have a mean value of zero and that the system will return to zero if left alone. No attempt has been made to introduce the effects of control behavior and the closed-loop transfer functions of the control system which the pilot is using into these models.

The simple model which was proposed in 1955 suggested that a precise quantitative relationship should exist between the frequency with which an instrument is observed and the bandwidths of the signals presented upon it. The model made no attempt to evaluate the relative importance of the signals or the costs of long observation, and it ignored the question of what the actual bandwidth of the signal might be considered to be. The early models assumed that a signal could be construed as possessing an infinitely sharp cut-off and that Shannon's sampling theorem would hold. Of course real systems do not have infinitely sharp cut-offs and as a result there exists the problem of determining at what frequency the signal power has dropped so far as to constitute a cut-off. For these simple models the determination is somewhat arbitrary and in the case of the simplest model we will use the frequency 3dB down from the break frequency obtained by fitting straight lines to the computer printouts of power spectral density.

The simplest model attempted to make absolute estimates. That is to say, it was predicted that if a signal had a bandwidth of 0.1 cps samples would be taken at the rate of one sample every five seconds (.2 samples/sec.). However, as pointed out earlier, we do not know what the effective loading of the pilot in this situation might have been and the model makes no attempt to estimate exactly what redistribution will be made or was made of the surplus attentive capacity of the pilots. In some of the laboratory studies (BBN Report 1246) it will be recalled that the surplus time seemed to be uniformly distributed over the remaining instruments rather than being distributed in proportion to their attentional demand a determined by actual measurement of the observing behavior of the operator.

For the purposes of testing this model our first requirement is the calculation of bandwidths of the signals presented on the various instruments. We will then convert these to proportions and examine the relationship between the proportion of the bandwidth of each signal to the sum of all the bandwidths and the proportion of total time spent observing that particular signal. Using the 3dB down frequency data obtained from Table 3 we can sum the bandwidths and re-express each of them as a proportion of the total. Then going to the eye-movement data we can, by multiplying the frequency and the durations of observation, calculate the total spent on each of the instruments in question and express these as proportions of the total available time spent on each of the instruments. There is some underlying assumptions which had best be made explicit; it is assumed that the duration of observation is relatively invariant within instruments; it is assumed that the differences between pilots can be expressed as alterations in the duration and frequency of observation in such a way to maintain a relatively constant total time spent on each instrument. Data supporting this notion have been obtained from the studies of pilot eye-movements at Wright-Patterson Air Force Base. The predictions to be made by these simple models are statistical in nature and do not claim that the sampling of each instrument will be periodic at the appropriate sampling frequency. Tables 4, 5, and 6 show the results for the simplest model for three pilots. As can be seen the results vary depending upon pilot, the particular run that he was making, and the phase of flight. The spectra were calculated for one run and predicted values obtained from this run compared with the actual values generated on each of the three runs. However, there is no indication that the run on which the spectra were obtained produces the highest correlation between the predicted and observed distribution of attention.

Let us consider the various pilots in sequence. Table 4 shows the data for pilot Memmolo in Phase I and Phase II. The correlations for both five or six instruments have been calculated for each of the three runs. For Phase I these range from .83 to .93. The suggestion is that the model is a good predictor of the actual behavior of the pilot. The greatest discrepancy appears in all runs in the large difference between the predicted observations of the heading indicator and the actual amount of time spent. For Phase II (descent and turn) the correlations are weaker, ranging from .64 for seven instruments to as high as .79 for five instruments. The major discrepancies occur in the rate-of-turn indicator in which the actual observations were very much below those predicted. The relationship for heading was closer than in the previous case although large discrepancies occurred here. The nature of the discrepancies is logical in that the instruments which are favored at the expense of others are in fact alternative ways of obtaining the necessary information for the execution of the turn maneuver.

Table 5 presents the data for Pilot MacNamara for Phases I, II, and III. Phase I correlations are very much lower, are all positive, and range from a low of .11 for six instruments to a high of .26 for five instruments. The greatest discrepancies are in

TABLE 4: PILOT MEMMOLO

PHASE I		PROPORTION OF ATTENTION							
		PREDICTED		OBSERVED					
				RUN 1		RUN 2		RUN 3	
Inst. Name	Bandwidth 3dB down in cps	6 inst.	5 inst.	6 inst.	5 inst.	6 inst.	5 inst.	6 inst.	5 inst.
Airspeed	.107	.135	.157	.121	.124	.124	.129	.085	.088
Compass	$5.57x10^{-2}$	.070	.082	.225	.230	.228	.237	.195	.200
Heading	.336	.461	.537	.469	.480	.425	.443	.530	.544
Altimeter	.121	.152	.178	.117	.120	.151	.157	.129	.132
Rate-of-Climb	$3.18x10^{-2}$	.040	.047	.044	.046	.033	.034	.035	.036
rpm.	.113	.142		.024		.040		.025	
	R=			.84	.898	.83	.887	.90	.93
PHASE II		7 inst.	5 inst.	7 inst.	5 inst.	7 inst.	5 inst.	7 inst.	5 inst.
Airspeed	$2.86x10^{-2}$	.038	.056	.081	.091	.070	.082	.055	.069
Compass	.186	.247	.336	.136	.154	.141	.165	.104	.129
Heading	.239	.317	.470	.550	.623	.447	.523	.465	.578
Altimeter	$8.99x10^{-3}$	.012	.018	.109	.124	.192	.225	.173	.215
Rate-of-Climb	$4.62x10^{-2}$	.062	.091	.002	.009	.004	.005	.007	.009
Rate of Turn	.236	.313		.105		.142		.190	
rpm.	$8.99x10^{-3}$	.012		.012		.003		.006	
	R=			.65	.79	.64	.71	.69	.69

TABLE 5: PILOT MAC NAMARA

PHASE I		PROPORTION OF ATTENTION							
		PREDICTED		OBSERVED					
				RUN 1		RUN 2		RUN 3	
Inst.Name	Bandwidth 3 dB Down in cps	6 Inst.	5 Inst.	6 Inst.	5 Inst.	6 Inst.	5 Inst.	6 Inst.	5 Inst.
Airspeed	1.27×10^{-2}	.022	.026	.130	.134	.134	.144	.130	.135
Compass	2.71×10^{-2}	.046	.056	.181	.188	.204	.219	.214	.222
Heading	15.44×10^{-2}	.263	.321	.487	.505	.404	.434	.459	.476
Altimeter	.193	.328	.401	.090	.093	.088	.095	.069	.072
Rate-of-Climb	9.39×10^{-2}	.160	.195	.078	.085	.099	.107	.918	.095
rpm	.107	.182		.034		.071		.036	
	R=			.21	.26	.13	.16	.11	.15
PHASE II									
		7 Inst.	5 Inst.	7 Inst.	5 Inst.	7 Inst.	5 Inst.	7 Inst.	5 Ints.
Airspeed	4.46×10^{-2}	.049	.077	.063	.069	.079	.084	.070	.074
Compass	1.50×10^{-2}	.016	.026	.098	.107	.122	.130	.106	.112
Heading	.347	.379	.598	.639	.696	.621	.662	.668	.707
Altimeter	6.68×10^{-2}	.073	.115	.076	.082	.076	.080	.068	.072
Rate-of-Climb	.107	.117	.184	.042	.046	.042	.045	.033	.034
Rate-of-Turn	.213	.233		.033		.023		.027	
rpm	.121	.132		.050		.038		.028	
	R=			.78	.95	.74	.93	.76	.94
PHASE III									
		8 Inst.	6 Inst.	8 Inst	6 Inst.	8 Inst.	6 Inst.	8 Inst.	6 Inst.
Airspeed	.159	.174	.223	.047	.485	.072	.075	.058	.059
Compass	.103	.113	.145	.170	.174	.166	.172	.149	.152
Heading	.209	.228	.293	.403	.412	.325	.338	.460	.470
Altimeter	8.99×10^{-3}	.001	.013	.060	.060	.036	.038	.046	.047
Rate-of-Climb	.113	.123	.158	.051	.053	.090	.093	.056	.057
Localizer + g.p.	.120	.131	.168	.247	.252	.274	.285	.211	.215
Rate-of-Turn	9.55×10^{-2}	.104		0		.008		.001	
rpm	.107	.117		.022		.030		.019	
	R=			.60	.62	.62	.65	.67	.69

TABLE 6: PILOT FISHER

PHASE II		PROPORTION OF ATTENTION							
		PREDICTED		OBSERVED					
				RUN 1		RUN 2		RUN 3	
Inst.Name	Bandwidth 3 dB Down in cps	7 Inst.	5 Inst.	7 Inst.	5 Inst.	7 Inst.	5 Inst.	7 Inst.	5 Inst.
Airspeed	2.86×10^{-2}	.025	.046	.116	.154	.124	.170	.115	.159
Compass	.207	.180	.333	.125	.166	.133	.182	.161	.223
Heading	.318	.277	.512	.333	.442	.302	.414	.295	.408
Altimeter	8.99×10^{-3}	.008	.015	.083	.110	.073	.100	.049	.067
Rate-of-Climb	5.89×10^{-2}	.051	.095	.965	.128	.098	.134	.104	.144
Rate-of-Turn	.294	.256		.207		.214		.242	
rpm	.232	.202		.040		.055		.033	
	R =			.63	.87	.67	.89	.69	.95
PHASE III									
		8 Inst.	6 Inst.	8 Inst.	6 Inst.	8 Inst.	6 Inst.	8 Inst.	6 Inst.
Airspeed	.299	.207	.356	.162	.171	.091	.095	.111	.116
Compass	9.39×10^{-2}	.065	.119	.134	.141	.141	.148	.173	.181
Heading	.277	.192	.330	.257	.271	.200	.209	.200	.209
Altimeter	4.46×10^{-2}	.031	.053	.044	.047	.036	.037	.042	.044
Rate-of-Climb	.107	.074	.127	.125	.132	.148	.154	.133	.139
Localizer + g.p.	1.80×10^{-2}	.013	.021	.227	.239	.342	.357	.299	.312
Rate-of-Turn	.224	.155		.025		.030		.025	
rpm	.379	.263		.027		.012		.018	
	R =			-.16	-.41	-.49	-.23	-.46	-.15

the very much lower attention paid to the altimeter compared to the predictions and the very much higher than predicted attention paid to air speed and to the gyro-compass as well. For this pilot certainly the utility of the model would not appear to be great unless further explanatory principles can be presented. In Phase II the correlations are much higher ranging from a low of .74 for seven instruments to a high of .95 for five instruments. The major sources of error are once again the heading indicator and the rate-of-turn indicator. In this case the pilot scans the rate-of-turn indicator much less than predicted, by a factor of about 7 and the heading indicator more often than predicted by about the same factor. Again these are alternative ways of getting similar and related information. In Phase III the correlations range from .60 for seven instruments to .69 for six instruments. The major discrepancies appear in the rate-of-turn indicator which apparently was virtually ignored and in the engine rpm indicator which was examined about one-fifth as often as predicted. The air speed indicator was undersampled, the attitude indicator over-sampled.

Table 6 shows the results for pilot Fisher, Phase II and Phase III. In Phase II the correlations range from .63 for seven instruments to .95 for five instruments. Attitude and APW contributed significantly to the discrepancy. Air speed was also very different from the predicted values. In Phase III, however, the results are very much at variance with the predictions of the model. The correlations become negative for virtually all comparisons. In particular, those involving eight instruments show negative correlations ranging from 0.16 to -.49. The large amount of attention paid to the heading indicator the rate-of-climb, and the localizer-and-glide-path and the small amounts paid to rate-of-turn and rpm, of course, make very good sense from the point of view of actual pilotage. It can only be concluded that the model does not effectively predict on the basis of the computed spectra what the pilot is going to do. To some extent the result is not too surprising. We have assumed in this elementary model that the pilot is engaged in monitoring behavior and that his attentional distribution will be a function of the nature of the signals that he monitors. What we have done in practice is to let control activity take place. Naturally the human operator tends to keep the error on each of the signals down to some acceptable boundary. Under these conditions we might find that in actual control behavior the error power will be far less than that which the pilot might consider to be a threshold. Since we have no way of knowing what the actual loading was we don't know to what extent the pilot exceeded his required accuracy limits in controlling the machine. In addition the pilot's awareness of the rate of divergence of the machine, based on prior experience, will lead him to consider not the actual signals presented but rather the signals that could have been presented. That is to say, he treats the signals that could have been presented. That is to say, he treats the signals which are available to him at any particular time as a sample from a broader distribution whose characteristics he knows and whose characteristics in fact guide his sampling behavior. These factors appear more strongly as components of the model in the case of the queueing model (See Section 6.3).

It is interesting also to speculate on the reasons for the high degree of correspondence for pilot Memmolo and the low degree for pilot Fisher. Whether these can in fact be attributed to the difference in current flying time is not clear. Other experiments and analyses will have to be done in order to determine this. However, the extraordinarily high values of correlation shown by Memmolo particularly for Phase I suggest that the model has great predictive power for some pilots at some times and that this predictive power may depend on the experience level of the pilot and the extent to which they are minimizing the uncertainty functions which are implicit in the model. This may hinge upon instructions or attitudes or experience. The answer to this question must also be reserved for further research.

6.2 Variable Nyquist Model

Certain difficulties arise in attempting to test the model against the available data. The simple sampling model and the conditional sampling model both depend on our ability to estimate two parameters of the signal: the bandwidth and the required accuracy of reading. In the former case the two interact over an appreciable period of time. That is, the ratio of signal power to permissible error power for any signal will determine the effective bandwidth in that as the permissible error goes down, the bandwidth of the information bearing portion of the signal goes up. The permissible error in reading of a signal must be considered as a variable in the short term, which is dependent upon the observed value of the signal on the previous fixation. Thus, for the former case, we can use information gathered from the human monitor by inquiry and questionnaire to fix the tolerable error. For the latter we depend on a measurable relationship between the power of the signal and its frequency.

Let us now see what we have in fact already tested and what more we can test. The signals recorded for analysis were specified to be the major flight parameters like attitude, altitude, and so on. These signals by their very nature are the permissible error functions themselves (if we assume that the pilot was not overloaded at any time). Thus the bandwidths calculated from the corrected signals more or less correctly represent the effective bandwidths and are an effective test of the "periodic sampler model."

The test of the variable Nyquist model can be accomplished in two ways and it is not clear that either one falls within the scope of the project as presently constituted. One of these depends on analysis of stored digital data which reveal the actual value of the various signals which were observed by the pilot on each fixation: given these various predictions can be made. For example, there should be a strong negative correlation between the observed deviation of a parameter from its mean and the subsequent interval of nonobservation. The exact value of the interval depends, of course, on the way in which the power of the signal falls off with frequency. An alternative approach would be based on an analytic solution, that is, if one could assume a gaussian signal and some arbitrary power spectrum one could in theory compute an expected distribution of intervals between observations. From these distributions of intervals one could then easily compute the proportional distribution of attention across instruments, based on some simple assumptions about the way in which the duration of observations is related to the observed deviation of the signal.

The former analysis, which predicts correlations within the fine structure, requires computer data processing which was not contemplated at the outset and must await further analysis of the cumulated data base. The latter problem decomposes into two parts. One involves the computation of the expected distributions of observed values; the other calculations of intervals. On first approaching the former problem one is lulled into thinking that observations of a signal with a gaussian amplitude density distribution will also be distributed in gaussian fashion. However, careful consideration reveals that the distribution of observed values will deviate markedly from the density distribution of signal values. If signals were in fact independent (i.e., were taken at the *largest* appropriate Nyquist interval) the two distributions would in fact be identical. Since the interval varies between zero and the maximum interval the observations are dependent. Further, the dependency is such as to increase the probability of observing larger deviations. Thus, if a signal is observed at some large value Y(0), the next observation will be close in time and the correlation function between Y(0) and Y(1) greater than zero.

Let the logarithm of the low-frequency (below cut-off) power density of the uncontrolled (and unobservable) signal be A. Let E be the corresponding value for the controlled (observed) signal. The shift in cut-off frequency produced by the observation-control loop is then given by:

$$\log_2 f = \frac{1}{S}(A-E) \tag{1}$$

where S is the slope of the spectrum beyond cut-off in dB/oct.

If the cut-off of the uncontrolled signal is f and that of the controlled one is f_n, we have:

$$f_n = f \cdot 2^{1/S(A-E)} \tag{2}$$

Let Y_L be the applicable threshold; let Y_0 be the last observed value; and let Z_L and Z_0 be the corresponding normalized values. Then:

$$\Delta Z = Z_L - Z_0 = \frac{1}{\sigma}(Y_L - Y_0). \tag{3}$$

It is reasonable to relate E to ΔZ. Let us make the tentative assumption that

$$E = K \cdot \Delta Z \ , \ 0<K<1 \tag{4}$$

Then replacing (4) into (2) we obtain

$$f_n = f \cdot 2^{A/S - K \cdot \Delta Z/S} \tag{5}$$

$$\therefore \ E(\Delta t) = E\left(\frac{1}{f_n}\right) = E\left[\frac{2^{K \cdot \Delta Z/S}}{f \cdot 2^{A/S}}\right] \tag{6}$$

We also have a relationship between the conditional expected value of ΔZ, and Δt, under the assumption of a first-order control. It is given through the autocorrelation.

$$E(\Delta Z|\Delta t) = \Delta Z_0 \cdot R(\Delta t) = \Delta Z_0 e^{-\alpha\Delta t} \tag{7}$$

We would like to obtain a distribution for Δt. Analytically, this task does not seem to be a trivial one because of two reasons:
(a) the dependence of Δt on unobservable values f and A;
(b) more importantly we have the mutual dependence between ΔZ and Δt. Hopefully, careful work with equations (6) and (7) should yield the desired result, but this is beyond the scope of the present work.

6.2.1 Discussion

It can be seen upon reflection that for controlled loops the power of the signal at any frequency will diminish as a function of the amount of control which is exercised, that is, as a function of the amount of attention which is paid. Thus we might find that for a man exercising very tight control, the greater the high frequency power the less often was the signal attended to. In fact, this did occur for one of the subjects (see Table 6, Phase III). Presumably pilot Fisher operating in Phase III, final approach, was paying more attention and in consequence controlling more tightly than were the other two pilots. We find that during cruise (steady state maneuvers) the correlations are high and positive; whereas for pilot Fisher during the final approach the correlations are negative. For pilot MacNamara the correlations are always positive. We can infer that he controlled less closely. The data for pilots MacNamara and Fisher show that the signal power for MacNamara for the localizer and glide-path was .120 and for Fisher was .018.

The predicted values of attention for pilot Fisher are very much higher on the localizer and glide-path whereas the predicted values for the altimeter are of approximately the right size. For pilot MacNamara the relationship is approximately the reverse.

We consider then that the displayed signals in a tightly controlled situation represent the residual error below which the pilot either will not or cannot go. And we can only infer from the amount of attention required to reduce the error to this level what the signal power might have been had not control been exercised. We would expect that, in general, relatively relaxed pilots will conform more closely to the sampling model discussed in section 6.1 and that relatively hard-working pilots will conform less well. The variable Nyquist model would predict the more closely a parameter is controlled (and the less the power displayed upon the instrument), the greater the amount of attention devoted to that signal. For sampling behavior related to functions which are under direct control, other models must be used which take into account the closed loop transfer function of the man-machine combination.

The work of McRuer and Clement (32) is an example of this kind of application. These authors used man-machine systems analysis to make predictions of displayed signals and then applied a modified sampling model to predict the sampling behavior exhibited by controllers in multi-degree of freedom systems. It would appear from examining their results and ours that the simple sampling models are restricted in their application to monitoring processes only and that models which involve control are needed to handle sampling behavior of pilots engaged in control activity.

In essence the visual monitoring behavior of pilots engaged in closed loop control is predicated on "possibility" rather than on "actuality" and these possible displayed signals are computable on the basis of the man-machine system dynamics. Where control processes are under automatic control either one of two possible strategies for a monitor can be conceived of. One might be that he would monitor as closely as he would as if he were engaged in control process himself. The alternative is that he would treat the displays as displays to be monitored and would conform to the simple bandwidth or the variable bandwidth model. If the former were the case, automation would not result in a diminution of visual load although it would result in the diminution of motor load. We predict therefore that the behavior of a monitor of an automatic system would behave more like a monitor than like a controller in so far as his visual work is concerned.

6.3 An Economic Queueing Model of Visual Sampling: Detailed Discussion and Experimental Validation

6.3.1 Most Important Features of this Model

In Section 6.3 we discuss the theory and some validation studies of the queueing model of a pilot's visual sampling behavior in an instrumented flight recently proposed by Carbonell (22). It is pertinent here to start by summarizing the main assumptions on which the model is based because of their effect on several experimental details in the validation process. These assumptions are as follows:

(1) The instruments compete for the pilot's attention; each time he looks at one instrument he is postponing the observation of others, which form a queue.

(2) The queue discipline stems from an intelligent decision made by the pilot at each time. We assume that he tries to minimize the total risk involved in not observing the other instruments.

(3) This risk is given, for each instrument, by a unitary cost times the probability that the displayed value may, while not being observed, exceed a certain threshold that could lead to some catastrophic result.

(4) The pilot's task in visually sampling his instruments is part of a feedback loop closed through his control actions.

(5) If the pilot does not exert control, displayed values are not zero-mean Gaussian signals; the mean will be given by the last reading of the instrument, while the variance monotonically increases with time. This increase is due to the autocorrelation and also to a divergence term accounting for forgetting and fear of a sustained drift.

(6) If the pilot exerts control, he will be concerned not with the absolute reading of each instrument, but rather with the variations with respect to the reading he had expected to obtain at that time.

In the actual implementation of the simulation, we were forced to make, by necessity, some further assumptions which are not, however, an intrinsic part of the model. Some of them are:

(7) The pilot looks at each instrument a fixed amount of time, namely 0.4 seconds. Longer looks are accounted for as consecutive selections of the same instrument in 0.4-sec quanta of time.

(8) Control actions and autocorrelation functions of the signal are of the form exp(-kt).

(9) The divergence is accounted for by means of a linear term in t subtracted from the square of the autocorrelation.

As said above, assumptions (7) to (9) are not necessary components of our model. As a matter of fact, it is our feeling at the time of this writing, that more elaborated conditions should be used instead of those stated above. And indeed, condition (8) was partially modified in the validation experiments as the result of trial runs (see Section 6).

6.3.2 Detailed Formulation of the Model

A. General Discussion

Let us define:

M	Number of instruments
t	Observation time
$C(t)$	Total cost of not looking at any instrument at some instant of time
C_i	Cost associated with exceeding the established threshold for instrument i. It is supposed to be independent of time for a given mode of operation (normal horizontal flight, take-off, landing, etc.), but the observer changes it when he switches from one mode of operation to another.
$P_i(t)$	Probability that an instrument i will exceed its threshold L_i at instant t

Probabilities and cost will be related by

$$C(t) = \sum_{i=1}^{M} \frac{C_i P_i(t)}{1 - P_i(t)} \tag{8}$$

The nonlinear dependence on the P_i's forces the observer to eventually look at all instruments that reach a high probability, P_i, even if the unitary costs may be low.

The total cost C above is really the cost of not scanning, of not looking at any instrument. The total cost of looking at a particular instrument j at instant t will be:

$$C'_j(t) = C(t) - \frac{C_j P_j(t)}{1-P_j(t)} \tag{9}$$

The postulated strategy is to select the instrument j that makes $C'_j(t)$ a minimum, i.e., the subtrahend a maximum. Equation (9) is justified on the basis that by looking at instrument j, one removes the uncertainty with respect to it; but because of the finite time taken by each observation the uncertainty with respect to the other instruments may increase.

From observation $(t - \Delta t)$ to observation (t) there is only a change in the probability vector. All probabilities are modified at each observation; those from instruments not observed because of greater uncertainty; and that of the one just observed because we have obtained new information about it. The probabilities P_i are conditional probabilities

$$P_i(t) = P[y_i(t) \geq L_i | y_i(t - \Delta_i(t)) = YO_i(t)] \tag{10}$$

where

t = Observation time
$y_i(t)$ = Value indicated by instrument i
$\Delta_i(t)$ = Length of time units since last sampling of instrument i before time t

$YO_i(t)$ = Last reading of instrument i before time t

L_i = Threshold value (danger mark)

In the next sampling instant we will have

$$P_i(t + \Delta t) = P[y_i(t + \Delta t) \geq L_i | y_i(t + \Delta t - \Delta_i(t + \Delta t)) = YO_i(t + \Delta t)] \quad (11)$$

where

$$\Delta_i(t + \Delta t) = (1 - \delta_{ij}) \Delta_i(t) + \Delta t, \quad (12)$$

where j corresponds to the instrument sampled at instant t and δ_{ij} is the Kronecker delta function.

In order to apply and test the model described above, we must compute the vectors $\underline{C}$ and $\underline{P}(t)$ whose components are respectively the C_i's and P_i's. The cost vector may be approximated by making some weighted estimations of the risks involved in exceeding the thresholds under a certain mode of operation for the different actual instruments. On the other hand, matching the model with the actual behavior of the pilot may provide a means for obtaining the C_i's that provide the closest replica to his strategy.

Let us study now the probability vector $\underline{P}$. Let us deal with a "desired value" instrument like the altimeter in normal horizontal flight. Let us zero the instrument on the desired value and assume that there are symmetrical thresholds around that desired value ($\pm L_i$), though the model could also handle "one-sided" or non-symmetrical instruments. If it is of the type which can be called a "threshold instrument," we would only have to modify our knowledge of the present state of the instrument (a threshold value instrument does not give a reading until the limiting value is reached).

Let us assume (Fig. 8) that a value YO is read on instrument i at a given instant. According to the type and reading of the instrument, the pilot may, or may not, exert control to bring YO back to zero. If he does not, and in the absence of any self-stabilizing process (this is the case for high performance planes), YO will become the mean of the distribution of probable values in subsequent intervals. There is a distribution, which we assume Gaussian, because of random effects on the aircraft or aircraft system. So, in this case, with no control, the process is a sort of random walk, with equal probability of being above or below the value just read, by values that are Gaussianly distributed. But the whole process is not a zero-mean Gaussian process; only the conditional probability of the increments can be considered Gaussian, conditional both on last reading, and on time elapsed since.

If control is effected, we can assume, for simplicity, that it is of the form $\exp(-K_1 t)$. This control will affect the mean but not the variance of the distribution. But even when control actions are taken, the resultant long-term signal that will then be zero-mean will not be Gaussian since the control coefficient K_1 is predictable and quite independent of the random fluctuations due to the atmosphere, etc. For a zero-mean low-pass Gaussian signal the mean is zero at the next "Nyquist interval," no matter what YO was read. But, we have shown that the mean might be expressed as $\exp(-K_1 t)$ (or any other controlling function), and if K_1 is small, this mean will not be zero for t equal to the "Nyquist interval" of the random fluctuations.

In conclusion, the random fluctuations are a type of random walk which only control actions bring down to equilibrium. But, both effects are independent and though the overall result may look as

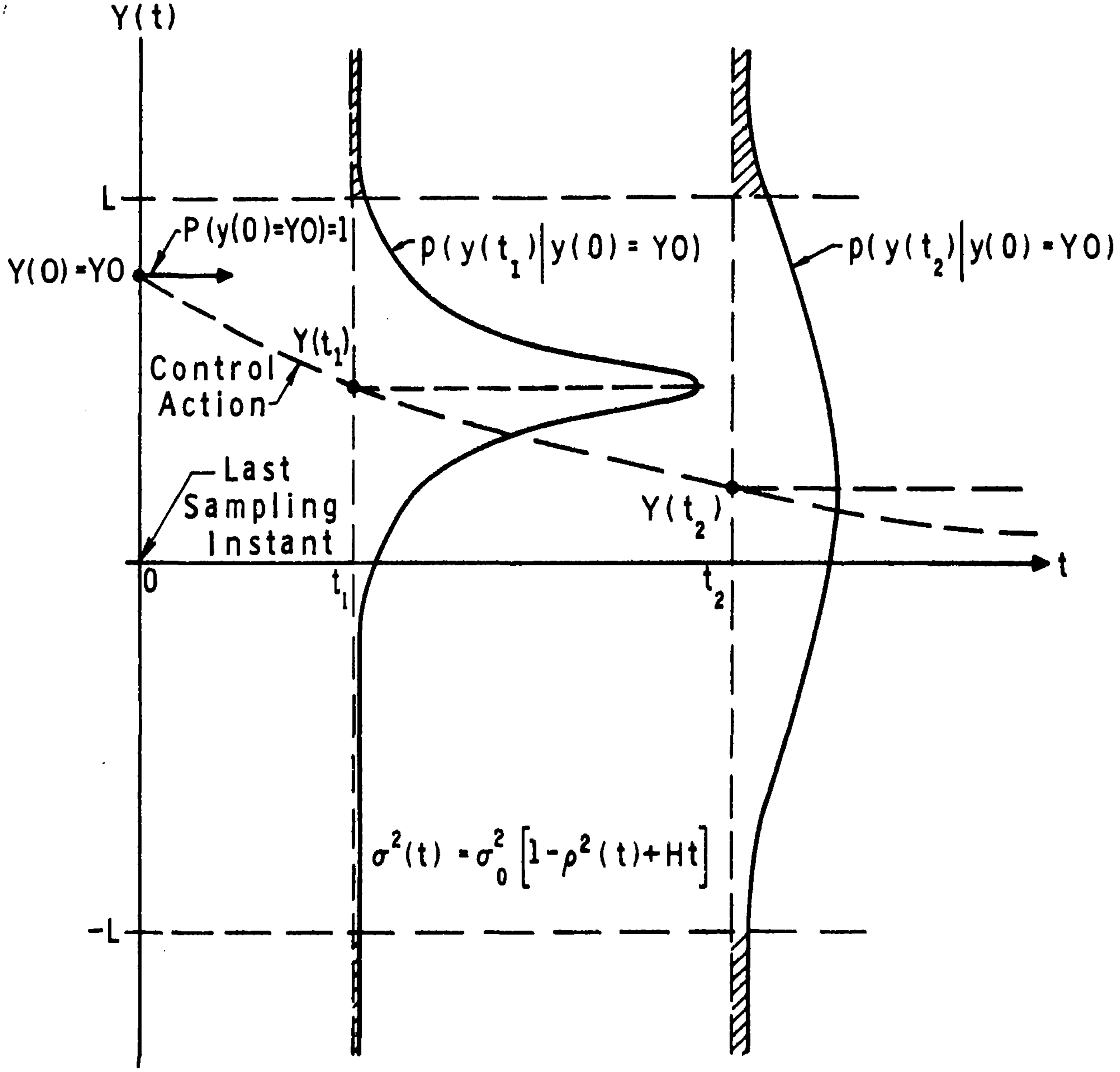


FIG. 8 RISK GROWTH WITH TIME FOR A GIVEN INSTRUMENT

if it were a zero-mean Gaussian signal, it is not. This is especially critical between observations when no new control action is initiated.

The above discussion can be summarized by Eq. (13)

$$P_i(y_i(t) = Y_i | y_i(t_o) = YO_i) = \frac{1}{\sqrt{2\pi[\sigma_i(t - t_o)]^2}}$$

$$\exp\left[-\frac{(y_i - YO_i \exp[-(t - t_o)K_i])^2}{2[\sigma_i(t - t_o)]^2}\right] \tag{13}$$

If we had a zero-mean Gaussian process, instead of $\exp[-(t-t_o)K_i]$, we would have had $\rho_i(t - t_o)$, ρ_i being the autocorrelation function of the random process.

Let us now consider the variance σ_i^2. This variance is obviously zero for $t = t_o$, and so the distribution is an impulse on YO for $t = t_o$ (see Fig. 8). As time t increases, the values of $y(t_o) = YO_i$ and $y_i(t) = Y_i$ become less and less correlated. Assuming the fluctuations to be Gaussian, the variance at t is given by:

$$\sigma^2(t) = \sigma_o^2[1 - \rho^2(t)] \tag{14}$$

where $\rho(t)$ is the normalized autocorrelation function and σ_o the sample variance of the fluctuation process. For a (low-pass simple RC) filtered noise, $\rho(t)$ is of the form $\exp(-K|t|)$.

With the present modified approach (as compared to some presented in Ref. 20), the (conditional on last reading) probability of exceeding the threshold in the absence of control never decreases with time until a new sample is taken; the variance σ^2 will, however, asymptotically approach the constant value σ_o^2 for large values of t. A better fit with experimental results may be obtained if we allow a divergence term to appear in Eq. (14). We can conveniently make, for instance:

$$\sigma^2(t) = \sigma_o^2[1 - \rho^2(t) + Ht] \tag{15}$$

where H is a (small) divergence factor. This divergence takes account of both forgetting on the part of the pilot and a sustained drift on the part of the aircraft or aircraft system. It makes the probability of exceeding the threshold tend to 1 if we wait for a sufficiently long period of time (see Fig. 8).

Finally, the probability of exceeding either one of two symmetrical thresholds (±L) is:

$$P_i(t) = 1 - G_i^{c+}(t) - G_i^{c-}(t) \tag{16}$$

where

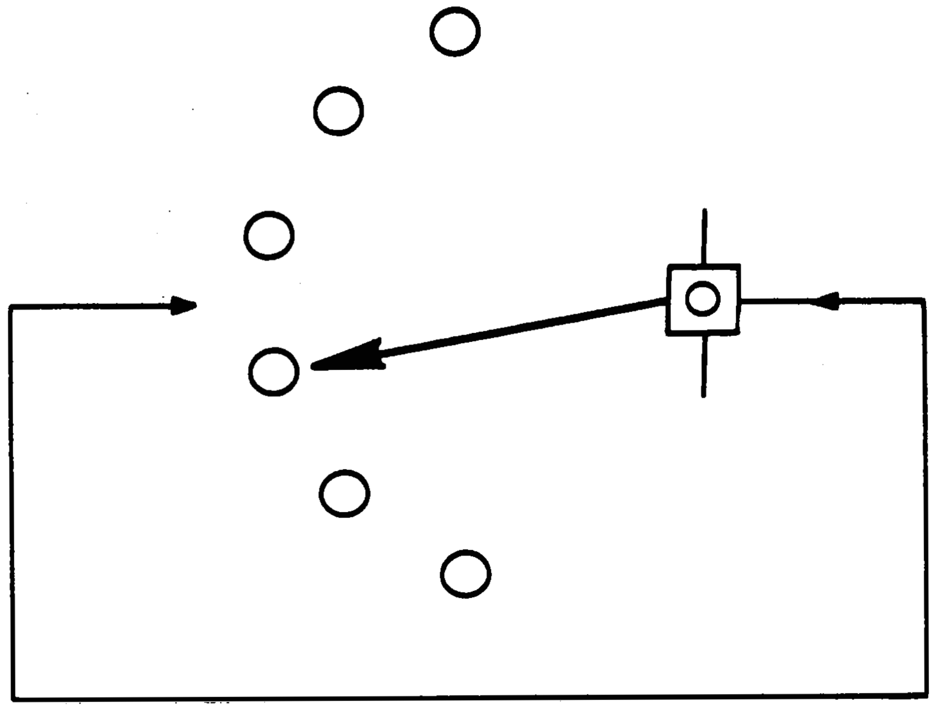
$$G_i^{c+}(t) = \int_L^{\infty} P_i(y(t)|y(t_o) = YO)\, dy \tag{17}$$

$$G_i^{c-}(t) = \int_{-\infty}^{-L} P_i(y(t)|y(t_o) = YO) dy \tag{18}$$

i.e., cumulative Gaussian distributions (complementary for G_i^{c+}).

B. Theoretical Approaches Through Queueing Theory

In Fig. 9 we have defined our problem as a queueing problem. As we see, it is a single server cyclic queue. Once an instrument

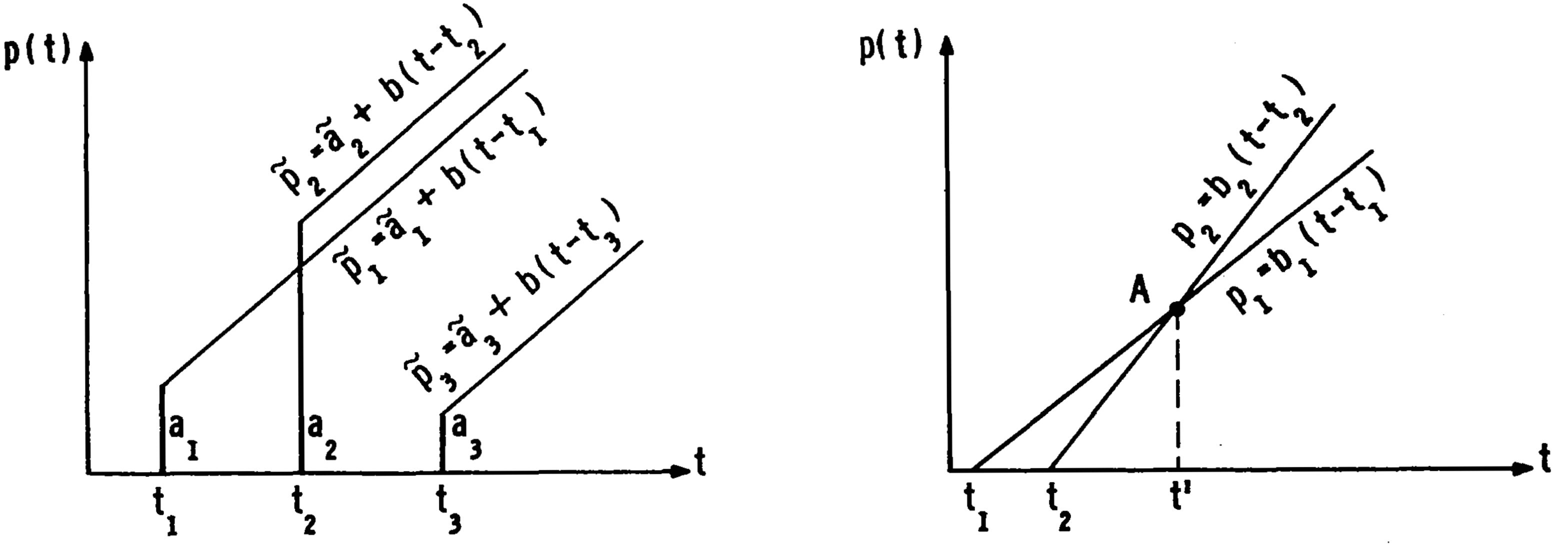


Type: Cyclic queue
Arrivals: Coincident with service completions
Service: General (First approximation: Regular)
No. of servers: 1
Queue limit: M
Queue discipline:
  Variable priorities depending on:
    a) Conditional random distribution(last reading)
    b) Waiting time
    c) Risk factor

FIG. 9 QUEUEING MODEL FOR INSTRUMENT SAMPLING has been served, it returns to the queue (we may consider it out if the risk probability is very low). So, the number of customers is M, and the queue length is M - 1. We assume regular service time and, therefore, regular arrivals. The delicate and difficult point appears in stating the queue discipline. We have a priority system of a very complicated nature since priorities should depend nonlinearly on last readings of the instruments, on time elapsed since last reading of each, and on a risk-cost hierarchy. The concept of priority queues has received wide attention in the past few years after the pioneering work of Cobham (26) on fixed priorities. Good reviews of the subject can be found in Cox and Smith (27) and Saaty (28). Jackson (29,30) was the first to study a queue with variable priorities with his introduction of the concept of dynamic priorities in 1960. Another type of variable priority has been described by Kleinrock (31). Fig. 10 presents a comparison of both types of time-dependent priority disciplines. Jackson's model assumes that each customer is assigned on arrival a random number $\tilde{a}_i$. His priority is then measured as a random variable $\tilde{p}_i(t)$ given by:

$$\tilde{p}_i(t) = \tilde{a}_i + b(t - t_i) \tag{19}$$

where b is a constant common to all customers (namely 1) and t, is the time of arrival. So, "a newcomer takes precedence over a unit in the queue if, and only if, the difference between the former's class number and that of the latter is not less than the time the latter has spent waiting (29)." Obviously, at any time t between arrivals we have a fixed-priority type of queue. Jackson gives, under certain hypotheses, bounds for expected waiting times.



a) Jackson's Dynamic Priorities

b) Kleinrock's Priorities

FIG. 10 TIME DEPENDENT PRIORITY MODELS

Kleinrock's model assumes a customer-dependent proportionality between priority and wait in queue:

$$p_i(t) = b_i(t - t_i) \tag{20}$$

where, as before, p_i is the priority of customer i, and t_i its time of arrival. Now, the b_i's are constants assigned to the different customers. In the case of Fig. 10, customer 1 has a higher priority for $t < t'$, while priority of customer 2 takes over for $t > t'$. We see that Kleinrock's model allows changes in the priority hierarchy between arrivals. A customer can "pass" another when both are in the middle of the queue.

In our application we could use a combination of Kleinrock and Jackson approaches. Jackson's random number assignment could be related to the last value read, and Kleinrock's constants, b_i, to the costs, the variances, and divergences. In a linearization of our problem, we could assume priorities given by

$$\tilde{p}_i(t) = \tilde{a}_i + b_i(t - t_o) \tag{21}$$

where $\tilde{a}_i$ is a random number and b_i is a constant particular to each instrument. An attempt to find a closed-form solution to this problem has two inherent drawbacks. First, it would be mathematically a very difficult, if not impossible, problem, since Jackson, in his reduced version, can give only bounds to some solutions. Secondly, it would very likely represent an oversimplification of the complex highly nonlinear decision process described in Section 6.3.2-A. At this stage in our study, it seemed then that the only practical line of attack was to simulate the problem on a digital computer.

C. Early Computer Simulations of the Model

The model was first simulated using the BBN Telcomp time-sharing system. The basic assumptions were essentially the same as stated in Section 6.3.1. Only double-sided symmetrical desired-value instruments were considered, but no difficulties are foreseen to including other kinds if so required. No coupling among instruments was included in the present version; coupling could be incorporated by correcting the expected as well as the read values. The observer is assumed to be a first-order observer, reading only positions of a pointer on an instrument. All time intervals are made equal to 1, though continuous time and variable observation durations could be easily incorporated. Finally, the decision process is assumed to proceed along the lines described in Section 6.3.2-A, as the result of evaluation of the risks involved if given thresholds are exceeded. Within the above assumptions, the simulation program has been made as versatile as possible to be able to easily modify parameters and conditions and even the program itself.

Figure 11 is a block diagram of the simulation. At initiation, the computer demands the value of the different constants. These are the number of instruments, M, threshold for not considering an instrument in queue, TH, and these values for each instrument i: standard deviations of the random fluctuations, autocorrelation decay constant K_i, threshold of risk L_i, threshold of accident L'_i, unitary cost, divergence constant, and control decay constant (which measures the convergence towards the desired value). If this last constant is zero, no control is exerted.

Next, we read initial values for all the instruments using our Gaussian random number generator subroutine. At this initial stage we take the mean equal to zero (we might also initially

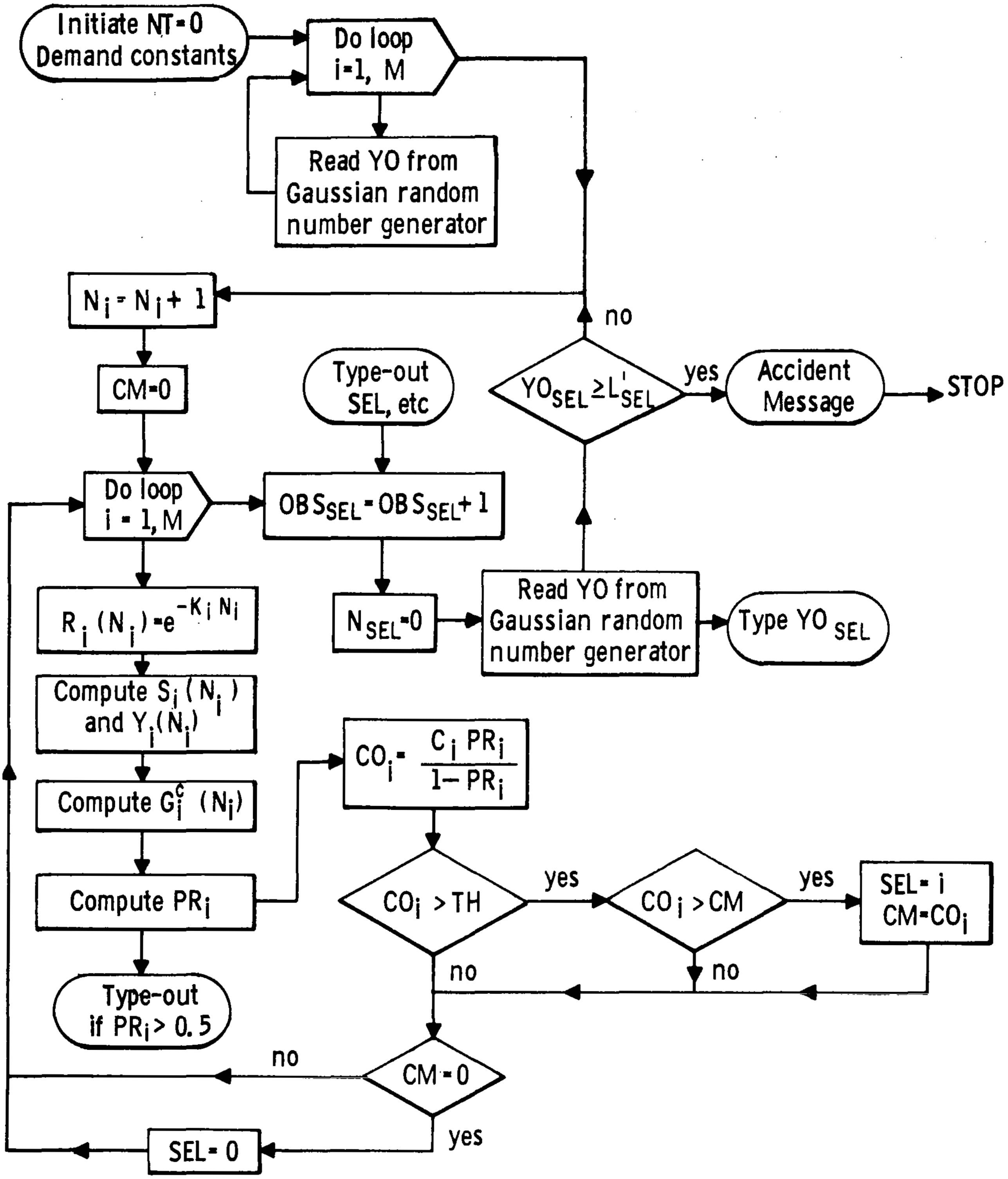


FIG. 11 BLOCK DIAGRAM OF SIMULATION

randomize the means). Then we let one time unit pass and enter the evaluation process on the different instruments. To this end, the autocorrelation for each instrument is computed as a function of time elapsed since last reading, and from it and the divergence the variance S_i^2 is found. From it and the expected mean (the value last read corrected by the control), we can find the probability is very large, a type-out ("instrument in red") is produced. That instrument may not be observed yet, however, if some other instrument has higher cost. Next, the cost CO_i is computed. If the cost is too low, the instrument is eliminated from the queue for the present interval; it will certainly enter the queue in some future interval because the divergence term forces the probability to monotonically increase toward 1. If the cost is higher than TH, we compare it with the cost from previous instruments. In this way, the instrument with the highest cost at the end of a given interval will be selected (SEL). When all the instruments have gone through the above process, SEL gives the instrument to be read (or a zero if all costs are below TH), and a type-out is produced. The value YO_{SEL} of the selected instrument is then "read" from our Gaussian random number generator, using Y_{SEL} as mean and $S^2{}_{SEL}$ as the variance. If $YO_{SEL} > L'_{SEL}$ the computer types an accident message and stops. If $YO_{SEL} < L'_{SEL}$ we recycle again by adding another time unit, and so on, in an infinite loop mode. This infinite loop can be interrupted any time the operator may desire (to stop the computation, change some constants, etc.). Statistics of observations are then produced.

By playing with the different constants in the simulation program, we have been able to place our simulated pilot in a very easy, a normally hard, or a jammed situation. The model shows, as was to be expected, good sensitivity to the parameters representing control actions. It is also sensitive to bandwidths (through the autocorrelation function) and to amplitudes (through the variance SO_1^2 and thresholds). Again, as expected, it is very sensitive to relative costs; increasing or decreasing their relative costs; increasing or decreasing their relative values produces important changes in the decisions.

Detailed results of this early simulation have been reported in the literature (22). These results indicated that if divergence constants were high, the model can simulate a quasi-periodic sampling behavior. On the other hand, by increasing the effect of past information (previous readings) versus that of the divergence, fully aperiodic sampling behavior could be generated, a case which has been observed in actual flights under conditions of stress.

6.3.3. The Experiments Modeling Actual Pilots

A. Objectives of this Validation

The simulations mentioned above (Section 6.3.2-C) could not test the model under realistic conditions. In order to do this, data recorded during the experiments on the Link trainer were used. As mentioned elsewhere in this report, we have digitized files stored in the computer of all instrument signals sampled at 0.1-sec intervals for three pilots under typical flight conditions, including descent, turn and landing approach.

The purpose of the simulation experiments to be discussed below was to make a statistical comparison of the normalized percentile number of fixations on a set of key instruments when they were obtained from actual eye-movement records versus those predicted by the model. The signals used for the simulation runs were based on the actual signals stored in the computer during Link trainer sessions, while the constants necessary to run the model were derived as precisely as possible from both signals and pilot questionnaires (see below). Each simulation produced 240 "looks" (i.e., 96 sec real time flight) aportioned among the various instruments.

A total of seven cases were thus analyzed, three for one pilot and two for each of the other two pilots. Three of the cases, one to each pilot, corresponded to what was called Phase II of the Link trainer flights during a 360-degree turning maneuver. Two cases involving two pilots corresponded to level flight and beginning of a descent maneuver (Phase I); while the two remaining cases, again involving two pilots, corresponded to a landing approach (Phase III). Instruments in Phases I and II were heading, air speed, altitude, pitch and roll, and rate of climb. For the simulation of Phase III, localizer and glide-path indicator were added.

B. Supporting Effort Necessary to Prepare the Simulation Runs

The constants for each instrument necessary to run the model are: cutoff frequency, mean square or variance of the corrected signal, threshold of risk in departure from a desired value, unitary cost, and divergence constant.

The first two sets of constants are derived from analyses of the signals described above. The mean-square value is obtained from the mean and mean-square values printed by the computer after the polygonal corrections for "desired" values are effected. In the first trial runs, the cutoff frequency was set equal to that derived graphically from the computer-produced power spectrum plot. At that time, no use was made of the measured slope beyond the cutoff value. By comparing the results of those early runs with the signal parameters, a significant correlation was found to exist between deviations in the percentage of observations between model and ocular data, and slope beyond cutoff. This pointed to our simplifying hypothesis assuming first-order control for all instruments. Instruments with slopes less than 6 dB/oct gave in those first runs more observations than they should, while those with slopes greater than 6 dB/oct (there are some with slopes up to 13 dB/oct) produced fewer observations. This was interpreted as meaning that the subjects payed attention not only to the cutoff frequency of the signals, but also to the general shape of the power spectrum. As a heuristic solution to this, it was assumed that the subjects behaved as if they extracted from each power spectrum the point 6 dB down from the cutoff, and then adopted that cutoff frequency obtained by the intersection a 6 dB/oct line from this point and the flat portion of the spectrum. Thus, a lower effective cutoff frequency for instruments with slopes greater than 6 dB/oct, and higher in the opposite case, is obtained.

Since the above discussed correction yielded satisfactory results, it was systematically incorporated in all runs. We are, of course, convinced that if more work with the model is to be done, the first-order hypothesis even with corrections should be replaced by control mechanisms of a more general nature. Because of the modularity of the model, that change would not affect other sections of it. It is interesting to note, incidentally, how the correction defined above shows the capacity of a good model to bootstrap itself by pointing to its own imperfections.

The sets of constants that we have called thresholds or risk, unitary costs, and divergences, were essentially derived from questionnaires (mentioned previously in this report) handed to the pilots after their Link trainer flights. It must be said that we were initially rather skeptical about the usefulness of their answers, but results of the simulation suggest that, with a few minor exceptions, the pilots' answers were truly representative of their strategies.

With respect to threshold for control, the pilots were asked to state for each instrument the minimum deviations they could perceive, the deviations they would like to stay within, and the emergency action deviations. In most runs we made some exploration by interpolating between these values; in any case, the set of values finally used represented a consistent way of selecting the thresholds regardless of the individual instrument.

The costs were selected by making use of two sets of costs, one numerical, the other graphical, declared by the pilots. They were not totally correlated, and though the numerical version was generally used, the graphical one proved better in a couple of instances. An interesting fact about the questionnaires was suggested by our experiments. One of the pilots (P.M.) probably after hard thinking, had changed his set of costs for Phase III by erasing and writing over. His final values produced quite distorted results when applied to the model. Fortunately, the original values could still be read, and when applied, results were better. In another instance, pilot D.M. had qualified an instrument (rate of climb) as not applicable during a certain phase of the flight and later changed his answer to give it a non-zero cost. It turned out that he practically never looked at that instrument, and a zero cost is the best approximation. (This is not absolute since another example of the same behavior corresponded to a substantial number of looks.) The above observations suggest that too much thinking by the subjects about highly complicated and mechanized behavior may impair the accuracy of their answers.

Finally, the divergence constants provided one of the most interesting and yet partially uncovered issues in the operation of the model. We started by asking the pilots how their concern would grow as a function of time if different individual instruments would become inoperative. They were required to draw separate curves for each instrument and each flight phase, with ranges between zero concern and total panic. Most of the curves were nonlinear with time and very few reached the "total panic" level, both perhaps because of the pilots' reluctance to declare so, and probably more importantly, because of their use of redundancy among instruments to replace an inoperative one with others. In any case, it was felt that the initial slope of the curves was related to their initial concern. A constant with the dimensions of time was derived by measuring how many seconds had passed before the tangent at the origin cut the midpoint between no panic and total panic. Then the divergence constant was the number that, after so many seconds, produced an N-fold increase in the variance of the signal, normalized with respect to the control threshold. The main uncertainty was originated by the value N, which forced considerable experimentation in trial runs by multiplying all divergence constants by different constant factors in order to vary that N. The proper way to account for the divergence effect still requires considerable experimentation.

C. How the Simulations Were Run

After the program was loaded into the PDP-1 computer, the program asked for the number of channels (instruments) to be used, and the corresponding constants (five constants per instrument). The channels were then identified by number and the corresponding signals transferred from magnetic tape to rapid access drum storage. Next polygonal corrections were inserted and everything was ready for the simulation run. At the beginning of the simulation, each corrected channel is read to initialize the system. Every 0.4 sec the decision process in the model selected an instrument to be read from the corrected signals. The normal output of the model was the number of observations of each channel after an overall total of 240 observations. Optionally, the number of observations can be typed out every twenty 0.4-sec samples. Another option not regularly used because of the time involved is the production of a type-out with the channel selected at each time, its cost, and the value read from the corresponding corrected signal. The computer also types when a threshold has been exceeded.

6.3.4 Presentation of Results

The general results of the simulation runs are presented in Tables 7 to 9 where both data and model results are normalized in percent. Though pitch and roll are separate instruments in the model, they have been lumped together since they appear together on eye-movement data. The same is true for localizer and glide-path indicator.

The general agreement between pairs of data and model columns in Tables 7 to 9 is surprisingly good. Table 7 refers to Phase II of the flight analyzed by the model, by pilot P.M. Table 8 presents results for the three pilots in Phase II; obviously Phase II by P.M. appears duplicated in Tables 7 and 8. Finally, Table 9 presents the results of the two simulation runs not covered in Tables 7 and 9.

Next, Tables 10 to 12 present statistical analyses of the results already presented in Tables 7 to 9. Two types of statistical figures of merit have been applied. The first is simply the average of the absolute differences between model results and data when these are expressed in percent. The second is the ratio of the sum of the squares of the differences between data and model (both in percent) to the sum of the squares of the observed data (in percent). This last measure is an indication of the relative importance of the deviations of model versus data when compared with data.

Table 10 refers specifically to averages across instruments, giving the above defined indices for each of the even simulation runs. Table 11 gives the indices, averaged across runs, for each individual instrument (or instrument pair); observe that localizer plus glide-path indicator is only pertinent for Phase III of the flight, and therefore appears in only two runs. Finally, Table 12 gives

TABLE 7

OBSERVED AND SIMULATED EYE FIXATIONS IN PERCENT FOR THREE PILOTS DURING PHASE II OF FLIGHT (TURN)

INSTRUMENT		PILOT					
		P.M.		J.F.		D.M.	
No.	Name	Data	Model	Data	Model	Data	Model
1	Heading	19.6	15.4	22.7	21.7	25.0	27.9
3	Air Speed	8.5	10.4	12.1	8.0	7.0	7.9
4	Altimeter	15.9	13.7	14.4	15.8	25.4	20.4
5 & 6	Pitch & Roll	46.0	45.9	36.4	39.1	41.9	43.8
13	Rate of Climb	10.0	14.6	14.4	15.4	0.7	0.0
	TOTAL	100.0	100.0	100.0	100.0	100.0	100.0

TABLE 8

OBSERVED AND SIMULATED EYE FIXATIONS IN PERCENT FOR A GIVEN PILOT (P.M.) DURING THREE DIFFERENT PHASES OF THE FLIGHT

INSTRUMENT		PHASE I Beginning of Descent		PHASE II Turn		PHASE III Landing Approach	
No.	Name	Data	Model	Data	Model	Data	Model
1	Heading	18.2	16.3	19.6	15.4	22.4	14.6
3	Air Speed	13.6	13.3	8.5	10.4	4.3	6.3
4	Altimeter	11.4	13.3	15.9	13.7	7.6	5.4
5 & 6	Pitch & Roll	41.5	41.3	46.0	45.9	31.9	32.9
13	Rate of Climb	15.3	15.8	10.0	14.6	8.6	14.2
14 & 15	Localizer & Glide Path	—	—	—	—	25.2	26.6
	TOTAL	100.0	100.0	100.0	100.0	100.0	100.0

TABLE 9

OBSERVED AND SIMULATED EYE FIXATIONS IN PERCENT FOR RUNS NOT INCLUDED IN TABLES 7 AND 8

INSTRUMENT		D.M. PHASE I Beginning of Descent		J.F. PHASE I Landing Approach	
No.	Name	Data	Model	Data	Model
1	Heading	25.3	22.5	20.5	19.0
3	Air Speed	16.9	11.7	10.5	12.0
4	Altimeter	16.9	22.1	6.5	8.2
5 & 6	Pitch & Roll	40.0	43.7	24.5	25.8
13	Rate of Climb	0.9	0.0	17.0	9.0
14 & 15	Localizer & Glide Path	—	—	21.0	26.0
	TOTAL	100.0	100.0	100.0	100.0

TABLE 10

STATISTICAL COMPARISON BETWEEN MODEL AND DATA, AVERAGED ACROSS INSTRUMENTS

CASE		Average of Absolute Differences (In Percent) Between Model and Data	Sum of Squares of Observed Data	Sum of Squares of Differences Between Model and Data	Ratio of Two Preceeding Columns
Pilot	Phase				
P.M.	I	0.96	2602.50	7.60	0.0029
	II	2.60	2925.22	47.26	0.0162
	III	3.33	2304.62	104.00	0.0451
J.F.	II	2.04	2401.38	28.06	0.0117
	III	3.17	1903.00	98.08	0.0516
D.M.	I	3.56	2812.12	76.42	0.0272
	II	2.28	3075.26	38.32	0.0125

TABLE 11

STATISTICAL COMPARISON BETWEEN MODEL AND DATA FOR INDIVIDUAL INSTRUMENTS AVERAGED ACROSS PILOTS AND PHASES

INSTRUMENT		No. of Runs Included	Average of Absolute Differences (In Percent) Between Model and Data	Sum of Squares of Observed Data	Sum of Squares of Differences Between Model and Data	Ratio of Two Preceeding Columns
No.	Name					
1	Heading	7	3.16	3417.79	101.59	0.0297
3	Air Speed	7	2.27	866.97	54.61	0.0630
4	Altimeter	7	2.80	1620.91	70.18	0.0433
5 & 6	Pitch & Roll	7	1.56	10136.68	27.33	0.0027
13	Rate of Climb	7	3.04	905.71	119.07	0.1310
14 – 16	Localizer & Glide Path	2	3.20	1076.04	26.96	0.0251

TABLE 12

STATISTICAL COMPARISON BETWEEN MODEL AND DATA, OVERALL RESULTS FOR PILOTS AND PHASES

NAME		No. of Runs Included	Average of Absolute Differences (In Percent) Between Model and Data	Sum of Squares of Observed Data	Sums of Squares of Differences Between Model and Data	Ratio of Two Preceeding Columns
PILOT	P.M.	3	2.36	7832.34	158.86	0.0203
	J.F.	2	2.65	4304.38	126.14	0.0293
	D.M.	2	2.92	5887.38	114.74	0.0195
PHASE	I—Beginning of Descent	2	2.26	5414.62	84.02	0.0155
	II—Turn	3	2.31	8401.86	113.64	0.0135
	III—Landing Approach	2	3.25	4207.62	202.08	0.0480

the indices globally, first for individual pilots, and next for flight phases—averaged in each case over everything else.

The differences between pilots are not really significant. Results for Phase III are poorer than those for the two other phases, though they are still quite acceptable. This probably reflects the presence of two new signals, as well as some coupling between instruments that may be higher than in the other phases.

With respect to accuracy in representation of individual instruments, rate of climb is the worse; incidentally, different pilots assign widely different costs to this instrument in relation to others. It is interesting to note that some of the instruments that have a high degree of information coupling are the three with less accurate predictions, namely rate of climb, air speed and altimeter, in that order. On the other hand, we have been able to quite accurately represent pitch and roll, an encouraging fact since it is the instrument that clearly was observed the highest number of times.

6.3.5 Conclusions and Suggestions for Further Work in Relation to This Model

The results discussed above prove that the basic assumptions involved in this model (Section 6.3.1) are essentially sound. The model has shown itself capable of accurately representing the behavior of pilots visually sampling their instruments during an instrumented flight. In our judgment, however, it would be premature to extrapolate our present success and attempt at this stage to use the model for designing purposes. We think that before that, some further research is needed to clarify some of the questions that our work with the model has raised. At this stage, we have fulfilled our original research goals which were to demonstrate the validity of this approach.

There are many interesting avenues open for improvements both from an applied and a basic research point of view. In the remainder of this Subsection we want to discuss some of them.

(a) The model has been validated in a statistical sense, based on percentage of observations of each instrument during a near 100-sec flight phase. It would be of great interest to study the behavior of the model at a more microscopic level by observing short trains of its decisions in comparison with actual short trains of eye-movements. It may be possible to automate this whole process by developing procedures (probably of a heuristic nature) for automatically decoding eye-movement waveforms stored together with instrument waveforms.

(b) At present, the model requires some tuning to particular pilots and particular phases of the flights. This is motivated by our need to adjust factors like the multiplicative constant of all the divergences and the method of interpolation between thresholds. It would be desirable to develop methods and criteria capable of avoiding that tuning process. The adjustment runs really represent our incomplete knowledge of what the constants in the model mean as a quantification of a highly complex human decision process. [Points (c) and (d), following, suggest an approach to this problem.]

(c) As mentioned before, the questionnaires are an essential input in order to run the model. Improved techniques for selecting and presenting the questions, as well as for interpreting the answers are needed.

(d) An alternative way to improve our handling of the constants, though not necessarily a substitute for the questionnaires, is to measure some of these constants by means of controlled experiments that would separately test particular aspects of the pilots' behavior. Of course, running the model is in itself an experiment (though a highly complex one) from which values of some constants could be empirically derived in order to be systematically applied later to similar cases.

(e) It is clear, of course, that some carefully planned and controlled experiments are also desirable not to directly measure the values of the constants, but rather to assess the nature of the psychological phenomena underlying them. Individual experiments can possibly be formulated to separately characterize phenomena like divergence, use of thresholds for deviations from a desired or expected signal, etc.

(f) The model can be improved in several of its subroutines. For example, we would like it to be able to handle the various slopes in the power spectrum without applying the heuristic correction described before in this Report. We would also like to modify the way in which the divergence constant affects the expected variance of the signal (we are not satisfied with the subtraction of Kt to the square of the autocorrelation).

(g) Other kinds of modifications could and should be made to the model to eliminate some of the restrictions and simplifying assumptions. For example, the duration of an observation could be made a function of the instrument being looked at, its value, and the overall situation. The simplifying restriction of having symmetrical thresholds could be lifted. The problem of information coupling and redundancy between different instruments could probably be modelled and incorporated into the general model. What happens when an instrument displays two or more signals (like pitch and roll) should also be investigated, as well as some allowance for taking into account peripheral vision. Finally, the configuration of instrument panel layout may have some influence that could be accounted for by means of a variable duration "blank" switching time.

The above items represent possible future lines of research. It may not be possible to consider all of them, which should lead to the establishment of a priority scale. In any case, we consider that those suggestions related to improvements in the model could be implemented under favorable conditions due to the modular nature of the model and the solution by simulation that we have adopted. Any analytical solution, in closed form, of a very complex mathematical problem, is usually highly affected by changes in minor components. A modular information processing model, such as ours, which yields answers through simulation, can, if its basic assumptions are sound, suffer considerable alterations and improvements without requiring drastic changes in the method of solution or in its other components.

6.4 A Comparison Between Two Models

Table 13 presents a comparison of correlation coefficients computed for the Nyquist and the queueing models. The queueing model as can be seen in the Table gives consistently higher correlations. This suggests that there is a trade-off between degree of correlation and complication of conceptualization which implies more complex procedures for obtaining results.

TABLE 13

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN DATA AND MODEL FOR NYQUIST MODEL (LEFT) VERSUS QUEUEING MODEL (RIGHT)

PILOT	PHASE		
	I	II	III
D.M.	0.905/0.966	0.730/0.974	———
P.M.	0.190/0.994	0.940/0.983	0.653/0.917
J.F.	———	0.903/0.984	-0.263/0.837

7. GENERAL CONCLUSIONS AND RECOMMENDATIONS

We have undertaken and completed a program involving extensive instrumentation of a simulator and of the pilots who operate it. The investigation has involved the recording and subsequent transformations of many varieties of data relating to both system and human performance. The basic goals were the testing of models and the examination of the ways in which human operators and system dynamics interact in a more or less deterministic way. The models which were tested range in sophistication from the very simple to the very complex and in their predictive efforts from the very statistical to the very deterministic. Only some of them could be tested within the framework of the present project and, indeed, as it turned out, some which we thought we could test seem to be basically untestable given the nature of the task presented to the pilot and the kinds of recording and analysis which we could undertake. Detailed discussion is limited to two different models of where people look and why they look there when flying an aircraft through a variety of more or less routine maneuvers. One of these models is basically a sampling theorem application, the other an extension of a cyclic queue with the addition of certain cost factors which guide the sampling behavior. In both cases, of course, the signal characteristics enter in strongly in determining the behavior of the model, so that we would expect both models to relate fairly well to the same body of experimentally obtained data. The results indicate that the statistical predictors are weaker than the queuing model which depends strongly upon individual pilot's responses about the cost of making certain control activities and the cost of exceeding certain limits. This we might expect. In general, the more detailed the examination of the basis of behavior, the more closely should the model so constructed fit the actuality.

The statistical model supposes that eye movements are purely a function of signal frequency characteristics and the desired accuracy of readout. However, due to the nature of instrument displays in which only error signals, i.e., deviations from desired setting, are displayed, it is very difficult to extract ratios of signal power to error power. Instead, these must be inferred. What we have, then, is a fitting of straight lines to the error spectrum and an inference from the break frequency of the error spectrum to the signal characteristics which governed the pilots' behavior. The results for the three pilots, as presented in Table 6, show that, in all but one case, the correlations between the actual distribution of attention and the observed data are positive and, with one exception, high. The defect of this model arises from the fact that, under certain flight conditions, in particular the final approach, the signal power for certain of the instruments becomes very small and our estimates of the bandwidth of the signal presented rather dubious at best. Thus our estimates of the attention paid to the localizer and glide path during final approach are greatly at variance with the observed behavior of the pilot as are the estimates for rate-of-turn and rpm. However, for Phases I and II of flight the correlations are generally high and in the expected direction and suggest that the model has power for predicting the distribution of attention when the pilot is not engaged in effectively continuous control of the signals that he is presented with.

The results of the queuing model presented in Section 6.3.3 are uniformly good and show that the application of this model to the distribution of attention results in variance reductions which in all but one case are greater than 90%. Overall, more than 95% of the variance is accounted for by the model behavior for all phases of flight and for all pilots.

Where it is possible to use data obtained from pilots to enter into a queuing/cost-effectiveness model, very powerful estimates can be obtained which would be useful for engineering decisions. Where such data cannot be obtained, as for example, in the analysis of systems which are new and for which no great experience exists, the statistical predictions of the sampling model also yield powerful predictions, although by no means as powerful as those yielded by the queuing model. In either case, the predictions of these models will be useful in making estimates of the loading imposed on the pilot by a well-defined system, and in calculating instrument panel configurations based on the way in which people must use the signals which are to be presented. An interesting application of the principles and theory which have been developed within this program is presented in Ref. 26. The results of that application appear to support the general notion that these analyses have immediate and direct engineering utility. Further work will explore the fine structure of visual attention. This will make it possible for the systems designer to make more reliable and rational decisions about problems of loading, function allocation, and cockpit layout.

Suggestions are also made for specific future research.

REFERENCES

1. R. E. Jones, J. L. Milton, and P. M. Fitts, "Eye fixations of aircraft pilots: A review of prior eye-movement studies and a description of a technique for recording the frequency, duration, and sequence of eye fixations during instrument flight." USAF Tech. Rept. 5837 (ATI 65996), September 1949.

2. J. L. Milton, R. E. Jones, and P. M. Fitts, "Eye fixations of aircraft pilots: Frequency, duration, and sequence of fixations when flying the USAF instrument low approach systems (ILAS)," USAF Tech. Rept. 5839 (ATI 65996), October 1949.

3. P. M. Fitts, R. E. Jones, and J. L. Milton, "Eye fixations of aircraft pilots: Frequency, duration, and sequence of fixations when flying Air Force ground controlled approach system (GCA)," AF Tech. Rept. 5967 (ATI 69038), February 1950.

4. R. E. Jones, J. L. Milton, and P. M. Fitts, "Eye fixations of aircraft pilots: Frequency, duration, and sequence of fixations during routine instrument flight," AF Tech. Rept. 5975 (ATI 73422), December 1949.

5. J. L. Milton, R. E. Jones, and P. M. Fitts, "Eye fixations of aircraft pilots: Frequency, duration, and sequence of fixations when flying selected maneuvers during instrument and visual flight conditions," AF Tech. Rept. 6018 (ATI 84010), August 1950.

6. J. L. Milton, B. B. McIntosh, and E. L. Cole, "Eye fixations of aircraft pilots: Fixations during day and night ILAS approaches using an experimental instrument panel arrangement," USAF Tech. Rept. 6570 (ATI 128350), October 1951.

7. J. L. Milton, B. B. McIntosh, and E. L. Cole, "Eye fixations on aircraft pilots: Fixations during day and night GCA approaches using an experimental instrument panel arrangement," USAF Tech. Rept. 6709 (ATI 158203), February 1952.

8. J. L. Milton and F. J. Wolfe, "Eye fixations of aircraft pilots: Fixations during zero reader approaches in a jet aircraft," WADC Tech. Rept. 52-17 (ATI 149050), February 1952.

9. J. L. Milton, B. B. McIntosh, and E. L. Cole, "Eye fixations of aircraft pilots: Routine maneuvers under day and night conditions, using an experimental panel arrangement," WADC Tech. Rept. 53220 (AD 35161), March 1954.

REFERENCES (Cont.)

10. J. W. Senders, "Tracking with Intermittently Illuminated Displays," WADC Tech. Rept. 55-378, October 1955.

11. J. W. Senders, "Tracking with Intermittently Illuminated Stimuli," WADC Tech. Rept., 1955.

12. J. W. Senders, "Man's Capacity to Use Information From Complex Displays," in Information Theory in Psychology, H. Quastler, Ed., The Free Press, Glencoe, Illinois, 1955.

13. J. W. Senders, "Information Input Rates to Human Users: Recent Research Results," WADC Symposium on Air Force Flight Instrumentation Program, Wright-Patterson Air Force Base, Ohio, 1958.

14. W. B. Knowles, "Operator Loading Tasks," Human Factors, pp. 155-161, April 1963.

15. A. I. Siegel and J. J. Wolfe, "A Technique for Evaluating Man-Machine Systems Designs," Human Factors, Vol. 3, pp. 18-27, 1961.

16. P. Ekstrom, "Analysis of Pilot Workloads in Flight Control Systems with Different Degrees of Automation," Proc. Inter. Congress on Human Factors in Electronics, IEEE, Long Beach, California, May 3-4, 1962.

17. H. Lindquist and R. Gross, "Human Engineering Man-Machine Study of a Weapon System," Minneapolis-Honeywell Aero Report R-ED 6094, Minneapolis-Honeywell, Minneapolis, Minnesota, October 1958.

18. J. W. Senders, "The Human Operator as a Monitor and Controller of Multi-Degree of Freedom Systems," IEEE Trans. on Human Factors in Electronics, HFE-S, No. 1, September 1964.

19. J. W. Senders, "A Re-Analysis of the Pilot Eye-Movement Data," Transactions on Human Factors in Electronics, IEEE, HFE-7, No. 2, June 1966.

20. J. W. Senders, J. I. Elkind, M. C. Grignetti, and R. Smallwood, "An Investigation of the Visual Sampling Behaviour of Human Observers," BBN Report No. 1246, May 1965.

REFERENCES (Cont.)

21. L. J. Fogel, "A Note on the Sampling Theorem," Transactions of the Professional Group on Information Theory, IRE, December 1956.

22. J. Carbonell, "A Queueing Model of Many Instrument Visual Sampling," Transactions on Human Factors in Electronics, Vol. AFE 7, No. 4, December 1966.

23. R. Smallwood, "Internal Models and the Human Instrument Monitor," Symposium on Human Factors in Electronics, IEEE, Minneapolis, May 1966.

24. E. C. Kris, "A Technique for Electrically Recording Eye Position," WADC Tech. Rpt. 58-660, Wright-Patterson Air Force Base, December 1958.

25. M. C. Grignetti, R. V. Payne, and J. I. Elkind, "Signal Analyzer I. A Signal Analysis System for Manual Control Studies," IEEE Int. Convention Record 13: Part 6, pp. 143-151, 1965.

26. A. Cobhan, "Priority Assignment in Waiting Line Problems," J. Operations Research Soc. Am., Vol. 2, pp. 70-76, 1964.

27. D.R. Cox and W.L. Smith, Queues. London: Methuen and Co., 1961.

28. T.L. Saaty, Elements of Queueing Theory. New York: McGraw-Hill, 1961.

29. J.R. Jackson, "Some Problems in Queueing with Dynamic Properties," Naval Research Log. Quart., Vol. 7, pp. 235-249, March, 1960.

30. _____,"Queues with dynamic Priority Discipline," Management Science, Vol. 8, pp. 18-34, Oct. 1965.

31. L. Kleinrock, "A Delay Dependent Queue Discipline," Naval Research Log. Quart., Vol. 11, pp. 329-342, Dec. 1964.

32. D. McRuer, H.R. Jex, W.F. Clement, and D. Graham, "Development of a Systems Analysis Theory of Manual Control Displays," Tech. Rpt. No. 163-1, Systems Technology, Inc., Oct. 1967.

NASA-Langley, 1968 —— 5 CR-1258