NASA CONTRACTOR REPORT

NASA CR-1746

STUDIES OF MULTIVARIABLE MANUAL CONTROL SYSTEMS: A MODEL FOR TASK INTERFERENCE

by William H. Levison, Jerome I. Elkind, and Jane L. Ward

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

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1971

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-1746		

4. Title and Subtitle	5. Report Date
STUDIES OF MULTIVARIABLE MANUAL CONTROL SYSTEMS: A MODEL FOR TASK INTERFERENCE	May 1971
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
William H. Levison, Jerome I. Elkind and Jane L. Ward	
	10. Work Unit No.
9. Performing Organization Name and Address	11. Contract or Grant No.
Bolt, Beranek and Newman Cambridge, Massachusetts	NAS 2-3080
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Contractor Report
National Aeronautics and Space Administration Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The results of a study of multivariable manual control systems are presented. The objectives of this study were to investigate human performance in multivariable control and monitoring situations, to develop models for the controller in these situations, and to develop a metric for pilot workload.

An extensive experimental study of multi-axis tracking performance is described. Data are presented for: (a) four-axis tracking tasks with visual scanning permitted among the four separate displays, (b) multi-axis tracking with fixation maintained on a single display, and (c) a complimentary set of single-axis tracking tasks with foveal and peripheral viewing of the display. Multi-axis behavior is shown to be consistent with single-axis results.

A model is presented for interference among multiple control tasks. This model is based upon the assumption that multiple tasks are performed in parallel and that the human must share a fixed amount of central-processing capacity among the tasks. Validation is provided by a comparison of model predictions with experimental results obtained from multi-axis control situations and from a single-axis, multivariable control situation. A metric for pilot workload is suggested which is based upon the model for task interference.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Manual control multivariable control systems task interference workload peripheral vision	Unclassified-Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	229	$3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

CONTENTS

CONTENTS (Cont.)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF TABLES

LIST OF TABLES (Cont.)

LIST OF TABLES (Cont.)

1. INTRODUCTION

The pilot's task in the operation of flight vehicles is a complex of control and decision tasks. Models for predicting pilot performance and limitations must take into account the effects of interaction among several such tasks. For the past few years we have been conducting a series of experimental and theoretical studies of these interactive effects as part of a continuing program to develop models for human performance in realistic flight situations. This report describes our most recent work on this problem which has been concerned with the interaction or interference among several different control tasks.

INITIAL TWO-AXIS STUDIES

We began this work with studies of simple two-axis control systems without coupling between axes. Integrated and separate displays, several different controlled element dynamics, and foveal and peripheral viewing of the displays were examined in these studies. The theoretical basis for this work was provided by the describing function models for the human operator available at that time (Refs. 1,2). These had been developed by several investigators over a ten-year period from studies of single axis control systems. Our approach was to attempt to elaborate these models so that they would include the effects of interference between tasks, visual scanning of displays, and the effects of viewing conditions (foveal vs. peripheral). The results of these studies have been reported and documented (Refs. 3,4) and are summarized below.

In the experiments with an integrated display, one in which the horizontal and vertical displacement of a single dot indicated the errors in the two axes, we found that the human controller transmitted almost twice as much information in certain two-axis tasks – those in which the dynamics on the two axes are the same – as he did in a single-axis task. We did not find evidence in the describing function results of a switching mechanism such as would be expected if the controller were single-channel and switched his attention from one task to the other. This result suggested that the two channels were processed in parallel.

To determine the extent to which these results were a consequence of the integrated display, a series of two-variable manual tracking experiments was next conducted in which subjects were required to view two separated displays and operate two control devices to control the system. Performance was measured as a function of the display separation, the forcing function bandwidth, the task difficulty, and the controlled-element dynamics. Human controller describing functions, eye movement distributions, and normalized mean-squared tracking error (NMSE) scores were obtained. Measurements were obtained when a single variable was viewed foveally, when a single variable was viewed peripherally, and when both variables were controlled simultaneously.

A simple model was developed which accounted for most of the results of this experiment. The key assumption of this model was that the human controller acts as a two-channel processor of information: one channel processes information obtained foveally while the other simultaneously processes information obtained peripherally. There was assumed to be no coupling, or interference, between channels. A switching mechanism was incorporated in the model to account for the switching of visual attention from one display to the other.

Although this model was able to reproduce with reasonable accuracy the effects of visual scanning upon system performance, it had at least two serious deficiencies. It lacked a scheme for predicting scanning behavior; system performance could be "predicted" only for a pre-specified scan pattern. There was no way to account for the performance degradation that was measured for certain cases in which the two displays were very close to each other so that they both were viewed foveally. It also became apparent that the describing function approach lacked sufficient generality to be the basis for an integrated model of complex control and decision tasks. We turned to modern control theory for a framework within which such tasks might be treated in a more straightforward manner.

THEORETICAL FRAMEWORK

Initial Optimal Control Theoretic Model

An optimal control model for human control, monitoring, and scanning behavior was developed in a separate but related research program (Refs. 5,6). Preliminary analysis of this model indicated that it had sufficient scope to account for human behavior in a wide variety of simple and complex tasks. We have adopted this model as the starting point for the theoretical and experimental investigations of task interference that are described in this report. The principal features of the model are discussed below.

A block diagram of the model is given in Fig. 1. The vehicle dynamics are assumed to be represented adequately by the linearized state equation

$$\underline{\dot{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{u}(t) + \underline{w}(t) \qquad (1)$$

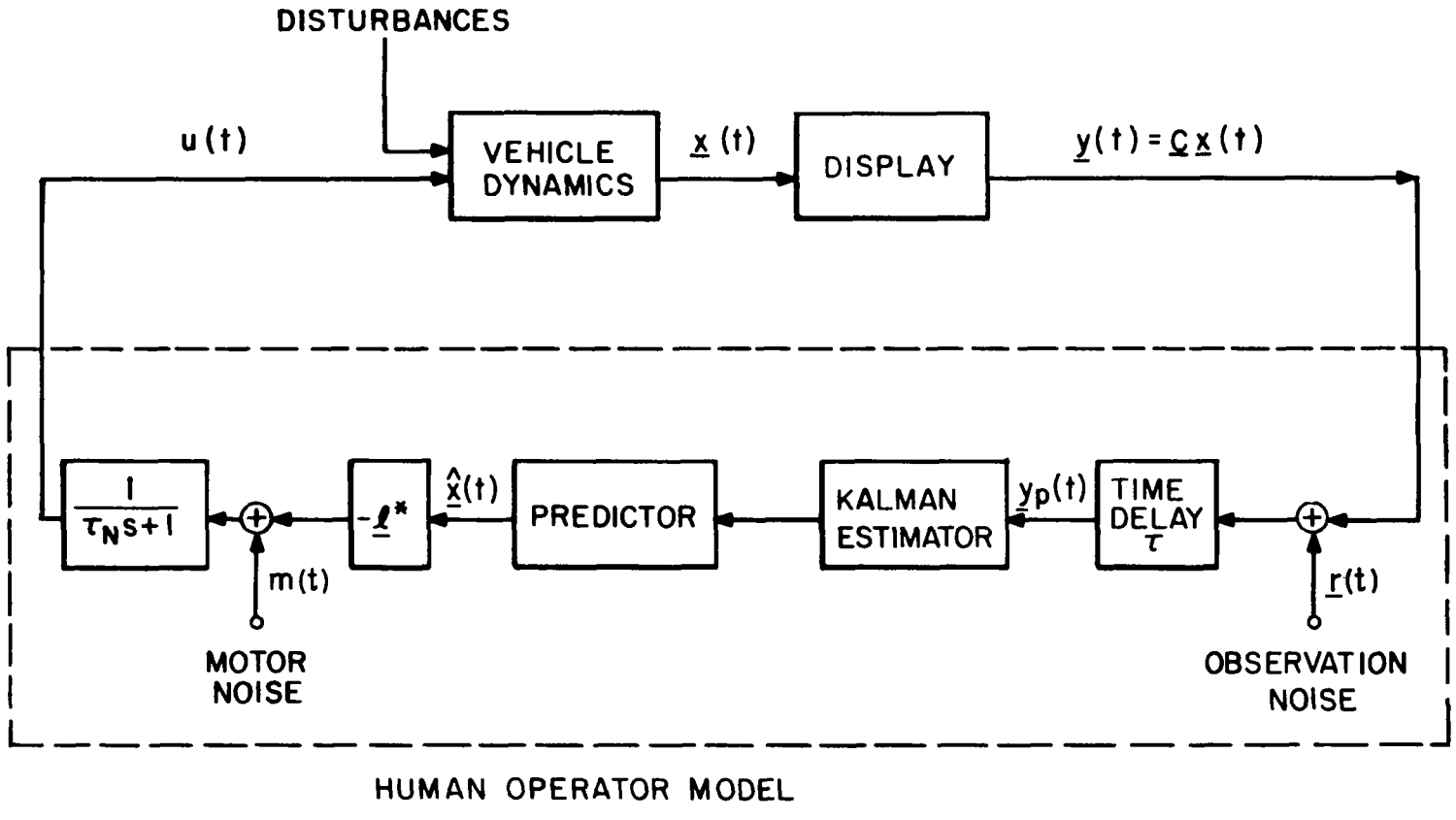


FIG.1 OPTIMAL CONTROL MODEL OF HUMAN BEHAVIOR

where $\underline{x}(t)$ is the vector which describes the state of the vehicle, $u(t)$ a scalar control, and $\underline{w}(t)$ a vector of white driving noise processes.* The pilot does not, in general, observe all of the state variables, but is provided with pertinent system outputs $\underline{y}(t)$, given as

$$\underline{y}(t) = \underline{C}\ \underline{x}(t) \tag{2}$$

The various psycho-physical limitations inherent in the human are represented by a lumped equivalent perceptual time delay τ and a model for remnant (discussed below) consisting primarily of an equivalent observation noise vector $\underline{r}(t)$. A separate (white) observation noise** is associated with each displayed output $y_i(t)$.

It is assumed that the overall control task is adequately reflected in the human's choice of a control input $u^*(t)$ which minimizes a weighted sum of averaged state and control variances

$$J(u) = \sum_{i=1}^{n} q_i \sigma_{x_i}^2 + r\ \sigma_u^2 + g\ \sigma_{\dot{u}}^2 \tag{3}$$

conditioned on $\underline{y}_p(t) = \underline{y}(t-\tau) + \underline{r}(t-\tau)$, the human's delayed, noisy observation of $\underline{y}(t)$.

---

* If the external forcing functions are rational noise spectra of first order or higher, the resulting "input states" are incorporated in the state vector $\underline{x}(t)$.

** In order to include the effects of visual scanning, statistics of the noises can be treated as time-varying. See Ref. 5.

Note that neuro-motor dynamics have not been included among the human's inherent limitations. However, included in J(u) is a cost which depends on control rate. This term can represent an actual cost on $\dot{u}(t)$ or it can be used to account indirectly for the physiological limitations on the rate at which a pilot effects control action. It can be shown that the inclusion of such a term results in a lag (often associated with the neuro-muscular system) being generated in the optimal controller.

In order to model certain control situations adequately, we have found it necessary to include a motor noise term m(t) in addition to the observation noise vector. This is helpful primarily when the input disturbance is not applied in parallel with the pilot's control signal. In this situation, motor noise serves to prevent the model from acquiring perfect knowledge of various system inputs or outputs which, in fact, are not known perfectly by the human. Use of motor noise here is strictly a mathematical convenience and does not imply that we are able to distinguish experimentally the various sources of remnant, which we cannot do at present.

With the above assumptions, the human's control characteristics are determined by the solution of an optimal regulator problem with time-delay and observation noise. We have shown (Ref. 5) that the resultant optimal closed-loop system has the structure shown in Fig. 1. The pilot model consists of the cascade combination of a Kalman estimator, a least-mean-squared predictor and a set of gains acting on $\underline{\hat{x}}(t)$, the best estimate of the system state $\underline{x}(t)$. The lag factor τ_N depends on the choice of control rate weighting, g. The features that are unique to this model are its representation of human limitations and the resulting optimally compensating elements.

Since the optimal feedback controller is linear, the pilot can be represented in the frequency domain by a transfer function,

$$u(s) = \underline{H}(s)\underline{y}(s) \tag{4}$$

Therefore, in a straightforward manner, one can predict human operator describing functions which are equivalent to those which could be measured in an experiment. Furthermore, the model allows us to predict the power spectrum (input and remnant related) of any system state, of any output, or of the human's control. Also available is a prediction of closed-loop performance in terms of the cost functional [Eq.(3)].

An Observation Noise Model for Controller Remnant

Although early studies of the human controller tended to ignore the effects of controller remnant on system performance, remnant has begun to command an increasing amount of concern. Recent attempts have been made by others to relate system performance to visual scanning behavior by appropriate treatment of the remnant process (Ref. 7), and we have found that an equivalent observation noise representation of remnant allows us to incorporate a model for task interference into our optimal-control framework. Because an understanding of remnant is crucial to the theoretical development pursued in this report, the pre-experimental model for controller remnant is reviewed below. This model was developed in a separate but related theoretical study and is discussed in detail in Refs. 8 and 9.

A number of multiplicative, or proportional, sources of human randomness were originally considered. Most of these processes were found to be indistinguishable in their effects on controller remnant and were therefore combined into a single noise process – an equivalent observation noise process.

Equivalent observation noise has been found to be a vector white noise process; i.e., each sensory input variable that the controller wishes to estimate is disturbed by white noise. The component noise processes are assumed to be linearly independent of each other and of the system forcing functions. When the display is viewed foveally, the noise processes appear to scale with signal variances. This scale factor has been found to be independent of the amplitude and spectral shape of the displayed variable, independent of vehicle dynamics, and does not appear to depend on the quantity (i.e., position or velocity) that the controller obtains from the display indicator. Foveal observation noise is thus given as

$$\underline{R} = P \cdot \underline{\sigma}_y^2 \tag{5}$$

where $\underline{R}$ is the power density level of the injected white noise vector, $\underline{\sigma}_y^2$ a matrix composed of the variances of the quantities displayed, and P is the power density level of a normative white noise process (which we shall refer to as the "noise ratio"). This observation noise model of remnant has been verified by extensive analysis of single-axis manual control data. We have found the noise ratio P to be approximately -20 dB (i.e., 0.01 units of normalized power per rad/sec, defined over positive frequencies) and relatively independent of the control situation. In fact the only departure from a -20 dB observation noise that we have seen in a large number of experiments was obtained in one of the last experiments discussed in this report. We do not yet fully understand the reasons for this singular result. Nevertheless, the relative invariance of the observation noise level suggests to us that foveal remnant – at least under the idealized tracking conditions that we have investigated – arises from a central-processing type of disturbance common to all tracking tasks (such as time-variational disturbances of controller gain or time delay). The

potential correspondence of equivalent observation noise to central-processing noise forms the basis for our models of task interference and pilot workload.

ORGANIZATION OF THE REPORT

In the following section of this report we present a model for task interference. Model predictions are compared with experimental results in the succeeding two sections: interference among independent tracking tasks is described in Section 3, whereas Section 4 is concerned with interference within a single axis of control. Our model for task interference leads to a metric for pilot workload which is discussed in Section 5. The major conclusions of this report are summarized in Section 6. The appendices provide details of experimental apparatus and procedures as well as complete documentation of experimental results.

2. A MODEL FOR TASK INTERFERENCE

Because the human being is inherently limited in the amount of physical and mental effort he can exert at any given time, his ability to perform psychomotor tasks will, in general, deteriorate as he is required to perform more and more tasks simultaneously. Some of our early experiments with two-axis systems showed evidence of task interference, as did the more recent studies of two- and four-axis systems described in this report. In order to provide a theoretical framework for the experimental results presented in the following sections of this report, we describe below a model for predicting interference among two or more continuous manual control tasks. Only central-processing sources of interference are considered; performance degradation associated with visual scanning or intermittent control activity are not treated at this time.

Interference which results from limitations on the human's central-processing capability have been most commonly attributed to single-channel behavior; i.e., a single channel has been postulated which must selectively "attend" to the various sensory inputs (Ref. 10). Parallel-channel behavior has also been considered in the literature; a multiband filter theory, for example, has been postulated to account for the behavior of the human *observer* when detecting multiple auditory signals (Ref. 11). In our model development we consider the human *controller* also as a parallel-channel processor of information.

The model for task interference is based on the notion that the controller possesses a fixed amount of information-processing capacity that must be shared among the various tasks to be performed. The fraction of "capacity" allocated to a particular task

is related to the equivalent observation noise ratio associated with that task. This relationship provides a mechanism by which we can incorporate the model of capacity-sharing into the framework of the optimal-control model of human behavior.

BASIC ASSUMPTIONS

The model is founded on the following primary assumptions: (a) Multiple tasks are performed in parallel, not in sequence. (b) The controller has a relatively large fixed number, N, of "information-processing channels" to distribute among his various tasks. (c) Each of these channels is perturbed by a white, Gaussian noise process which is linearly uncorrelated with all other noise processes and with system variables. The noise levels are proportional to signal variance.

There is some experimental basis for assuming parallel processing of tasks. Time-domain and frequency-domain analysis of our manual tracking data has consistently failed to show evidence that the controller time-shares among tasks. This does not rule out the possibility of a very rapid internal scanning mechanism. The scanning (i.e., sequential) model of information processing is not necessarily incompatible with the model of parallel processing, since in the limit of arbitrarily rapid scanning the two models appear to lead to the same results.

The assumption of a fixed number of channels available for tracking and other tasks is another way of saying that the controller's "channel capacity" is constant. Fortunately, we shall not have to determine N, nor shall we have to compute the noise ratio associated with each individual channel. The important point is that these numbers are assumed invariant.

The assumption of white noise processes which scale with signal variance is a direct extension of our model for controller remnant and is consistent with earlier psychophysical data which show that estimation errors tend to scale with the magnitude of the stimulus.

For mathematical convenience we refer central processing sources of interference to the input and treat them as if they were perceptual sources of interference. Accordingly we consider interference to occur between perceptual tasks. However, we assume that position and velocity information can be obtained from the same display indicator without interferring with each other. Thus we assign a perceptual task to each *indicator* provided on the display rather than to each *variable* used by the controller. This assumption is justified by our previous results (Refs. 6 and 9) which indicate that the noise ratio associated with estimation of indicator position does not depend on whether or not the pilot must also estimate indicator velocity. The noise ratio associated with estimation of velocity is similarly invariant to the requirement to estimate position. Thus, if the controller is provided with two single-axis tasks, each having a simple display of system error, then he must divide his capacity between the two displays (aside from any scanning behavior that might take place). Similarly, if the subject is provided with a single (but complex) axis of control and a display consisting, for example, of a pitch indicator and a position indicator, he must again divide his capacity between the two perceptual tasks of estimating pitch and estimating position.

EQUIVALENT OBSERVATION NOISE FOR SINGLE-VARIABLE TASKS

Let us consider a tracking task in which the subject is required to estimate (primarily) a single variable. We shall derive an expression for equivalent observation noise which explicitly

shows the dependence of the noise level upon the number of central-processing channels assigned to the task.

The flow of information is diagrammed in Fig. 2. The displayed variable [x(t)] is corrupted by the noisy, N-channel percaptual pre-processor to yield the perceived variable [y(t)]. Both x(t) and y(t) are presumed to have zero mean. The signal suffers a delay due to limitations of the pilot's neuro-motor system and is then processed by the pilot's equalizer to yield the control signal u(t). Since all sources of randomness have been reflected to the perceptual pre-processor, we need consider here only the relationship between y(t) and x(t).

The output of the n^{th} information-processing channel is

$$y_n(t) = x(t) + r_n(t) \tag{6}$$

where $r_n(t)$ is a Gaussian white noise process, having power density level $Q \cdot \sigma_x^2$, which is injected onto the n^{th} channel. (The "noise ratio" Q is the same for each channel.) Since all N channels are presumed to be allocated to this single task, the total output of the subject's perceptual pre-processor will be

$$y(t) = \sum_{n=1}^{N} [x(t)+r_n(t)] = N\ x(t) + \sum_{n=1}^{N} r_n(t) \tag{7}$$

Noting the linear independence of the $r_n(t)$, we compute the following spectrum for the perceived variable:

$$\begin{aligned} \Phi_{yy} &= N^2 \Phi_{xx} + N\ \sigma_x^2\ Q \\ &= N^2 [\Phi_{xx} + \frac{1}{N}\ \sigma_x^2\ Q] \end{aligned} \tag{8}$$

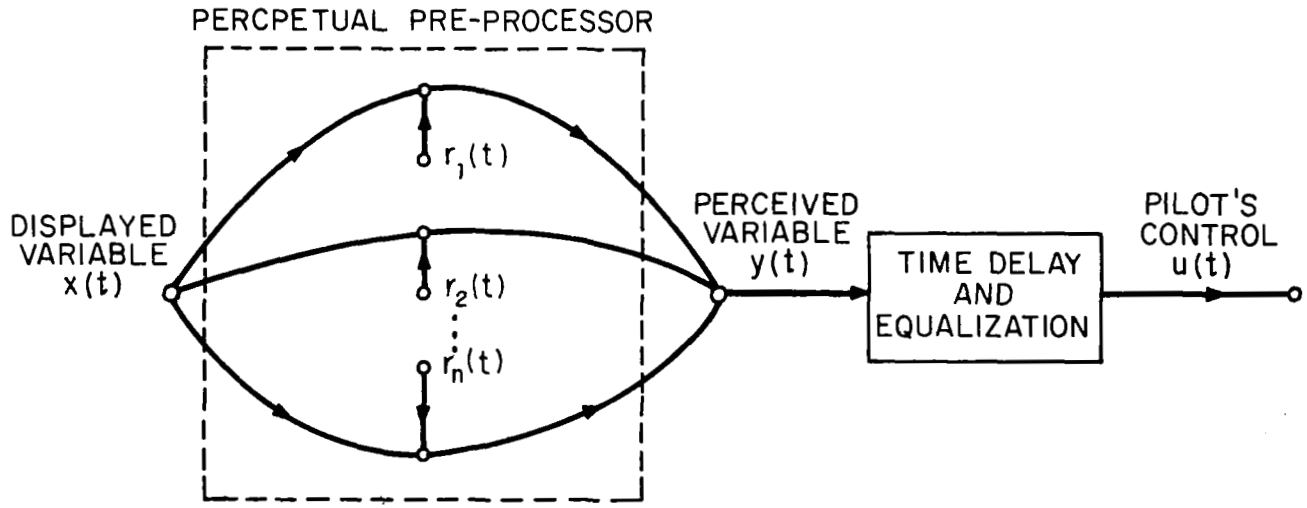


FIG.2 FLOW OF INFORMATION THROUGH THE HUMAN CONTROLLER FOR A SINGLE-VARIABLE TRACKING SITUATION

The noise-related portion of the spectrum Φ_{yy} may be referred to an equivalent observation noise process whose power density level is

$$R^{(1)} = \frac{Q}{N} \sigma_x^2 = P_o \sigma_x^2 \tag{9}$$

where the constant, P_o replaces Q/N. The unity superscript indicates that this relation holds only when the task is performed alone.

We have thus rederived the expression for equivalent observation noise which we have found to represent controller remnant in a variety of single-axis control situations. In addition, we have now seen how the noise level relates to the number of "channels" devoted to the tracking task. This relationship is crucial to our model for task interference, as shown below.

Let us now consider a situation in which M perceptual tasks are performed in parallel. Since the N information channels must now be distributed among the M tasks, the subject will be able to devote only the fraction $f_m^{(M)}$ of his channels to the m^{th} task. The equivalent observation noise is now given as

$$R_m^{(M)} = \frac{Q}{f_m^{(M)} N} \sigma_x^2 = \frac{P_o}{f_m^{(M)}} \sigma_x^2 \ . \tag{10}$$

We thus show that the primary effect of requiring the subject to perform a multiplicity of tasks is to increase the effective observation noise ratio associated with each component task. The observation noise ratio for the m^{th} task when M tasks are performed simultaneously is simply

$$P_m^{(M)} = P_o / f_m^{(M)} \quad . \tag{11}$$

Transformation of this equation yields

$$f_m^{(M)} = P_o / P_m^{(M)} \quad . \tag{12}$$

Equation (11) represents a predictive model for task interference. Given that we know the observation noise ratio P_o that corresponds to "full capacity", we can predict the increase in noise ratio associated with each displayed variable for a particular distribution of channel capacity. The optimal control model then allows us to predict the performance measures that accompany this distribution of capacity. Conversely, we can compute the fraction of capacity devoted to a given subtask when multiple, independent, tasks are performed in parallel. Provided that the subtask consists of a tracking task requiring a single displayed variable, we can readily measure the observation noise ratio under single-axis and multi-axis conditions. Equation (12) shows that the fraction of capacity associated with this task is given by the ratio of the single-axis to multi-axis observation noise ratio.

PREDICTION OF MULTIVARIABLE TRACKING PERFORMANCE

We now combine the model for task interference developed above with our optimal-control model for human performance to obtain a model for multi-task performance. The following computational procedure may be used to predict overall system performance and pilot behavior in a multivariable control situation in which the pilot's goal is to minimize a quadratic, total-task performance measure of the form given in Eq. (3):

a. Find the set of parameters for the optimal-control model of human behavior which best reproduces the controller's single-axis behavior. This is a "calibration" procedure which allows one to determine the observation noise ratio that corresponds to full capacity, as well as to determine other basic parameters such as controller time delay, subjective cost functional, and, in the case of peripheral tracking, additional internal noise processes related to peripheral viewing.

b. Compute the total performance cost as a function of the distribution of capacity among the displayed variables, where the fraction of capacity assigned to a given indicator is reflected by an increase in observation noise in accordance with Eq. (11). Since the total capacity must remain fixed, readjustments of the various observation noise ratios must meet the constraint

$$\sum_{m=1}^{M} f_m^{(M)} = \sum_{m=1}^{M} \frac{P_o}{P_m^{(M)}} = 1 \tag{13}$$

c. Select the capacity distribution that yields the minimum total performance cost. Under the assumption that the subject has been given sufficient training to learn the optimal distribution of capacity, the minimum cost is then the predicted cost. If the total task comprises a set of independent subtasks, pilot behavior and system performance for each subtask may be computed on the basis of the observation noise ratios associated with the corresponding displays. The amount of "interference" occuring on a given component task may be defined as the performance measure (score) predicted for that task in the M-task situation minus the score when the task is performed alone.

The model for task interference is validated against experimental data in the following two sections of this report. In Section 3 we consider interference among independent axes of control when two or four axes are controlled simultaneously. Interference between two display indicators related to a single axis of control is illustrated in Section 4.

3. INTERFERENCE AMONG INDEPENDENT AXES OF CONTROL

In this section we compare data obtained from several two- and four-axis manual control experiments with predictions obtained from the model for task interference. The experimental and analytical procedures used in obtaining these data are described in detail in Appendix A, and a complete presentation of the experimental results is given in Appendix B. We shall consider here only those experiments which relate directly to the study of central-processing sources of task interference.

EXPERIMENTAL CONDITIONS

The subjects were provided with the four-axis display configuration shown in Fig. 3. Each component display consisted of a moving error bar and a stationary reference line presented on an oscilloscope. This configuration provided four viewing conditions: (a) foveal, (b) 16° peripheral with reference extrapolation possible – as, for example, when fixating the upper left display and tracking a signal on the upper right, (c) 16° peripheral with no reference extrapolation possible, and (d) 22° peripheral also with no reference extrapolation. Since the stationary reference line became imperceptible a few seconds after peripheral viewing was initiated, the subject's tracking performance was appreciably enhanced whenever he could extrapolate the zero reference from his fixation point to the peripheral display.

A single display was fixated during the entire run length, and two or more axes were controlled simultaneously. Each display used in a given multiaxis experiment was also tracked singly to provide a set of baseline measures to allow each subject to serve as his own control. Two two-axis manipulators were provided – one

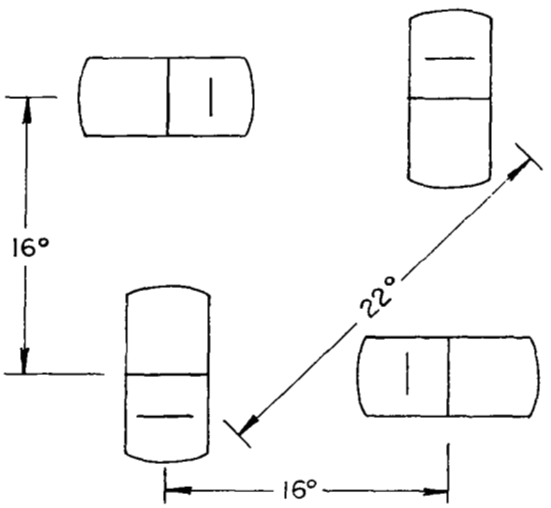


FIG.3 DISPLAY CONFIGURATION USED IN THE EXPERIMENTS

Dimensions Shown in Degrees of Visual Arc

controlled by each hand – in order to provide control-display compatibility. Velocity control was provided on each axis, and each input signal was constructed from a number of sinusoidal components chosen to simulate a first-order noise process having a break frequency at 2 rad/sec. The input was applied in parallel with the controller's output (i.e., as a disturbance on vehicle velocity). Different input waveforms were used on each axis so that the subject would perceive no linear correlations among the inputs. The subjects – all of them instrumented-rated aircraft pilots – were instructed to minimize mean-squared system error when performing a single task and to minimize the sum of the component MS error scores when tracking multiple axes.

FOUR-AXIS TRACKING: ONE FOVEAL AND THREE PERIPHERAL AXES

We first analyze the results of an experiment in which the subjects were required to fixate the upper left display while tracking all displays simultaneously. No attempt was made to equalize the component task difficulties. On the contrary, since all four input signals were statistically identical, the component task difficulty (in terms of the error score) increased with decreasingly favorable viewing conditions. In order to obtain measures that were most likely to be representative of subject behavior in general, and also to conserve modelling effort, the performance measures of the four subjects were averaged together for comparison with model predictions. The various performance measures were averaged in a way that was both convenient and internally consistent: we computed a geometric* average of the error variance scores along with algebraic averages of the normalized observation noise spectra (in dB), controller amplitude ratio (in dB), and controller phase shift (in degrees).

*We define the "geometric average" of the variable (x) as $\left[\prod_{i=1}^{N} (x_i)\right]^{\left(\frac{1}{N}\right)}$.

The model parameters of relative weighting on control-rate variance, time delay, and observation noise ratio were adjusted to provide a good match between model output and single-axis foveal tracking results. The cost weighting on error variance was set to unity, and all other variables (except rate of control output) were given zero weighting in the cost functional. The motor noise variance was set at a low enough level so as to have a negligible effect on the model's output. The relative weighting on control rate was 0.0002, which introduced an effective lag time constant of 0.0841 seconds. A time delay of 0.17 seconds and an observation noise ratio of -21.0 dB were found to provide a good match. This observation noise ratio, therefore, was selected as the ratio representing "full attention" to the task.

Our criterion for a "good" match was that the predicted and measured error scores agree to within 10%, and the normalized observation noise spectra and controller amplitude ratios generally match to within 2 dB. (When the above requirements were met, the predicted and measured phase shifts generally agreed to within 10 degrees at mid-frequencies.) We did not attempt to match the error rate or control-related scores; the predicted values of these quantities tended to be on the order of 15 to 30 percent less than the measured values. Note that our primary concern here was to choose model parameters such that *changes* in model behavior corresponding to increased observation noise would be representative of the single-axis, multi-axis differences in controller behavior. We did not feel that it was necessary to match every detail of controller behavior to achieve this goal.

It was necessary to extend our model for equivalent observation noise to account for the effects of peripheral viewing on single-axis tracking performance. Since the single-axis experimental results (described fully in Appendix B) indicated that the

peripheral observation process was not proportional to signal variance, we adopted the following simple model to account for peripheral viewing effects:

$$\underline{R} = (\underline{\sigma}^2 + \underline{\sigma}_p^2) \cdot P_o \qquad (14)$$

where $\underline{\sigma}_p^2$ is the effective (vector) variance of an internal noise process and represents peripheral threshold effects. $\underline{\sigma}_p^2$ is assumed to be invariant with respect to the variance of the displayed signal $\underline{\sigma}^2$, but it will depend on display location and orientation. P_o is the noise ratio when full attention is devoted to the display. The peripheral vector variance has two components associated with each indicator – effective thresholds for position and rate – which are not assumed necessarily to have the same functional dependency on display parameters.

In order to quantify the components of $\underline{\sigma}_p^2$ for each of the three peripheral viewing conditions that we used, the control-rate weighting, time delay, and observation noise ratio were kept at their nominal (foveal) values, and the peripheral noise variances were adjusted to match the single-axis peripheral measures. Table 1 shows the parameter values that were found to match the 1-axis data corresponding to the four viewing conditions. Comparisons of measured and "predicted" single-axis normalized observation noise spectra* and controller describing functions are given in Figs. 4-7.

*Since we cannot measure the two component spectra of the *vector* observation noise process $\underline{r}(t)$, we have reflected controller remnant to an equivalent *scalar* noise process injected onto system error. See Appendix A for details of the computational procedure.

TABLE 1

Model Parameters to Match Single-Axis Measurements

Parameter	Viewing Conditions			
	Foveal	16°Periph Ref Ext	16°Periph No Ref Ext	22° Periph No Ref Ext
Time Delay τ(sec)	.17	.17	.17	.17
Lag Factor T_N(sec)	.0841	.0841	.0841	.0841
Obs. Noise Ratio P_o (dB)	-21.0	-21.0	-21.0	-21.0
Peripheral Error Variance $\sigma^2_{px}(deg^2)$	0.0	3.76	7.62	59.2
Peripheral Error Rate Variance $\sigma^2_{p\dot{x}}([deg/sec]^2)$	0.0	2.96	8.46	15.9

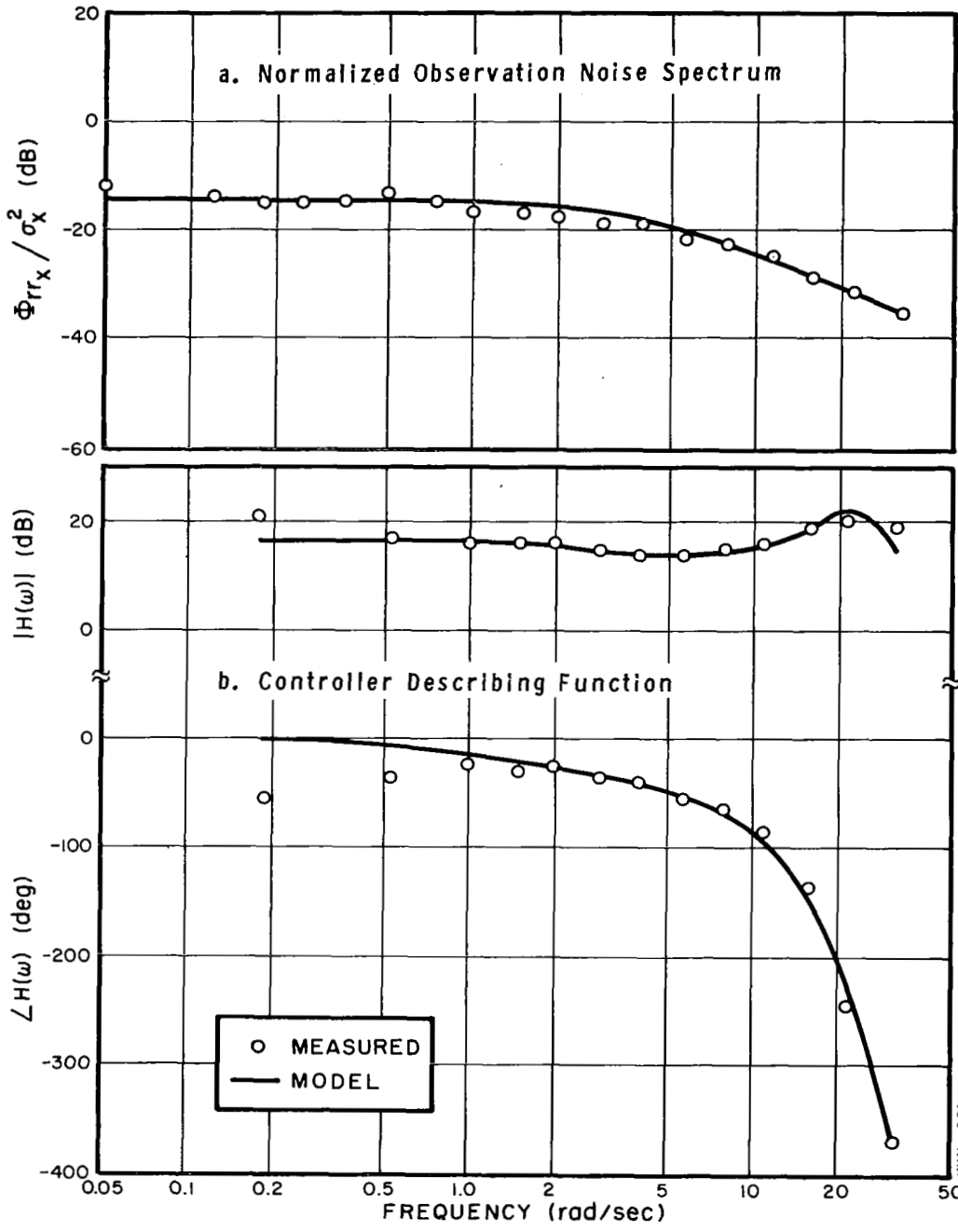


FIG.4 SINGLE-AXIS FREQUENCY-DOMAIN MEASURES: FOVEAL VIEWING
Average of 4 subjects, 2 trials/subject

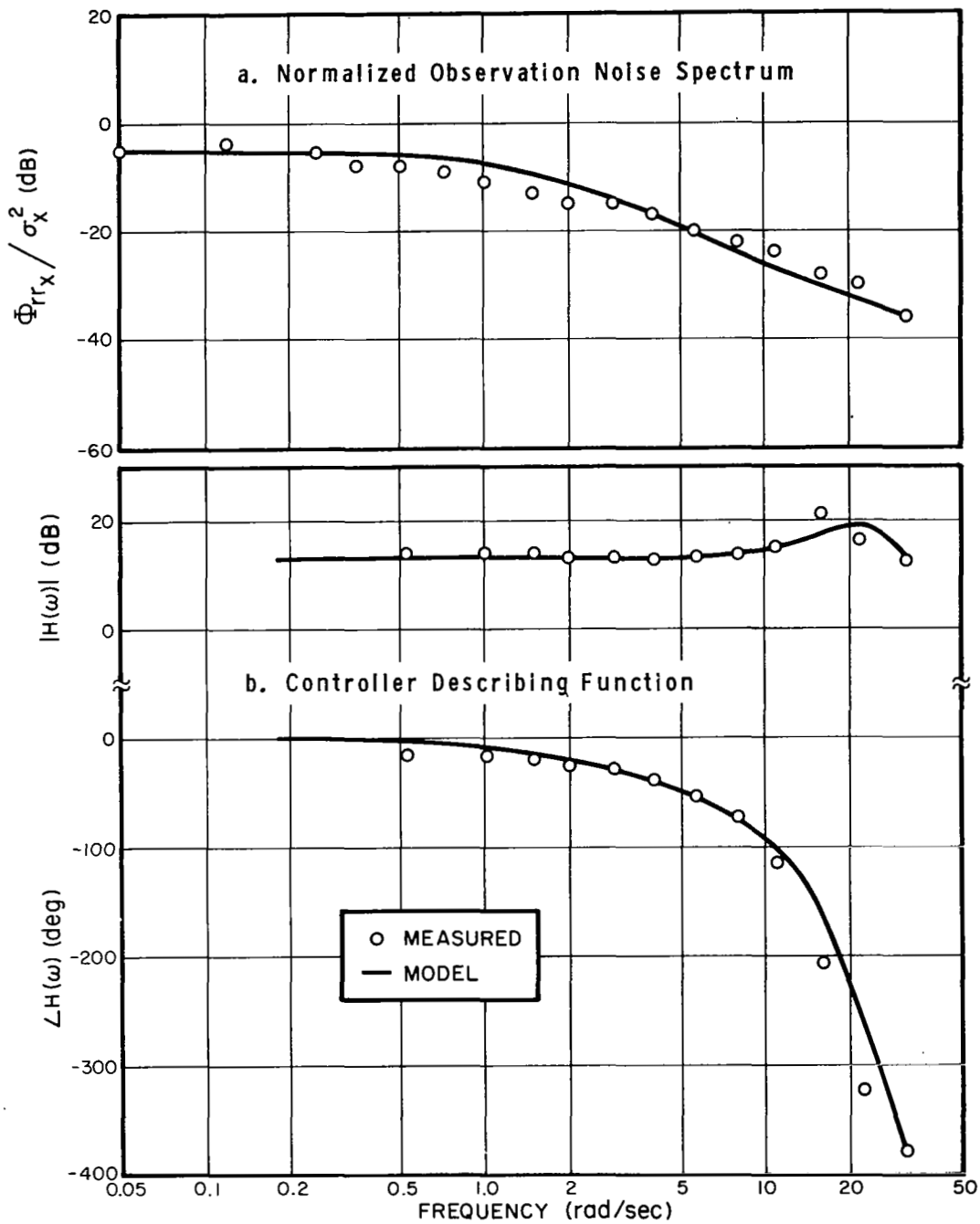


FIG.5 SINGLE-AXIS FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITH REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

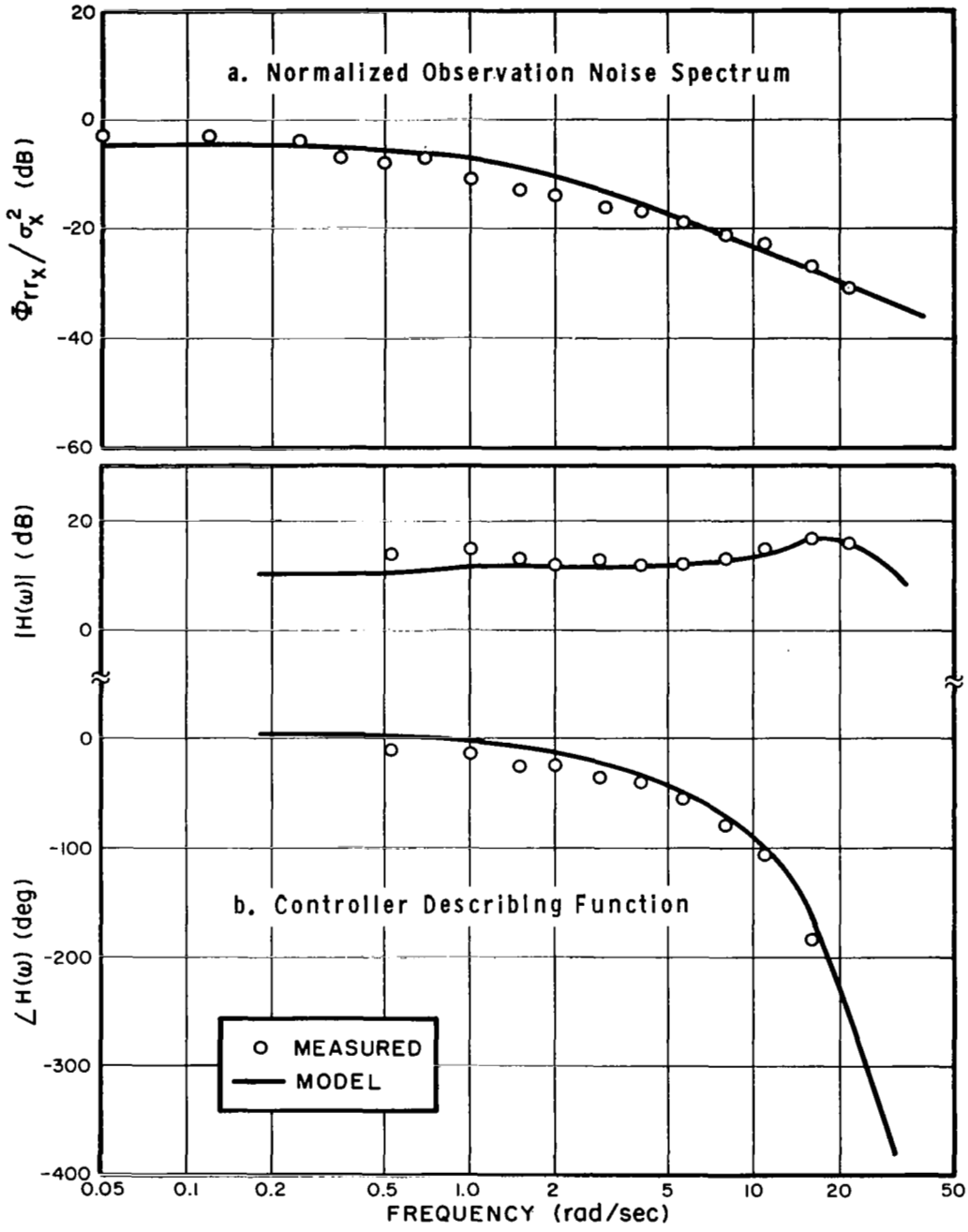


FIG.6 SINGLE-AXIS FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

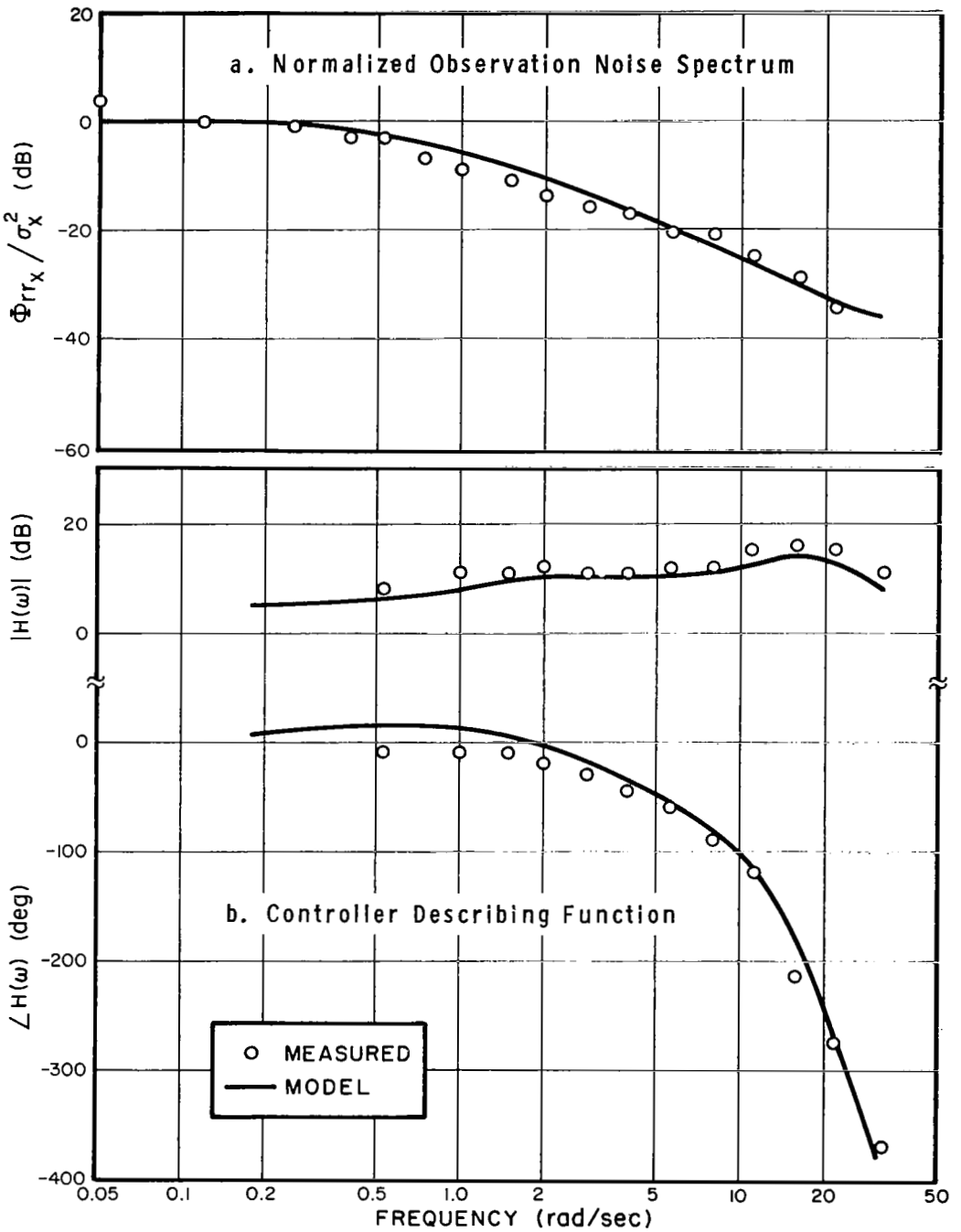


FIG.7 SINGLE-AXIS FREQUENCY-DOMAIN MEASURES: 22° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

Once the four single-axis conditions had been matched, our next task was to compute the total-task performance as a function of the allocation of capacity. Since the total task consisted of four independent tracking tasks, this part of the analytical procedure was performed in two steps. First, the optimal-control model was analyzed with several values of observation noise ratio to predict for each axis the relationship between performance score (in this case, error variance) and fraction of capacity. The remaining model parameters were kept fixed at the numerical values shown in Table 1. The predicted relationships are shown for the four viewing conditions in Fig. 8. Optimum four-axis performance was obtained by locating the operating point which yielded minimum total score – defined as the sum of the four component error variance scores – subject to the constraint that the fractions of capacity sum to unity.

Predicted and measured error variance scores are compared in Table 2. The 1-axis and 4-axis scores obtained from the manual control experiments are shown in Table 2a. Also shown are the ratios of the 4-axis to 1-axis scores for each viewing condition and for the total performance measure. Table 2b indicates the predicted optimum 4-axis performance: i.e., the performance corresponding to the distribution of capacity that would yield the minimum total score.

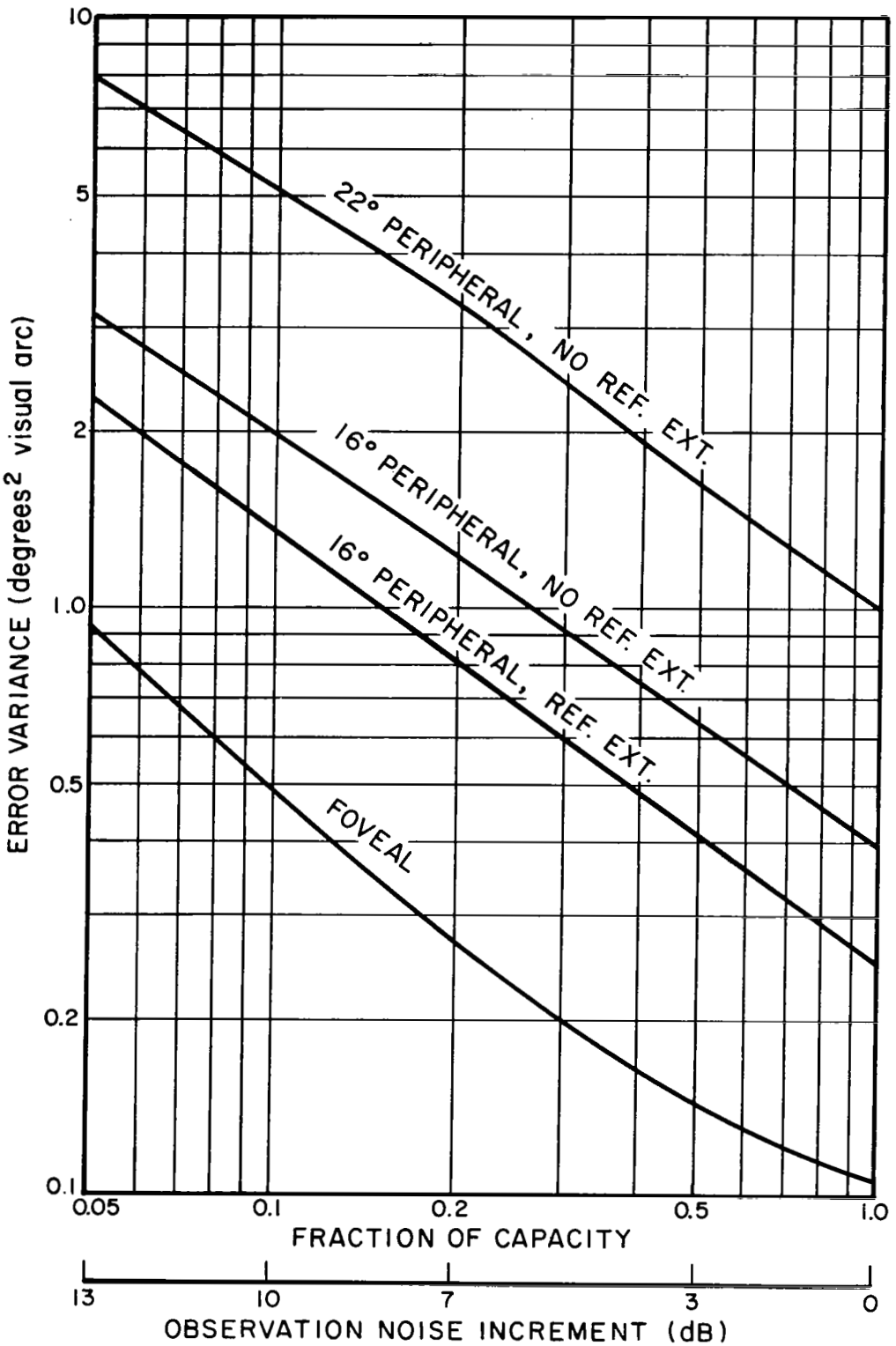


FIG.8 PREDICTED RELATIONSHIP BETWEEN ERROR VARIANCE AND FRACTIONAL ALLOCATION OF CAPACITY

TABLE 2

Comparison of Measured and Predicted
Error Variance Scores for 4-Axis Experiment

	Measurement	Viewing Condition				Total Score
		Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext	
(a) Measured	1-axis	.11	.25	.42	.96	1.7
	4-axis	.27	.94	1.3	1.6	4.1
	Ratio	2.5	3.8	3.0	1.7	2.4
(b) Predicted: Optimal Behavior	1-axis	.11	.25	.39	.98	1.7
	4-axis	.49	.82	1.1	1.8	4.2
	Ratio	4.6	3.3	2.7	1.9	2.4
	Fract. Cap.	.10	.20	.25	.45	1.00
	Noise Ratio (dB)	-11.0	-14.0	-15.0	-17.5	
(c) Predicted: Best Match of Subjects' Behavior	1-axis	.11	.25	.39	.98	1.7
	4-axis	.27	1.0	1.3	1.7	4.2
	Ratio	2.5	4.2	3.2	1.7	2.4
	Fract. Cap.	.20	.15	.20	.50	1.05
	Noise Ratio (dB)	-14.0	-12.8	-14.0	-18.0	

Error score in deg^2 visual arc

Average of 4 subjects, 2 trials/subject

Since the subjects were not instructed as to how to apportion the total error among the component scores, the most critical test of the interference model is its ability to predict total score. Table 2 shows that the predicted total performance score was within 3% of the measured score. The model predicts less well the performance on the component axes. The subjects achieved lower scores on the foveal and 22° peripheral tasks than the model would predict, whereas experimental scores were greater than predicted on the axes corresponding to 16° peripheral viewing. These results suggest that the subjects "traded" performance on one pair of component tasks for performance on the remaining two with little degradation in total performance.

The predicted fractions of capacity allocated to each component task are shown in Table 2b. Also shown are the corresponding observation noise ratios. The trend of the model predictions agrees with our intuitive expectations; the predicted fractions of capacity ranged from 0.10 for the foveal task to 0.45 for the 22° peripheral task.

Since the subjects' allocation of capacity was apparently different from that predicted by the model, a simple model-matching procedure was used to determine the actual allocation of capacity. The curves of Fig. 8 were used to associate a fraction of capacity with the error variance score obtained on each axis when the four axes were controlled together. The results of this procedure are given in Table 2c.* The fractional capacities obtained by this procedure were not subject to the constraint that they sum to unity; nevertheless, we note in Table 2c that the sum of the fractional

*Because we quantized our model results to the nearest integral multiple of 0.05 units of fractional capacity, we could not match the four-axis score perfectly. Comparison of Table 2a and 2c shows, nevertheless, that matching errors were less than 10%.

capacities is, in fact, nearly unity. It appears, then, that the subjects were operating within the constraint of a total fixed capacity.

Model predictions were obtained using the noise ratios shown in Table 2c in order to demonstrate that our model structure accounts for the kinds of 1-axis, 4-axis differences observed in the human controller's frequency-domain measures. (Note that the describing functions and observation noise spectra were not matched by the procedure described above — only the error scores were matched.) Experimental and measured controller describing functions and normalized observation noise spectra are compared for each of the viewing conditions in Figs. 9-12.

The model predicted the important trends of the differences; namely, the increase in normalized observation noise level, the decrease in controller gain, and the slight increase in high-frequency phase lag as the number of axes tracked was increased from 1 to 4. The most noticeable discrepancy between predicted and measured trends was observed in the observation noise results. Whereas the experimental data show the 1-axis and 4-axis normalized spectra nearly coinciding at high frequencies, the model indicates that the 1-axis, 4-axis differences should be small at low frequencies and larger at high frequencies. We suspect that this discrepancy may have resulted partly from a greater reliance on rate information (relative to position information) than was optimal.* That is, the subjects may have "traded" position information for rate information within a given axis of control, just as they apparently traded performance on one axis for performance on another, without appreciably affecting their total score. The

*In Ref. 9 we present a simple model for controller remnant which shows that the break frequency of the normalized observation noise spectrum, as well as the asymptotic low-frequency level, is determined by the ratio of the controller's gain on rate information to his gain on position information.

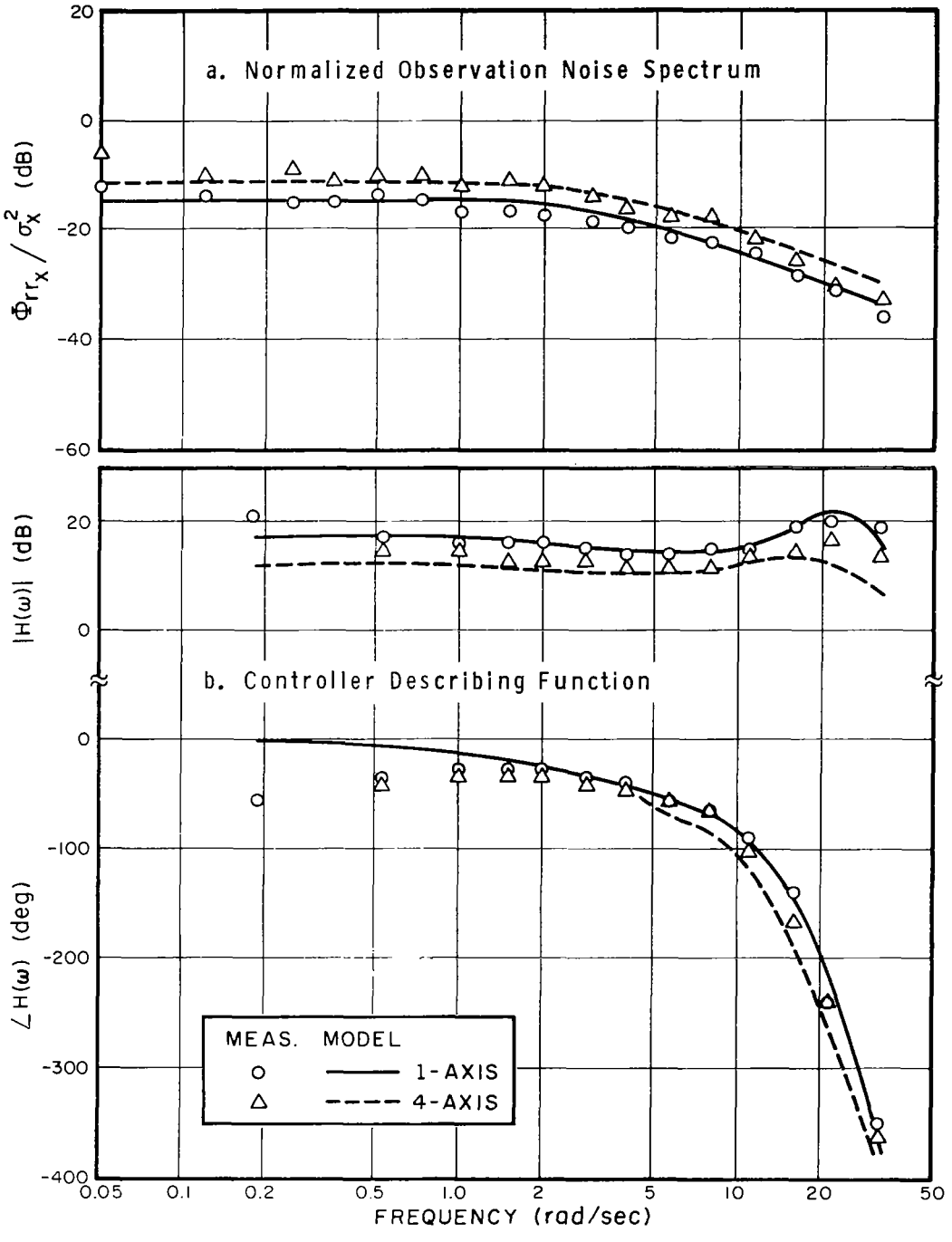


FIG.9 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: FOVEAL VIEWING

Average of 4 subjects, 2 trials/subject

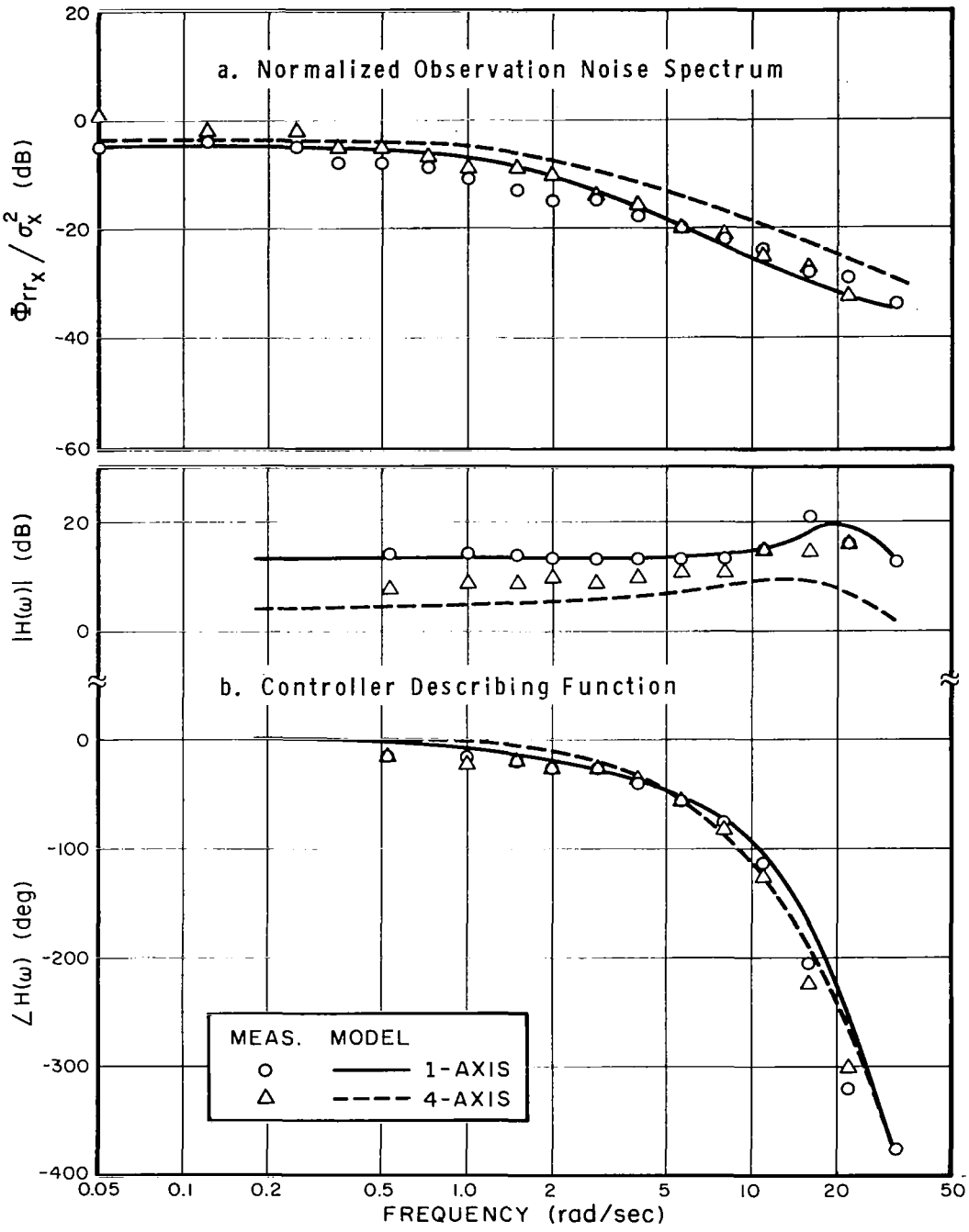


FIG. 10 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITH REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

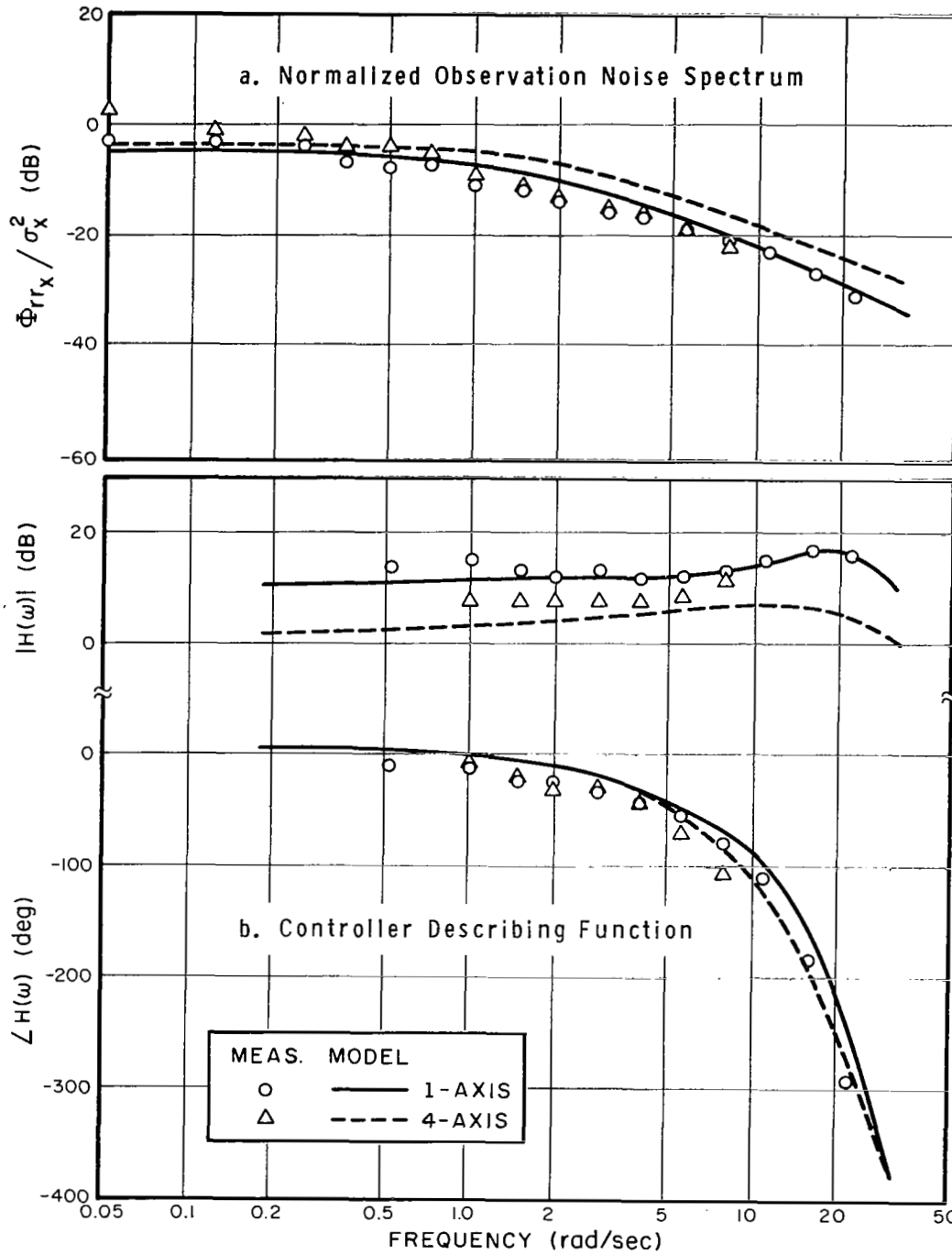


FIG.11 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

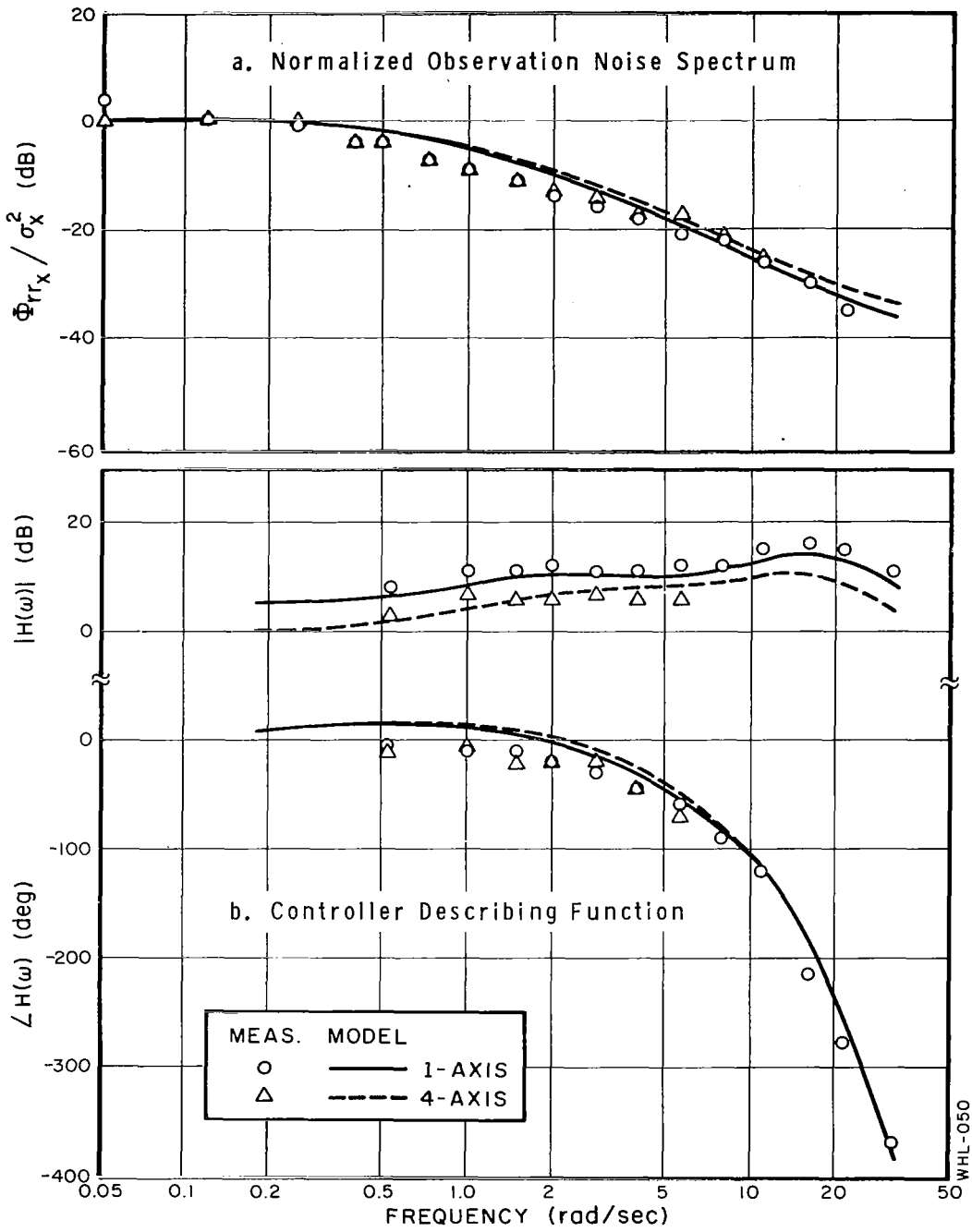


FIG. 12 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 22° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

high-frequency portions of the amplitude-ratio curves shown in Figs. 9-11 are consistently higher than those predicted by the model, which is consistent with the notion that the subjects were using somewhat more velocity information than predicted.

TWO-AXIS TRACKING: ONE FOVEAL AND ONE PERIPHERAL TASK

Model predictions are now compared with the results of a two-axis tracking experiment. Vehicle dynamics and forcing-function spectra were the same as for the four-axis task. For this experiment we adjusted the mean-squared inputs so that the foveal and peripheral tasks would be of approximately the same difficulty in terms of the single-axis mean-squared error score.

The display configuration of Fig. 3 was used, but only two adjacent displays were active during a given tracking run. The subject was required to fixate a single display throughout the run and to track that display foveally as well as to track peripherally the display in the nearest clockwise position. Single-axis foveal and peripheral measurements were also obtained. The subject was thus always able to extrapolate a zero reference to the peripheral display in this experiment. All four pairs of displays were tracked in sequence. When the two displays were in the same horizontal plane, two hands were needed for control, whereas only one hand operating a two-axis manipulator was needed when the displays were in the same vertical plane. In order to eliminate a direct source of motor interference, we consider only those trials in which two hands were needed for control.

Rather than recalibrate the model against the single-axis data obtained in this experiment, we used the single-axis results obtained in conjunction with the previous experiment to predict the two-axis total and component error scores. We assumed the model parameters

were the same as for the four-axis experiment and that therefore the relationship between error variance and fraction of capacity was as shown in Fig. 8. The lower two curves of this figure were translated along the ordinate so that the scores corresponding to full capacity matched the single-axis foveal and peripheral error variance scores measured in this experiment. (A change in scale factor was needed primarily to account for the increased mean-squared disturbance input applied to the foveal axis, and secondarily for learning effects.) Since only two axes were tracked in this experiment, the model results were readily combined to provide the curve of total error score versus the fraction of capacity on the foveal task shown in Fig. 13. The fractional distribution of capacity yielding the minimum total score was obtained from this curve.

Predicted and measured error scores, along with the predicted allocation of capacity to the foveal axis, are shown in Table 3. As is the case with the 4-axis results, the predicted total score is extremely close to the measured total score – a difference of less than 2 percent. The foveal and peripheral components of the score were also predicted reasonably well. Since differences between measured and predicted component scores were no larger than 10 percent, we did not attempt to achieve a finer match to the 2-axis results. The model predicted a capacity distribution of 50 percent to each axis. Figure 13 shows that the total performance score was relatively insensitive to distribution of capacity; the subjects could have allocated up to 60 percent of their capacity to one axis or the other without increasing their total score by more than 5 percent above the predicted minimum.

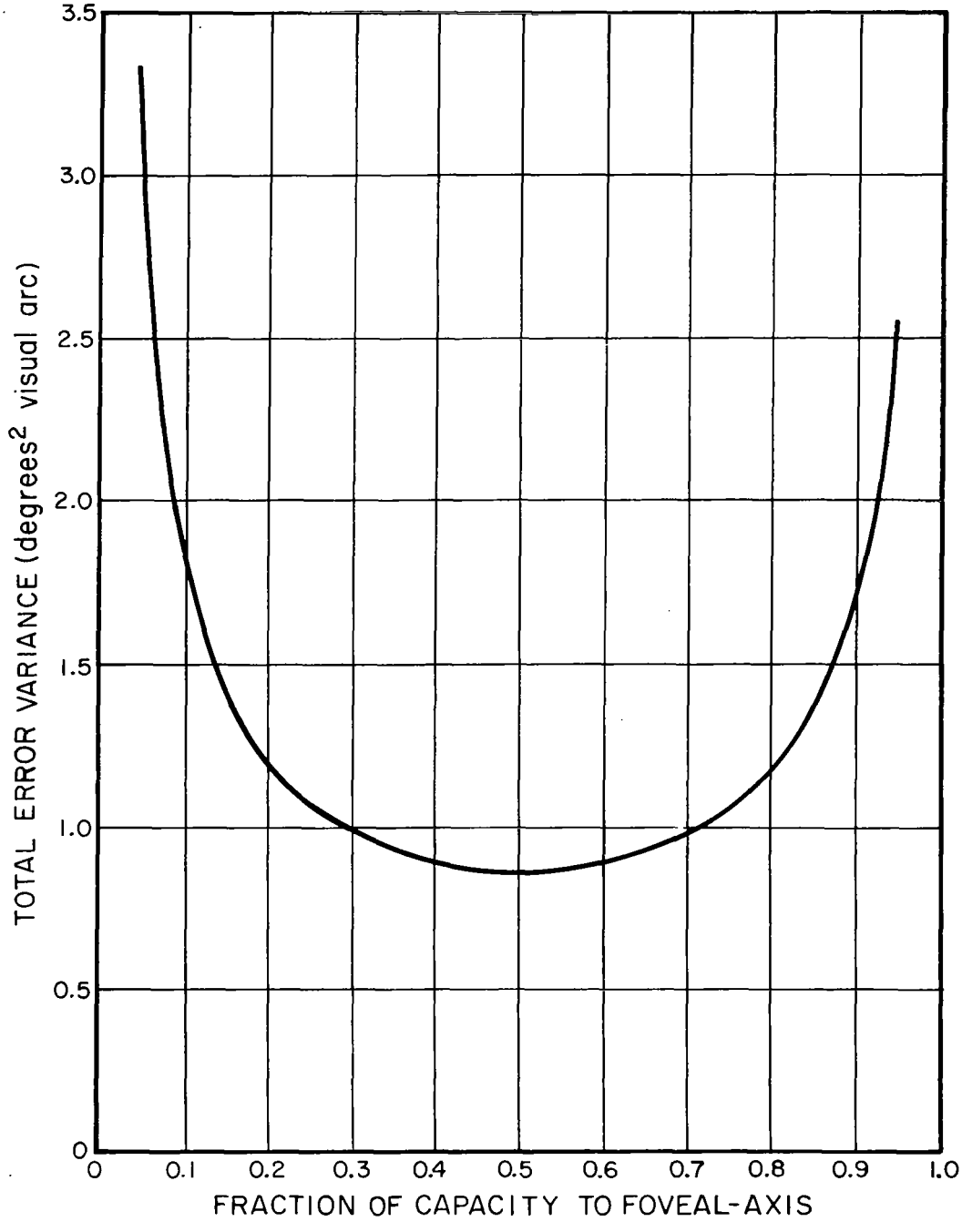


FIG. 13 PREDICTED RELATIONSHIP BETWEEN TOTAL ERROR VARIANCE AND FRACTION OF CAPACITY ON FOVEAL TASK: TWO-AXIS TRACKING

TABLE 3

Comparison of Measured and Predicted Error Variance Scores for the Two-Axis Task: Two-Handed Control Only

Measurement		Viewing Condition		Total Score
		Foveal	16°Periph Ref Ext	
Measured	1-axis	.35	.24	.58
	2-axis	.44	.42	.87
	Ratio	1.3	1.8	1.5
Predicted	1-axis	.35	.24	.58
	2-axis	.48	.39	.87
	Ratio	1.4	1.6	1.5
	Fract. Cap.	.50	.50	1.0
	Noise Ratio (dB)	-18.0	-18.0	---

Error score in deg^2 visual arc

Average of 4 subjects, 4 trials/subject

TWO-AXIS TRACKING: TWO FOVEAL TASKS

The model for task interference also accounts for interference observed in one of our earlier two-axis experiments (Ref. 4). Independent axes of K/s dynamics were used in this experiment, also. Each input signal (which was applied as a command signal rather than a vehicle disturbance) was composed of sinusoids designed to simulate a rectangular spectrum having a bandwidth of 2 rad/sec. In addition, a low-amplitude, high-frequency "shelf" was added to the input to provide measurements at high frequencies. Two error dots – one for each axis of tracking – were displayed to the subject on a single oscilloscope. The dots moved in the vertical dimension, and their axes of travel were separated horizontally by about 0.8° visual arc. Since the diameter of the fovea is about 2° visual arc, we assume that the dots were both viewed foveally most of the time. Separate controls were provided for each axis.

Model parameters were first chosen to provide a good match to the single-axis experimental results. The observation noise ratio was then doubled to simulate an equal division of pilot capacity between the two axes.* A second-order Butterworth representation of the input spectrum was found adequate to yield a reasonably accurate prediction of controller behavior. Model parameters were chosen which were found to be typical of the subjects who participated in this experiment: (a) the control-rate weighting was chosen to yield an effective lag time constant of 0.084 sec, (b) time delay was set at 0.17 sec, and (c) an observation noise ratio of -21 dB was adopted as the "full capacity" noise ratio.** In

---

*Observation noise spectra were unavailable for the computation of capacity allocation. We therefore assume 50 percent capacity on each axis. As we have shown in the previous examples, total performance is relatively insensitive to capacity allocation.

**The subjects used in this early experiment were the same ones who participated in the 4-axis experiment described earlier in this section. Hence, we use the same values for model parameters that we used to predict 4-axis performance.

addition, a nonzero motor noise level was needed because the input was a command signal and not a disturbance applied in parallel with the pilot's control. This noise level was set at about -26 dB relative to control power to be consistent with motor noise levels used in matching earlier data obtained with command inputs (Ref.6).

Measured and predicted normalized mean-squared mean-squared-error scores, averaged across the two axes, are compared in Table 4. The predicted scores differed by less than 10% from the measured scores. Most importantly, the predicted ratio of the 2-axis score to the 1-axis score (about 1.4) was very nearly that which was observed experimentally. Thus, the model for task interference accounts for the effect upon the performance score of adding the second axis of control.

TABLE 4

Effect of Number of Axes Tracked on Measured and Predicted Normalized Mean-Squared Error Scores

	Normalized Mean-Squared Error		
	1-axis	2-axis	Ratio
Measured	.050	.073	1.45
Predicted	.047	.065	1.40

Average of 4 subjects, 2 trials/subject

DISCUSSION OF THE INTERFERENCE MODEL

A model for task interference has been described which is based on the notion that the human controller possesses a fixed amount of capacity which he allocates among the various tracking tasks to be performed. By relating allocation of capacity to changes in equivalent observation noise ratio, we have employed the optimal control model of human behavior to predict total-task performance scores very accurately. In addition, the effects of interference on controller describing functions and observation noise spectra are reasonably well predicted, although less accurately than the effect on total performance score. The ability to predict correctly the trends along all measurement dimensions supports the assumption that interference effects can be related directly to a change in the observation noise ratio.

The effects of task interference on measurements obtained from a given axis of control were, in general, less well predicted than the total performance measure (which was the only measure that the subjects were instructed to regulate). This lack of precision may have stemmed partly from an inadequate representation of the effects of peripheral viewing on observation noise. However, we suspect that the major source of error in predicting component-task performance was the flexibility permitted in achieving near-optimum system performance. For example, studies performed with the model indicated that the total error score was relatively insensitive to allocation of pilot capacity about the nominal optimum. Similarly, we would expect a good flight control system to be somewhat insensitive to details of the pilot's control and monitoring strategy. In general, the accuracy with which we can predict performance along a given measurement dimension should be related directly to

the precision with which the measured variable must be regulated in order to achieve the required overall system performance.

A more serious limitation on the generality of our model for task interference is that we do not yet know how to predict what the pilot's total capacity (in terms of an equivalent observation noise ratio) will be in a given control situation. Although we have found single-axis noise ratios to be about -20 dB for control tasks involving stable vehicle dynamics, we show in the next section of this report that noise ratios as low as -26 dB were measured when the dynamics were unstable. (The subjects apparently were induced to achieve a particularly low noise level because of the high sensitivity of system performance to the observation noise ratio.) Until we understand better how the pilot's apparent capacity depends upon the nature of the control situation, single-variable "calibration" experiments will be necessary so that nominal model parameter values can be determined. We continue to assume, on the basis of our results, that the pilot's capacity remains fixed for a given type of control situation (and for a given level of pilot proficiency).

Although model and experimental results are in good agreement, the phenomenon of task interference is clearly more complex than our model would indicate. For example, we show in Appendix B that two-axis interference is greater when a single two-axis manipulator is used than when two single-axis manipulators are employed. This result is suggestive of peripheral motor interference effects in addition to the central-processing interference that we have modelled. The near-perfect prediction of the total four-axis score is a bit surprising, then, when one considers that motor interference must have been present in this control situation. Apparently, there were compensating errors in our model structure (e.g., the subjects may have increased their total capacity in this very demanding control situation).

The data against which the model for interference has been tested have been obtained entirely from experiments in which the subjects were not allowed to scan visually. Although we have obtained a body of data relating to four-axis control with scanning permitted (see Appendix B), we have not been able to model this situation properly. The current implementation of the optimal-control model does not now allow us to analyze the situation in which both scanning and task interference occur. This does not reflect a conceptual limitation on the interference model, however. Consideration of observation noise processes whose statistics vary with time in accordance with the subject's fixation point (see Ref. 5) should enable the optimal-control model to yield predictions of pilot behavior which include the effects both of visual scanning and of central-processing limitations. Further development of the optimal-control model will be necessary before this capability is realized.

Since modern aircraft displays are tending to place multiple display elements in close proximity to one another, our ability to predict the performance of such flight-control systems may depend increasingly less upon our ability to predict scanning behavior and more upon the accuracy with which we can predict mutual interference among elements viewed foveally. We have shown that our model for task interference accounts for interference between two linearly independent, foveal displays. The experiment described in the following section of this report deals with the nature of interference between two highly coupled display elements.

the precision with which the measured variable must be regulated in order to achieve the required overall system performance.

A more serious limitation on the generality of our model for task interference is that we do not yet know how to predict what the pilot's total capacity (in terms of an equivalent observation noise ratio) will be in a given control situation. Although we have found single-axis noise ratios to be about -20 dB for control tasks involving stable vehicle dynamics, we show in the next section of this report that noise ratios as low as -26 dB were measured when the dynamics were unstable. (The subjects apparently were induced to achieve a particularly low noise level because of the high sensitivity of system performance to the observation noise ratio.) Until we understand better how the pilot's apparent capacity depends upon the nature of the control situation, single-variable "calibration" experiments will be necessary so that nominal model parameter values can be determined. We continue to assume, on the basis of our results, that the pilot's capacity remains fixed for a given type of control situation (and for a given level of pilot proficiency).

Although model and experimental results are in good agreement, the phenomenon of task interference is clearly more complex than our model would indicate. For example, we show in Appendix B that two-axis interference is greater when a single two-axis manipulator is used than when two single-axis manipulators are employed. This result is suggestive of peripheral motor interference effects in addition to the central-processing interference that we have modelled. The near-perfect prediction of the total four-axis score is a bit surprising, then, when one considers that motor interference must have been present in this control situation. Apparently, there were compensating errors in our model structure (e.g., the subjects may have increased their total capacity in this very demanding control situation).

The data against which the model for interference has been tested have been obtained entirely from experiments in which the subjects were not allowed to scan visually. Although we have obtained a body of data relating to four-axis control with scanning permitted (see Appendix B), we have not been able to model this situation properly. The current implementation of the optimal-control model does not now allow us to analyze the situation in which both scanning and task interference occur. This does not reflect a conceptual limitation on the interference model, however. Consideration of observation noise processes whose statistics vary with time in accordance with the subject's fixation point (see Ref. 5) should enable the optimal-control model to yield predictions of pilot behavior which include the effects both of visual scanning and of central-processing limitations. Further development of the optimal-control model will be necessary before this capability is realized.

Since modern aircraft displays are tending to place multiple display elements in close proximity to one another, our ability to predict the performance of such flight-control systems may depend increasingly less upon our ability to predict scanning behavior and more upon the accuracy with which we can predict mutual interference among elements viewed foveally. We have shown that our model for task interference accounts for interference between two linearly independent, foveal displays. The experiment described in the following section of this report deals with the nature of interference between two highly coupled display elements.

4. INTERFERENCE WITHIN A SINGLE AXIS OF CONTROL

One of the assumptions underlying the model for task interference described in Section 2 is that central sources of interference can be treated as perturbations of the equivalent observation noise. When system state information is to be presented visually, we assume that interference occurs among the display indicators viewed by the pilot. In the preceding section of this report we showed that the interference model provided reasonably good predictions of the effects of interference for a set of multi-axis control situations in which there were no linear correlations among the display indicators. In this section we discuss experiments conducted to determine whether or not the model also applies when the displayed variables are highly correlated. Our experimental strategy was similar to that described previously: (a) a set of single-variable calibration trials was run to determine nominal values for model parameters, (b) manual control data were obtained using two indicators, and (c) model predictions were compared with experimental data to determine the nature of the interference.

EXPERIMENTAL CONDITIONS

The basic experimental apparatus and analytical tools described in Appendix A were employed in this experiment. The specific experimental design is described below.

The subjects were required to control dynamics of the form

$$V = \frac{K}{s(s-1)} \tag{15}$$

where K was set to 1.0 degrees visual arc per second^2 per newton of control force. These dynamics were implemented as the cascade of two subsystems K_1/s and $K_2/(s-1)$, as shown in Fig. 14. Two display conditions were investigated. For half the training and data trails, only the system error x(t) was displayed; both x(t) and the auxiliary signal y(t) (the output of the first subsystem) were displayed for the remaining trials. Numerical values for the experimental parameters shown in Fig. 14 are tabulated in Table 5.

TABLE 5

Experimental Units and Parameter Values for the First Experiment

Parameter	Units	Value
i(t)	volts	(variable)
u(t)	newtons	(variable)
x(t)	degree visual arc	(variable)
y(t)	degree visual arc	(variable)
z(t)	volts	(variable)
K_1	(arc-degree/sec/volt	0.20
K_2	(degree/sec)/degree	2.50
K_I	(dimensionless)	0.157
K_u	volts/newton	2.00
$K_u \cdot K_1 \cdot K_2$	(degree/sec^2)/newton	1.00
σ_I^2	volts^2	87.2

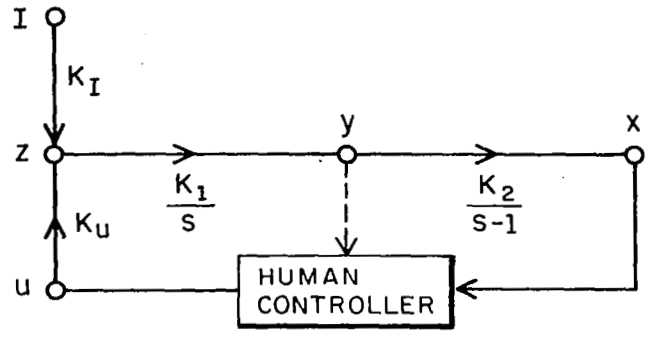


FIG.14 DIAGRAM OF THE SINGLE-AXIS, MULTIVARIABLE CONTROL SITUATION

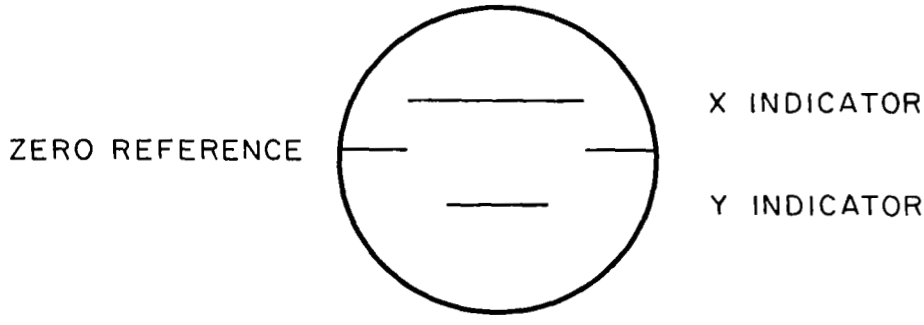


FIG.15 EXPERIMENTAL DISPLAY FORMAT

The vehicle dynamics shown in Eq. (15) were selected for this experiment for two reasons. First, the system was controllable when x(t) was displayed alone. An experiment could therefore be performed to determine pilot (i.e., model) parameters of time delay, subjective weighting, and observation noise ratio with the subjects controlling the same system that they would have in the multi-indicator situation. Our ability to perform this kind of calibration experiment was crucial to the test of our model for interference. Second, preliminary model analysis indicated that performance would be considerably enhanced if there were no interference between the subject's perception of x(t) and his perception of y(t). We thus could expect a sensitive test of the interference hypothesis.

Figure 15 shows the display format used when both x(t) and y(t) were displayed. These two quantities were presented electronically as two bars of light. The zero reference was identified by two short segments of tape placed on the horizontal axis of the 'scope face. We were initially uncertain of the appropriate format for displaying x(t) alone. In order to provide a consistent zero reference throughout this experiment, we desired to blank the Y indicator when y(t) was not displayed. On the other hand, consistency with previous experiments dictated that the Y indicator be displayed as an auxiliary zero reference. Since we did not know whether or not performance would be significantly affected by the presence or absence of the auxiliary zero reference, half of the x(t)-alone trials were conducted with each of these two display formats.

The input forcing function consisted of a simulated first-order noise spectrum having a break frequency of 2 rad/sec, as described in Appendix A. Input parameters shown in Table A-3 were used for this experiment. The input was applied as a vehicle disturbance, essentially in parallel with the pilot's control as shown

in Fig. 14. The input variance was chosen to provide an rms error for the single-indicator display situation that was approximately equal to the rms errors obtained in previous experiments (between 0.25 and 0.5 degrees visual arc).

Subjects were provided mixed training on the various display conditions throughout the training period, and a balanced experimental design was used. Half of the training and data trials were conducted with y(t) displayed. The subjects were instructed to minimize the mean-squared system error signal x(t) at all times. They were not told how to use the auxiliary signal y(t), nor were they assured that the latter signal would necessarily aid in their control of the signal.

Since the subjects who had participated in the first experimental program were transferred out of our locality during the interval between the two experimental programs, a new set of four subjects were used for this program. Two of the subjects (JM and KG) were helicopter pilots currently active in the Naval Reserves; the remaining two (WM and WR) were commercial instructors. All pilots had instrument ratings, and none (except for JM) had participated in previous experimental programs with us.

EXPERIMENTAL RESULTS

The results of the single-axis, multivariable experiments are presented in this section. Average results of four subjects are given here; performance measures for individual subjects are in Appendix D. A subsequent experiment, conducted to confirm the very low observation noise ratios obtained in this experiment, is discussed in Appendix C.

Validation of Interference Model

Model parameters were selected to provide a "good match" (as defined in the preceding section of this report) to the experimental data obtained with the single-indicator display. The relative cost weighting on control rate was found to be 7.5×10^{-5}, which produced a lag time constant of about 0.093 seconds in the controller's describing function. Controller time delay was 0.20 sec, and the observation and motor noise ratios were, respectively, -26.0 and -29.5 dB.* An observation noise ratio of -26 dB was thus associated with full pilot capacity for this experiment.

Predicted and measured human controller describing functions and normalized observation noise spectra for the X-only display condition are shown in Fig. 16. (Since the subject's primary perceptual task was to estimate indicator velocity, remnant was reflected to a scalar noise process injected onto error rate and normalized with respect to error-rate variance.) Except for the measurements at the ends of the frequency range shown, the measured and predicted amplitude ratios and noise spectra agreed to within about 1 dB, and the phase-shift was matched to within 10 degrees over most of this range.

Once the model was calibrated in this manner, we then tested it against two hypotheses: (a) no interference, and (b) full interference between the perceptual tasks of estimating $x(t)$ and $y(t)$. In terms of model parameters, "no interference" implied observation noise ratios of -26 dB associated with the variables $x(t)$, $\dot{x}(t)$, $y(t)$, and $\dot{y}(t)$; other model parameters were held constant. The

*Note that motor noise was not assumed to be negligible. We found that a substantial motor noise level was needed to match the data obtained in one of the subsequent experiments (see Appendix C). Since we do not yet fully understand the relation between motor noise and the various parameters of the control situation, we selected the minimum motor noise ratio that would provide a good match to all the conditions studied in this experimental program. Once chosen, this noise ratio was held fixed for all conditions.

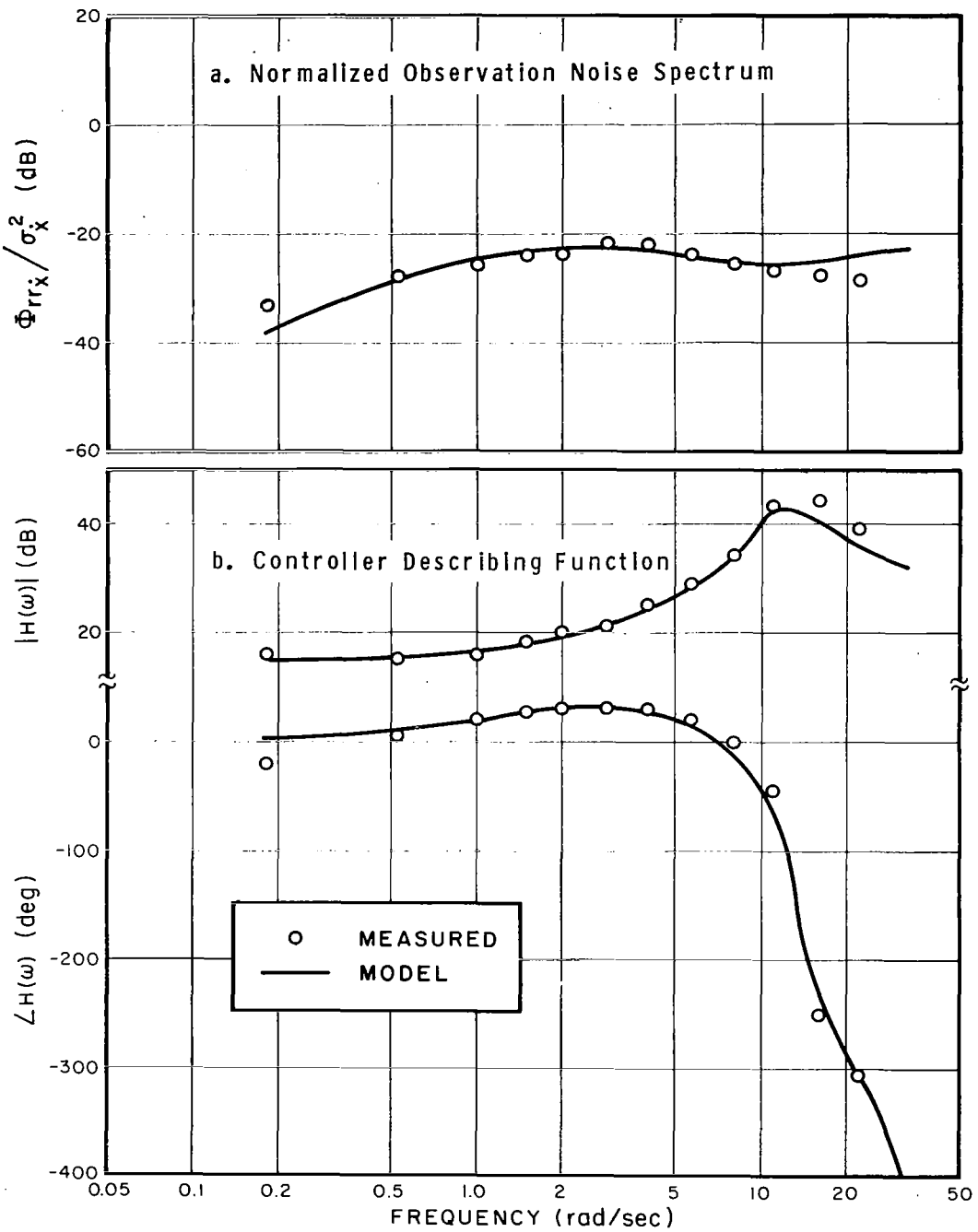


FIG.16 COMPARISON OF MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES: X INDICATOR ONLY

Average of 4 subjects, 4 trials/subject

hypothesis of full interference between the X and Y indicators implied that the noise ratios associated with the four perceived variables should increase so that the fixed capacity is shared between the two indicators. The allocation of capacity which minimized the system error variance was found and performance predictions were based on the associated observation noise ratios. (In order to minimize computational effort, capacity allocation was quantized to integral multiples of 0.1 per indicator.)

Observation noise ratios corresponding to the no-interference and interference hypotheses are given in Table 6. Also shown in this table is the predicted distribution of pilot capacity: 70% on the Y indicator, and only 30% on X.

Predicted and measured average variance scores are given in Table 7. (Scores for the individual subjects are shown in Table D-15 of Appendix D.) Variance scores are tabulated for system error [σ_x^2], system error rate [$\sigma_{\dot{x}}^2$], the optionally-displayed signal [σ_y^2], control effort [σ_u^2], and the rate-of-change of control effort [$\sigma_{\dot{u}}^2$]. Also shown are the ratios of 2-indicator score to 1-indicator score. Since the average scores obtained with the Y-indicator blanked differed by less than 5% from the scores obtained with the Y-indicator serving as a zero reference, scores from these two conditions have been combined into the 1-indicator category.

Performance was significantly improved by the explicit display of the signal y(t). Table 7a shows that the display-related scores were between 60% and 70% of the corresponding scores achieved with a display of only x(t). The control-related scores changed to a lesser degree: control variance was reduced to about 85% of its 1-indicator value, and the control-rate variance decreased by only

TABLE 6

Effect of Display Conditions on Observation Noise Ratios

Indicators	Assumption	Derivation	Observation Noise Ratios		Distribution of Capacity	
			P_x	P_y	X	Y
X	---	matched	-26.0	---	1.0	---
X,Y	no interference	predicted	-26.0	-26.0	1.0	1.0
X,Y	full interference	predicted	-20.8	-24.5	0.3	0.7
X,Y	full interference	matched	-23.8	-22.0	0.6	0.4

TABLE 7

Effect of Display Conditions on Average Performance Scores

Indicators	σ_x^2 deg^2	$\sigma_{\dot{x}}^2$ $(deg/sec)^2$	σ_y^2 $(deg)^2$	σ_u^2 $newton^2$	$\sigma_{\dot{u}}^2$ $(new./sec)^2$	Fraction of Capacity	
						x	y
a. Experimental Data							
x	.23	1.8	.33	49	2.0×10^3	--	--
x,y	.14	1.3	.23	41	1.8×10^3	--	--
Ratio	.62	.69	.69	.84	.94	--	--
b. Model: No Interference							
x	.24	2.0	.35	48	2.2×10^3	1.0	--
x,y	.097	.86	.15	26	1.1×10^3	1.0	1.0
Ratio	.40	.44	.43	.55	.51	--	--
c. Model: Full Interference (Predicted)							
x	.24	2.0	.35	48	2.2×10^3	1.0	--
x,y	.16	1.2	.22	33	1.4×10^3	0.3	0.7
Ratio	.64	.61	.61	.68	.65	--	--
d. Model: Full Interference (Matched)							
x	.24	2.0	.35	48	2.2×10^3	1.0	--
x,y	.17	1.4	.24	36	1.6×10^3	0.6	0.4
Ratio	.69	.69	.69	.75	.73	--	--

Geometric averages of 4 subjects, 4 trials/subject

about 6%. An analysis of variance of the system error score (Table D-16 of the Appendix D) indicates that: (a) the degree of improvement associated with the multiple display was, on the average, statistically significant, and (b) the amount of improvement varied significantly among the subjects.

Comparison of Tables 7b and 7c shows that the predictions based on the full-interference hypothesis were considerably more accurate than those based on the no-interference hypothesis. The reduction of the σ_x^2 score which accompanied the addition of the Y display indicator was predicted to within 5% by the interference model, and the predicted reductions of $\sigma_{\dot{x}}^2$ and σ_y^2 were in error by only about 12%. The no-interference model, on the other hand, predicted display-related scores for the 2-indicator task that were fully a third lower than those observed experimentally. The interference model also yielded better predictions of the control and control-rate scores, although these agreed less well with the experimental results than did the predictions of the display-related scores.

In Fig. 17 the describing functions and noise spectra predicted by the interference model for the two display conditions are shown along with the experimental measurements. The corresponding noise ratios are listed in the first and third rows of Table 6, respectively. The model predicts a rather substantial change in the normalized observation noise spectrum as the Y display is added; the spectrum is shown to increase by 4 to 5 dB at low and mid frequencies, whereas the trend is reversed at high frequencies. Smaller changes are predicted for the describing function: the amplitude ratio (AR) corresponding to the 2-indicator condition is about 2 dB greater than that predicted for 1 indicator at low frequencies, the AR's coincide at frequencies near gain-crossover (around 5 rad/sec),

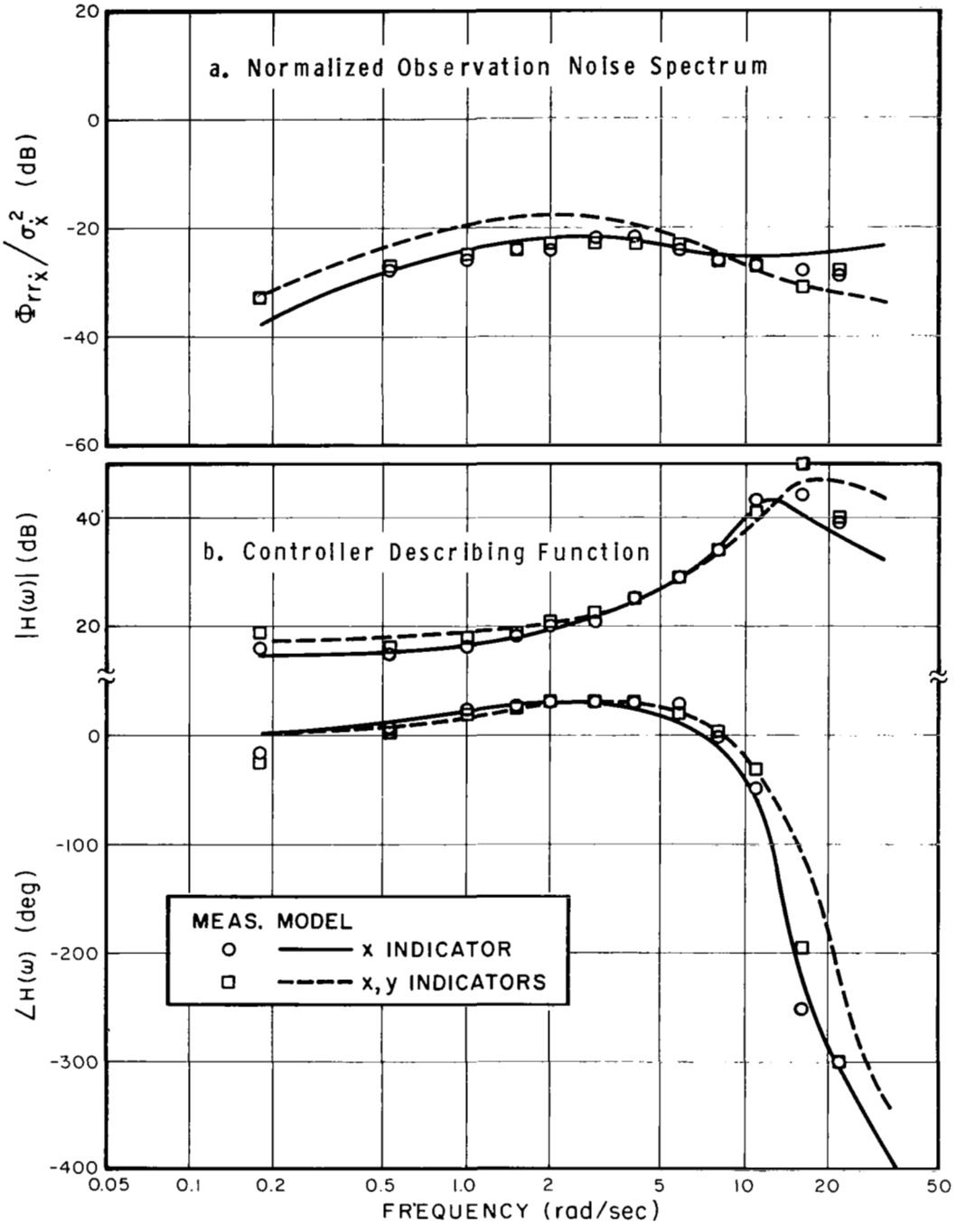


FIG.17 EFFECT OF DISPLAY ON MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES

Average of 4 subjects, 4 trials/subject

and a greater high-frequency peak is predicted for the 2-indicator curve. The model predicts negligible change of phase shift at low frequencies, whereas a substantial reduction of phase lag is predicted at frequencies above gain-crossover.

The prediction that the normalized observation noise spectrum will, on the average, *increase* when the Y indicator is added may come initially as a surprise, since the augmented display should provide better (and therefore less noisy) viewing conditions. Part of this increase is due to the normalization process, since the predicted error-rate variance is lower for the 2-indicator condition. At low frequencies the non-normalized observation noise spectrum for the 2-indicator display is still about 2 dB higher than that for the 1-indicator display, however.

Power spectra of the portion of system error resulting from controller remnant are shown for the two displays in Fig. 18a. Note that the 2-indicator spectrum was higher at low frequencies than the 1-indicator spectrum, a result consistent with the observation noise measurements. But the spectrum of the input-correlated portion of the error, presented in Fig. 18b, is lower for the 2-indicator display at low frequencies. Thus, the reduction of system error variance associated with the 2-indicator display results from the reduction of the input-correlated portion of the error spectrum, which in turn is produced by the higher controller amplitude ratio at low frequencies (Fig. 17).

The predicted effects of display conditions on the frequency-domain measurements are only partially supported by the experimental results. Figure 17 shows that the addition of the Y indicator produced no consistent or appreciable change in the measured normalized observation noise spectrum, as opposed to the appreciable change

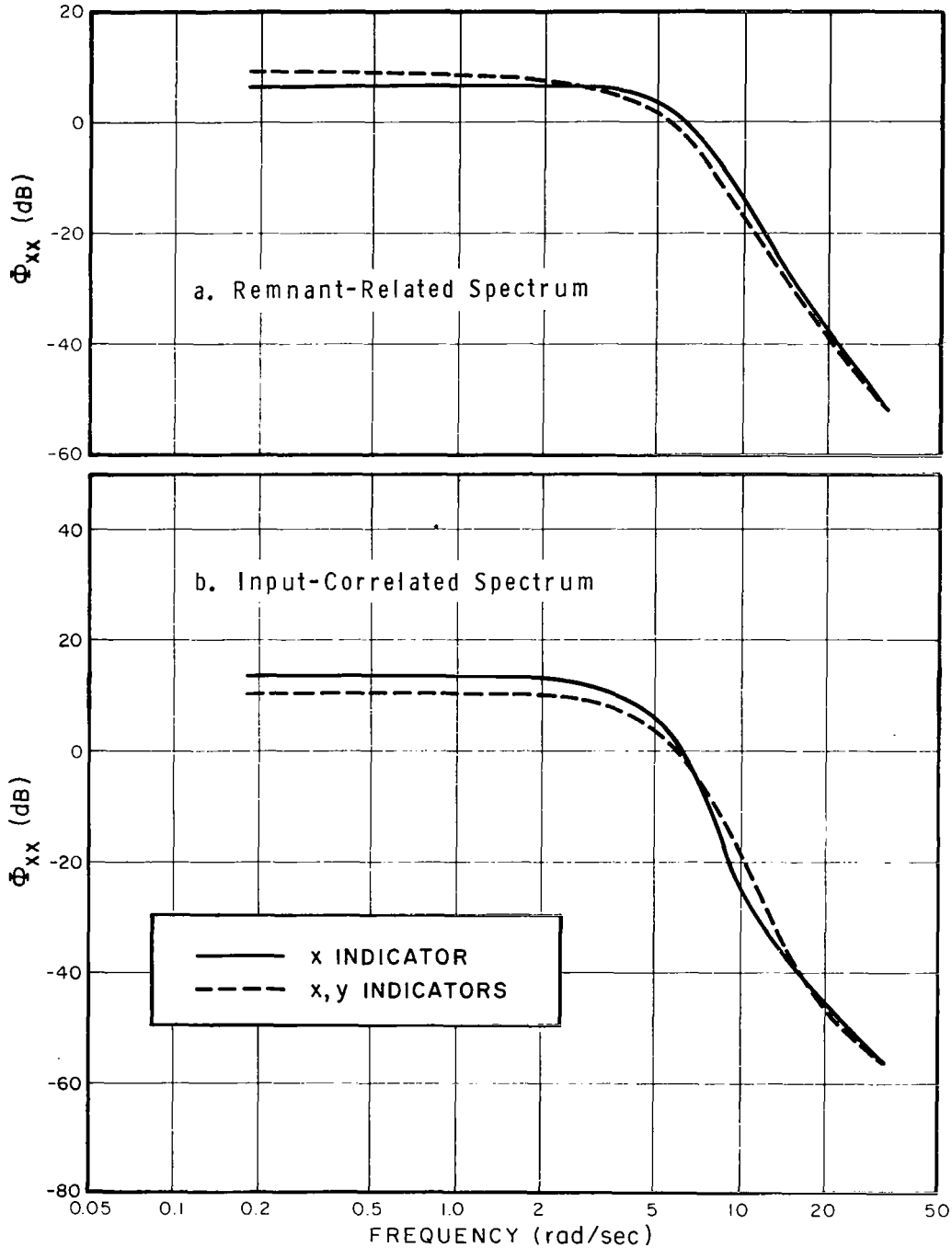


FIG.18 EFFECT OF DISPLAY ON PREDICTED ERROR SPECTRUM

predicted by the model. The controller describing function, on the other hand, was affected by the addition of the Y indicator in approximately the manner predicted by the model; the amplitude ratio increased at low frequencies, had a higher peak value at high frequencies, and was unchanged in the region of gain crossover; the phase lag was reduced somewhat at high frequencies.

The appreciable discrepancy between the predicted observation noise differences and the lack of experimental verification requires further comment. We suspect that this discrepancy may be a manifestation of the subjects' tendency to be slightly nonoptimal in their behavior so long as they achieve essentially the same overall system performance as they would if they had been truly optimal. The effect of distribution of capacity on predicted error variance [σ_x^2] shown in Fig. 19, for example, indicates a maximum variation of about 10 percent as the predicted fractional allocation of capacity on the X indicator is varied from 0.2 to 0.6. Since the subjects were required to devote full attention to the X indicator for half their training runs, the mixed training procedure that we used may have induced the subjects to devote as much capacity as was feasible to the X indicator when both indicators were displayed. This behavior would be consistent with the multi-axis behavior that we noted in our earlier 4-axis experiments; namely, that the subjects apparently allocated more of their capacity to the more familiar foveal display than would have been optimal without suffering much degradation in the total error score.

In an attempt to achieve a better match to the behavior of the subjects for the 2-indicator conditions, we assumed a capacity distribution of 0.6 on the X indicator and 0.4 on Y. The corresponding observation noise ratios are shown in the bottom row of Table 6.

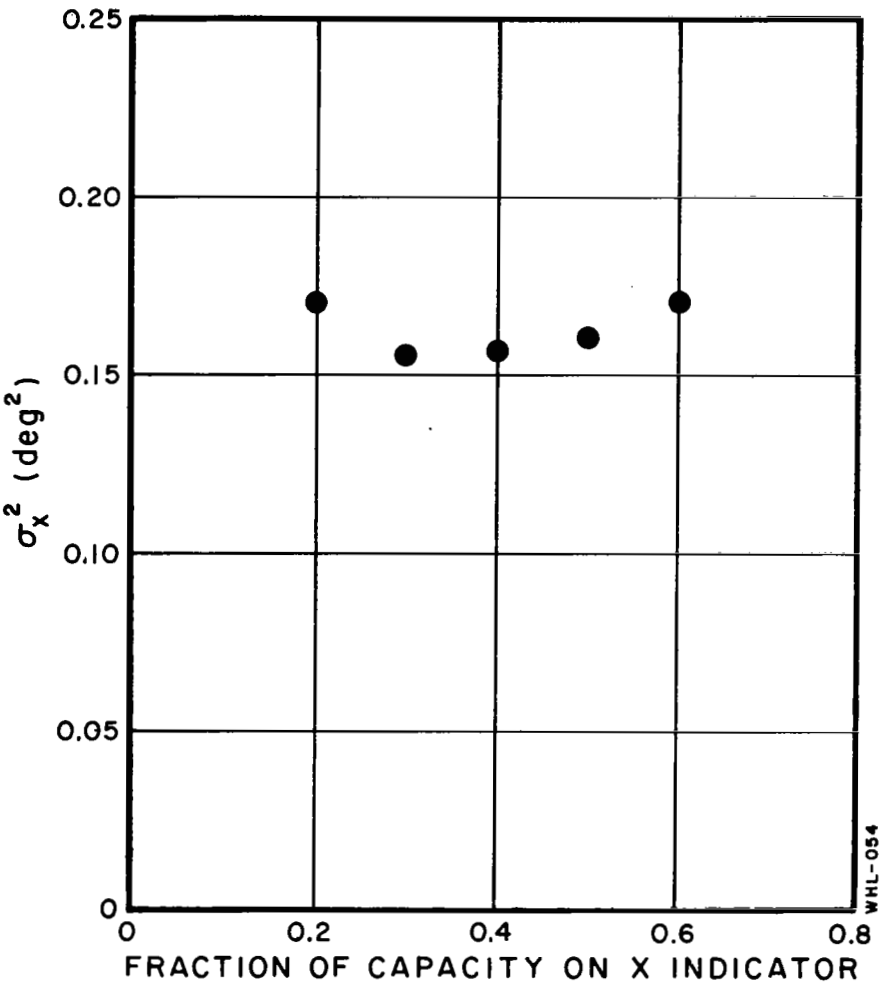


FIG.19 EFFECT OF DISTRIBUTION OF PILOT CAPACITY ON PREDICTED ERROR VARIANCE SCORE

Table 7d shows that an additional 10% error was introduced in the predicted σ_x^2 score, but better predictions were obtained for the control-related scores. Figure 20 shows that the frequency-domain measurements were matched more closely; in particular, the discrepancy between model and measured normalized observation noise spectrum was halved.

Subjective Responses

In an attempt to determine a correlation between the pilots' subjective impressions of task difficulty and objective measures of system performance, we required each subject to respond, in writing, to a questionnaire. When requested, in separate questions, to identify the display condition that was "easiest", "most preferable", and "best for doing the job", the four subjects unanimously chose the 2-indicator display. Although it is possible that the "subjective" opinions were partly based on the knowledge that lower error scores were achieved when the Y indicator was included in the display, several of the comments suggest that truly subjective impressions entered into their rating of the tasks. For example, two of the subjects commented to the effect that the Y indicator enabled them to predict, to some extent, the movements of the X indicator (which was correct, since the position and rate of the Y indicator provided information, respectively, about the first and second derivatives of the X indicator). One subject volunteered the comment that more concentration was required for the X-only display because of its decreased informational content. On the contrary, another subject felt that he worked harder when provided with *two* indicators precisely because *more* information was available. The fact that all subjects gave the 2-indicator task the most favorable rating – despite inconsistencies in their reasons for so doing – correlates favorably with the notions of pilot workload that we present in the next section of this report.

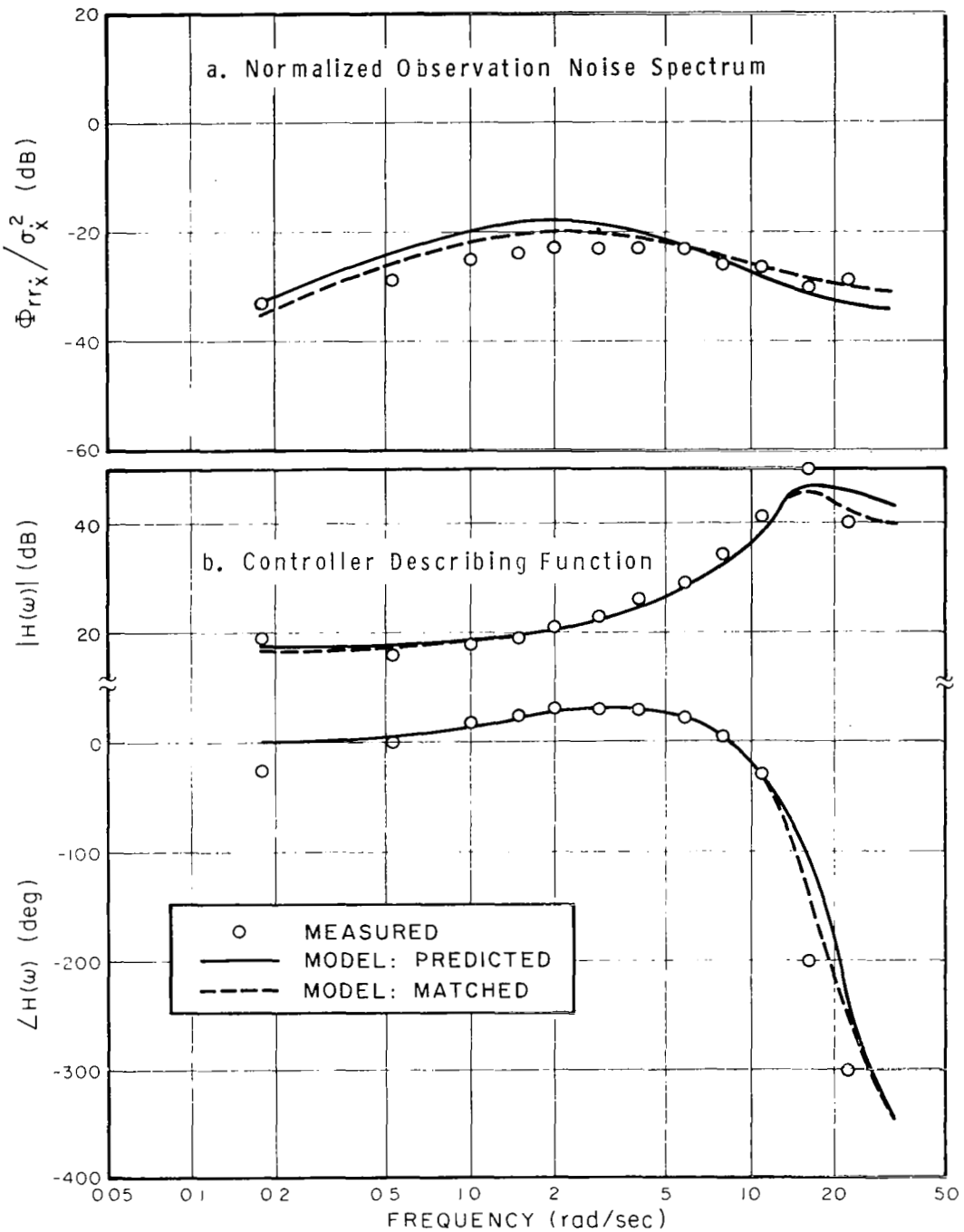


FIG.20 COMPARISON OF MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES: X,Y INDICATORS

Average of 4 subjects, 4 trials/subject

Validation of Observation Noise Results

Although this experiment provided further validation of our model for task interference, we obtained one experimental result that was unexpected; namely, the very low observation noise ratio associated with "full capacity." The 1-indicator "calibration" experiment yielded a ratio of -26 dB, as compared with the -20 dB level we have observed in all previous attempts at matching model output to experimental data. Accordingly, a subsequent experiment was performed to investigate the cause of this unexpectedly low noise ratio. This experiment is discussed in Appendix C; the results are summarized briefly below.

In order to determine whether or not observation noise ratio was related in a consistent manner to the nature of the task, we required the subjects to perform three single-axis tracking tasks in alternation: (a) the unstable-vehicle, $K/s(s-\lambda)$, task investigated in the experiment just discussed, (b) a K/s^2 tracking task of the type investigated in earlier studies, and (c) an additional K/s^2 task. We found the noise ratios to be on the order of -25 dB for all three tasks. On the other hand, the optimal-control model indicated that the error score was most sensitive to observation noise ratio when the vehicle dynamics were unstable and least sensitive when the previously-investigated K/s^2 task was simulated. We conclude from these observations that training with the unstable vehicle dynamics motivated the subjects to achieve a lower observation noise ratio than they would have achieved had they trained only with stable dynamics. Having trained thusly, they maintained the low noise ratio when the dynamics were more stable.

SUMMARY AND DISCUSSION

An experiment was designed to enhance our understanding of central-processing sources of task interference in multivariable manual control situations. Four subjects were required to perform a single-axis control task with either one or two system variables shown explicitly on a single display. Comparison of model predictions with experimental results indicated that a significant amount of interference was present when two indicators were shown, and that this interference could be accounted for by assuming that a fixed amount of "pilot capacity" had to be shared between the two indicators. Although a straightforward application of our optimal control model enabled us to predict the total performance measure very well, detailed predictions of pilot behavior were less accurate. By assuming a slightly nonoptimal allocation of attention on the part of the pilot, however, we were able to achieve an acceptable match to the 2-indicator experimental data without violating the notion of a fixed capacity.

Perhaps the most important result of this study was that our model for task interference correctly predicted the benefits of adding the second indicator to the display; that is, it predicted that the improvement in performance allowed by the additional information would more than offset the performance degradation caused by interference between display indicators. Accordingly, we feel that the model for interference may prove useful in the design of complex displays for real flight-control situations, particularly with regard to determining the point at which increased display complexity no longer aids the pilot.

5. A METRIC FOR PILOT WORKLOAD

In this section we define a metric for pilot workload and verify the reasonableness of this definition by a comparison of model predictions with pilot-opinion data and with other relevant experimental results reported in the literature.

DEFINITION OF THE WORKLOAD INDEX

The model of task interference discussed in Section 2 lends itself straightforwardly to the prediction of the amount of "workload" associated with a given task. The assumption that the pilot possesses a fixed amount of channel capacity to be shared among his tasks forms the very basis for the concept of workload. If capacity were not limited in some manner, the pilot would be able to perform an unlimited number of tasks with no performance degradation; hence, there would be no "loading" of one task upon another. It is clear, however, that the human is limited in the amount of work (either physical or mental) he can perform.

We define the "workload index" (WI) as the fraction of the controller's capacity that is required for him to perform a given task to some specified, or criterion, level of performance. This definition is consistent with those suggested by others in the past. Senders, for example, has defined the workload of a visual monitoring task as the fraction of time required to fixate the associated instrument and has attempted to employ information-theoretic concepts to enable him to predict workload (Refs. 12 and 13). The primary difference between our current definition of workload and those suggested previously is the distinction we now draw between distribution of total channel capacity (which may be thought of as an internal "attentional-sharing" mechanism) and overt visual

"attention" as defined by the pilot's fixation point. We have postulated a model of "central-processing workload," as distinct from "sensory workload" effects that arise when not all displays can be viewed foveally, or "motor workload" effects that would be present if the pilot had to time-share his hands among several manipulators. We are focussing upon central-processing workload because we wish to understand and predict interference effects between continuous control tasks and other kinds of flight tasks (such as conversation or auditory monitoring) that do not require the pilot to divert foveal attention from the flight displays. Furthermore, as head-up and multi-element displays are refined and put into practice, we suspect that central sources of task interference will become more important.

The workload metric suggested here – the fraction of capacity required for adequate performance – is one that we can predict quantitatively with the current implementation of our optimal-control model for human behavior. The procedure is similar to that used in Section 3 for predicting task interference: once the model is "calibrated" for single-axis behavior (either by doing a simple experiment or by using nominal values of parameters that have been found to match previous data), a curve of performance score versus observation noise ratio is obtained. By relating the observation noise ratio to fraction of capacity, as discussed in the preceding progress reports, a quantitative value of workload may be obtained. In terms of our model for the human controller, we may define the workload index alternately as

$$WI = P_o/P_c \tag{17}$$

where P_o is the observation noise ratio corresponding to the pilot's "full capacity," and P_c is the maximum noise ratio that allows the

pilot to achieve the required level of system performance. The noise level P_o reflects not only the pilot's inherent limitations, but also his state of training.

WORKLOAD INDEX AS A FUNCTION OF VEHICLE INSTABILITY

Jex (Ref. 14) and McDonnell (Ref. 15) have reported manual control experiments using vehicle dynamics of the form $K_c/(s-\lambda)$ as a secondary tracking task. They found that main-task performance and pilot opinion degraded as λ was increased. They concluded that the attentional demand of the secondary task increased as the instability increased, with the result that a decreasing fraction of pilot capacity remained for the primary task. If the connection between the unstable pole location and "attention" is correct, which we think it basically is, then our model should predict that the WI increases with increasing λ. Accordingly, in order to verify the reasonableness of our definition of workload index, we have determined the relationship between predicted WI and vehicle instability.

Since the WI is highly task-dependent, we had to specify not only the vehicle dynamics but the input spectrum, performance functional, and the criterion level of performance. The model was tested using a simulated first-order input spectrum having a break frequency of 2 rad/sec. The input was applied as a vehicle disturbance (i.e., in parallel with the control signal). The cost functional was of the form $J = \sigma_x^2 + G \cdot \sigma_{\dot{u}}^2$ where σ_x^2 is the variance of system error and $\sigma_{\dot{u}}^2$ the variance of control rate.

Using nominal values of parameters that we have found to be representative of single-axis manual tracking performance with first-order vehicle dynamics, we adopted -20 dB as the observation

noise ratio corresponding to full capacity,* fixed the effective time delay at 0.17 sec., and adjusted the relative weighting on control rate [G] to yield an effective lag time constant of approximately 0.1 sec. Motor noise was made negligibly small (on the order of -50 dB relative to the variance of the control signal). Curves of performance score versus observation noise ratio were obtained for λ's of 0.0, 0.5, 1.0, and 1.5 rad/sec. These curves, normalized to yield a performance score of unity for the 1/s, full-capacity condition, are shown in Fig. 21. For ease of interpretation we have shown the abscissa as the fraction of pilot capacity, given as P_o/P where P_o is the nominal full-capacity ratio of -20 dB and P is the actual noise ratio used in the model.

The relationship between WI and vehicle instability depends upon the particular criterion level of performance chosen: the more stringent the criterion, the greater will be the workload index. The WI can be obtained from Fig. 21 by noting the intersection of the performance-versus-noise-ratio curve with the criterion performance level.

The relationship between workload index and vehicle instability is shown in Fig. 22 for various values of criterion performance level. The curves show the expected trends. The WI increases as the instability increases, and, for a given set of vehicle dynamics, the WI increases as the criterion score is reduced. Figure 22 shows the trade-off between criterion and vehicle instability with respect to workload index. For example, a WI of 0.2 is predicted for a vehicle pole of about 0.65 rad/sec and a criterion performance level of 5, or for a vehicle pole of about 1.1 rad/sec and a criterion level of 10. In other words, a constant level of subjective

*This theoretical study was completed before we discovered that a substantially lower noise ratio could be obtained by training the subjects with unstable vehicle dynamics. Nevertheless, we suspect the trends shown in this section are relatively independent of the particular choice of minimum noise ratio.

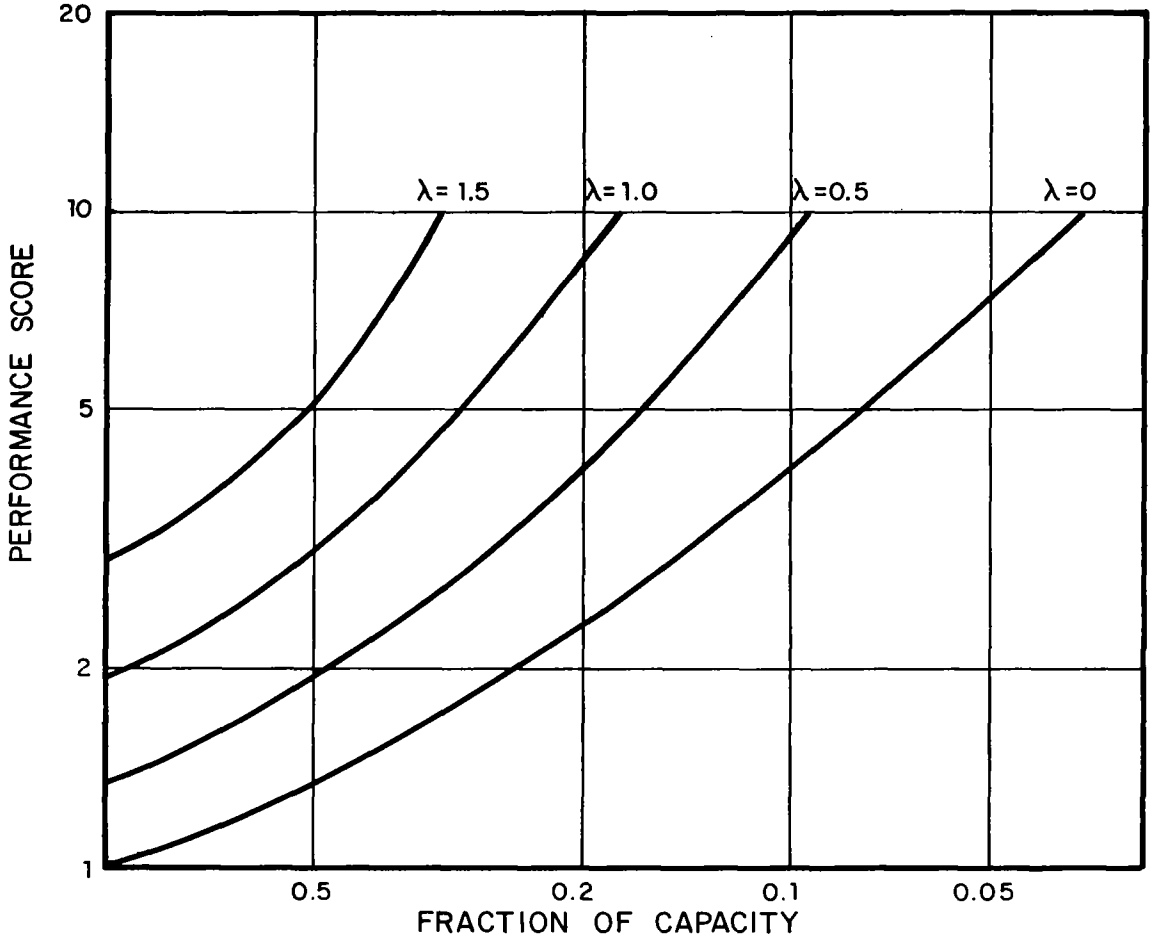


FIG.21 PREDICTED RELATION BETWEEN PERFORMANCE SCORE AND FRACTION OF PILOT CAPACITY

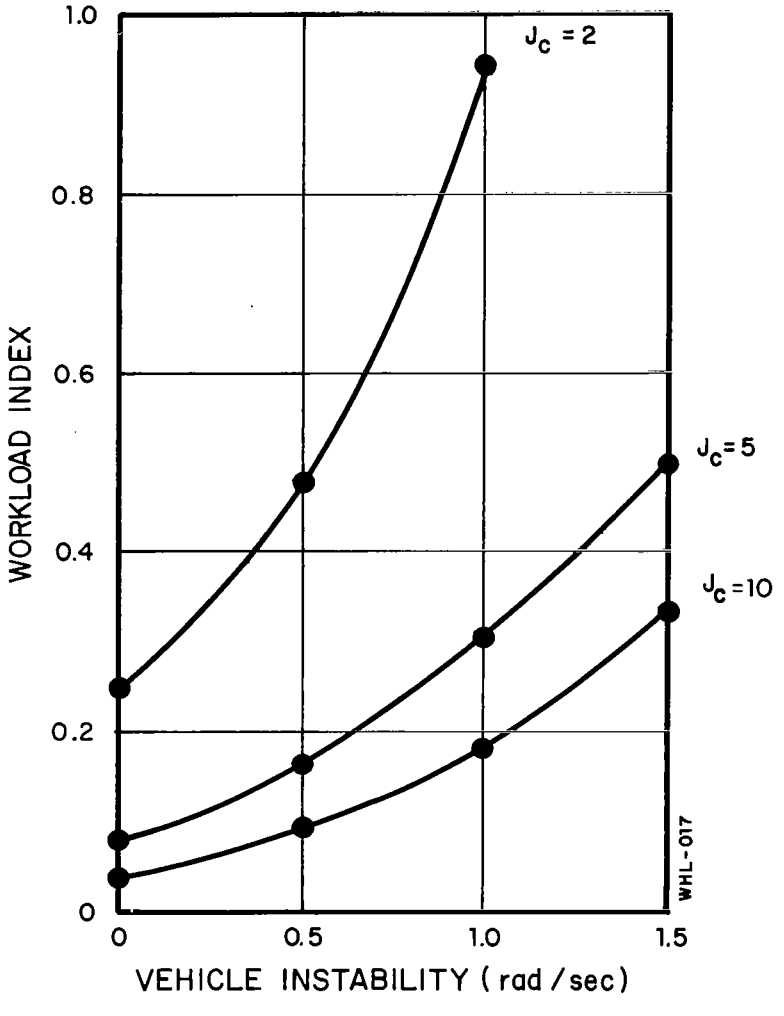


FIG.22 WORKLOAD INDEX AS A FUNCTION OF VEHICLE INSTABILITY

task difficulty should be obtainable if the required level of performance is relaxed as the vehicle dynamics become correspondingly more difficult.

The *absolute* value of the workload index is dependent not only upon the vehicle dynamics and the criterion performance level, but also upon the input spectrum, the cost functional, and the particular values assigned to controller time delay and nominal observation noise ratio. Consequently, some degree of arbitrariness will be associated with the WI predicted for a given task. We suspect that the WI will prove to be most useful as a measure of *relative* difficulty when two or more control configurations are considered. For example, Fig. 22 shows that the WI associated with a vehicle pole of 1.0 rad/sec will be about four times the index associated with a straight K/s plant for criterion performance levels ranging from 2 to 10.

Although our definition of the workload index is primarily a definition of *mental* workload, our method of computing the index also includes the effects of *physical* work. Note that the cost functional used in the above computation includes a term related to control activity. An increase in the simulated control effort (as might be produced by a decrease in the vehicle gain K_c) would, for example, shift upwards the performance-versus-noise-ratio curves of Fig. 21. A decrease in control gain would thus be accompanied by an increasing WI. Examples of predicted workload index versus vehicle gain are given below.

COMPARISON OF WORKLOAD INDEX WITH PILOT RATINGS

If the workload index as we have defined it is a measure of the overall difficulty of a specific control task, we would expect to find a high degree of correlation between pilot opinion ratings

obtained in real or simulated flight situations and the WI predicted by our model. McDonnell has recently reported a set of experimental results in which pilot ratings were obtained over a wide set of consistent experimental conditions (Ref. 15). We shall compare WI with the pilot ratings reported in that reference.

McDonnell and others (Ref. 16) have noted strong relationships between pilot opinion and pilot parameters such as pilot gain and pilot equalization. The relationship between rating and effective pilot time delay appears to be less consistent. In addition to these considerations, McDonnell states that pilot opinion is also influenced by the relative ease with which mission requirements can be fulfilled. Consequently, the parameters of the disturbance function and the required level of performance are important considerations.

All of these factors are incorporated into our procedure for computing the workload index. Although we have defined the WI in terms of pilot capacity or attention, the performance-versus-capacity relationship used in computing the WI is itself a function of input and vehicle parameters. Since pilot parameters are also dependent upon input and vehicle parameters, we would expect to predict some correlation between workload index (and, therefore, pilot opinion) and pilot behavior. Note that we do not consider this as a cause and effect relationship. Since the pilot describing function is predicted by the optimal control model, both pilot behavior and WI are model outputs which are based on a common set of underlying factors. We show below that the predicted workload index is effected in the same way as pilot rating when changes occur in: (a) control gain, (b) pilot lead, and (c) system performance.

McDonnell's primary task was a single-axis compensatory tracking task. A random-appearing command input was provided using a sum of sinusoids to simulate a rectangular spectrum with a low-amplitude, high-frequency shelf. Several vehicle dynamics were used, mostly of the form

$$V(s) = \frac{K_c}{s(Ts+1)} \tag{18}$$

where T ranged from 0 (for K/s) to infinity (for K/s^2). Pilot ratings were obtained for various values of input and vehicle parameters.

In order to provide a reasonable approximation to McDonnell's forcing function without requiring excessive computational effort, the model was programmed to simulate a second-order Butterworth input having a break frequency of 1 rad/sec. With one exception, nominal model parameters were chosen as in the previous example: time delay was set at 0.17, "full-capacity" observation noise ratio at -20 dB, and the weighting on control rate was chosen to yield an effective pilot lag of 0.1 sec. The motor noise, however, was not made negligible, but was nominally set at around -27 dB relative to control power.

We cannot, at this stage in our thinking, expect to obtain a very close match between Cooper rating and predicted workload index. We do not yet understand fully how to choose the appropriate cost functional, nor are we sure of the best way to handle motor noise. Since McDonnell apparently did not specify the criterion level of performance for all tasks, our choice of a criterion level is arbitrary. The value of the workload index, of course, depends upon all these parameters. In addition, we do not know the appropriate

functional relation between workload index and Cooper rating; for simplicity, we have assumed a linear relationship with a scale factor that seems to provide the best match on the average. Consequently, we must be content at this time with predicting the trend of the Cooper rating. As we show below, the trends are predicted rather well.

Effect of Control Gain

McDonnell and others have found that there exists a broad range of control gain that the pilot finds acceptable. When the gain is either increased or decreased beyond this range, a less favorable rating is assigned to the system.

In order to reproduce one of the conditions investigated by McDonnell, vehicle dynamics of K/s were simulated. The vehicle gain was first set to unity, and the control rate weighting and motor noise levels were found according to our procedure for choosing "nominal" values. This condition was interpreted as the "optimal gain" condition. The effects of control gain on workload index were then investigated by keeping fixed the absolute (not relative) values of rate weighting and motor noise as vehicle gain was varied. A performance level of 5 times the "full-capacity" performance of the optimal-gain system was chosen as the performance criterion.

Figure 23 compares predicted workload index and Cooper rating as a function of relative control gain. (A relative gain of unity indicates optimum control gain.) The pilot ratings were obtained from Fig. 20 of McDonnell (Ref. 15). Although the WI appears to be a more sensitive function of control gain than does the Cooper rating, the WI and rating exhibit the same trend: both increase numerically as the gain is either increased or decreased with respect to the "optimum gain."

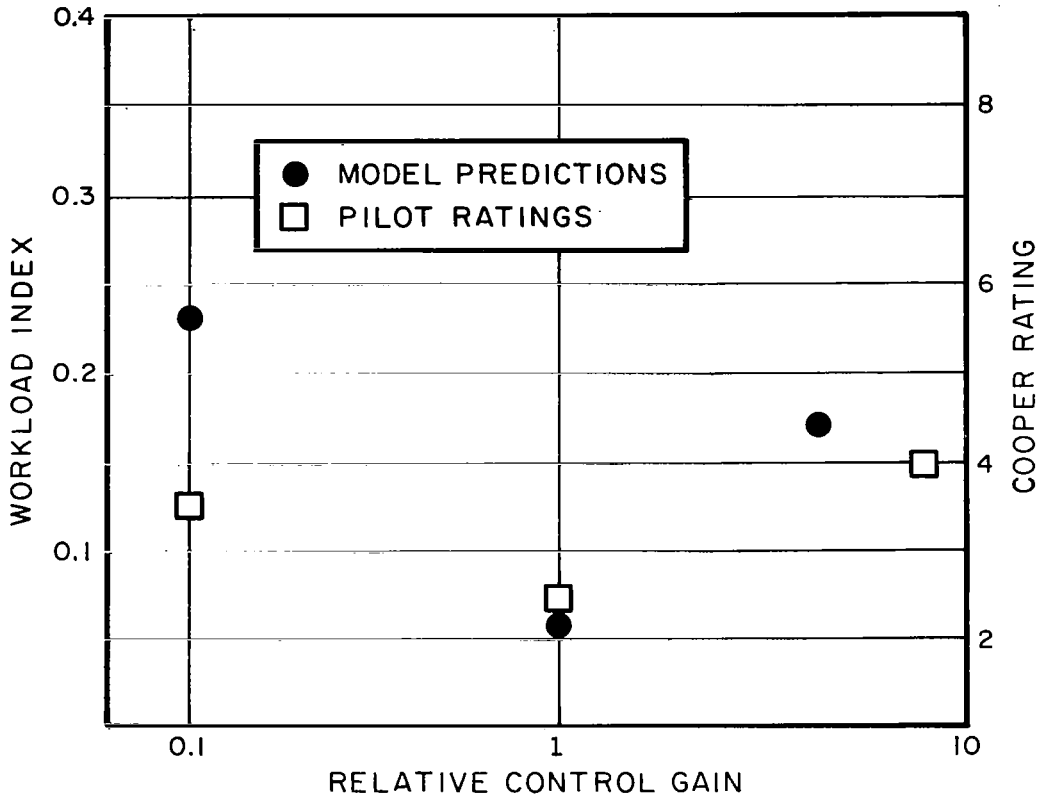


FIG.23 WORKLOAD INDEX AND PILOT RATING VERSUS CONTROL GAIN

Effects of Pilot Lead

McDonnell manipulated the amount of lead generated by the pilot via adjustment of the vehicle lag term T when vehicle dynamics were $K_c/s(Ts+1)$. He found that the pilot's lead coincided with vehicle lag for a lag time constant ranging from 0 to 1 second. For very large vehicle lags (i.e., for vehicle dynamics approaching K/s^2) he estimated the pilot lead time constant as 5 rad/sec. Vehicle gains were adjusted to be "optimal" for each vehicle configuration.

Predicted workload indices were obtained from the optimal-control model using the nominal parameter values described above. A second-order Butterworth noise process having a break frequency of 1 rad/sec was assumed for a command input. A performance level of 5 times the full-capacity performance attainable with the K/s vehicle was chosen as the criterion level.

The relationship between predicted WI and pilot rating versus pilot lead is shown in Fig. 24. (Pilot ratings were obtained from Fig. 22 of Ref. 15.) Rather than attempt to estimate pilot lead by fitting the predicted describing function with the crossover model, we have simply assumed the same relationship between pilot lead and vehicle lag as McDonnell. Figure 24 shows that both WI and pilot ratings increase with increasing pilot lead. The predicted WI shows a larger increment between the two highest pilot lead conditions than does the pilot rating. McDonnell, however, notes that other researchers have found relatively large decrements in rating associated with K/s^2 vehicle dynamics.

Effects of System Performance

McDonnell points out that pilot opinion is partially dependent upon the overall performance of the man-vehicle system. He finds

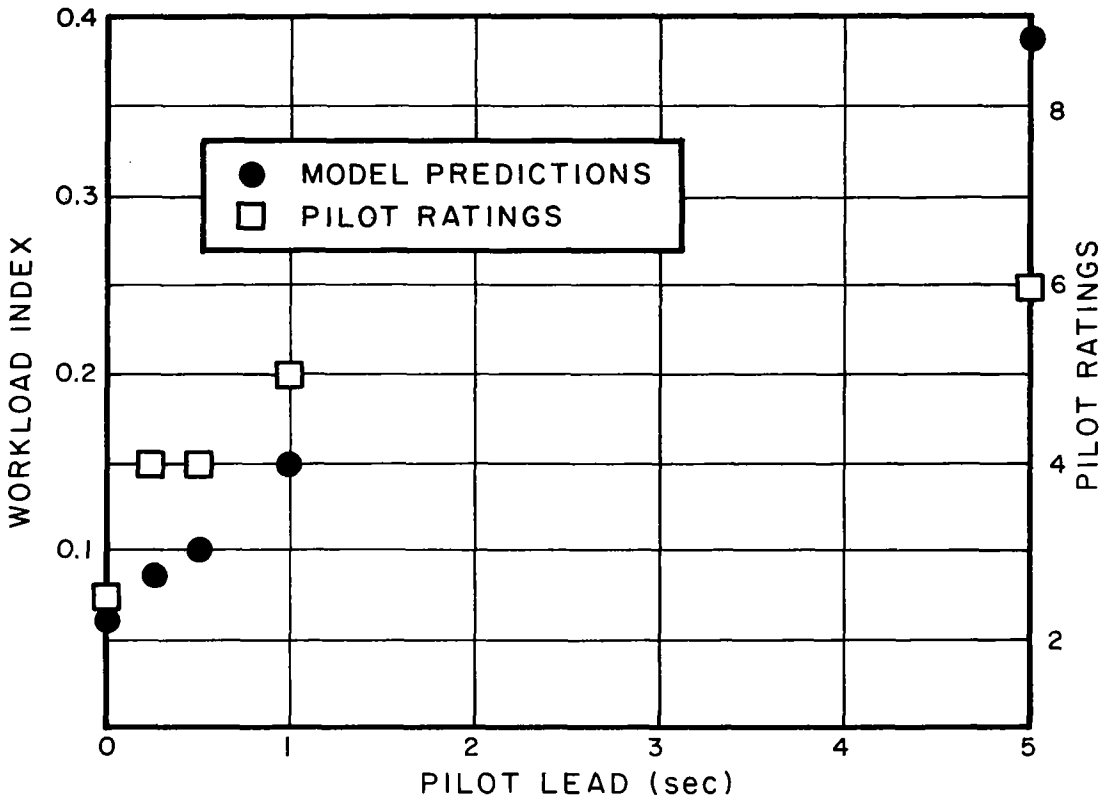


FIG.24 WORKLOAD INDEX AND PILOT RATING VERSUS PILOT LEAD

a reasonably consistent trend between normalized error score (given as the integral absolute error divided by integral absolute input) and pilot rating. These data, obtained from Fig. 25 of Ref. 15, are shown below in Fig. 25. Also shown in that figure is the relationship between predicted WI and normalized error score, which was determined by reinterpreting the model results obtained in the previous example. A normalized model score of $2 \cdot (\sigma_x/\sigma_i)$ was found to provide a good match between predicted WI and pilot ratings as a function of system performance. (Predicted error scores were consistently lower than those achieved by McDonnell, possibly because of input differences or because the experimental subjects may not have been trained sufficiently to reach near-optimal levels of performance.)

SUMMARY AND DISCUSSION

A metric for pilot workload – the "Workload Index" – has been defined as the fraction of pilot capacity (or attention) required to perform a given task to a criterion level of performance. By relating pilot capacity to the equivalent observation noise ratio, we have developed a procedure for predicting the workload index in a given control situation.

We have seen that the predicted workload index increases with vehicle instability (provided other control system parameters remain constant). Since vehicle instability has been shown to correspond to some extent to attentional demands on the pilot, these results support our approach of relating workload to attention. Further analysis has shown that the predicted workload index reproduces the trend of pilot ratings with respect to changes in: (a) control gain, (b) pilot lead, and (c) system performance. These results suggest that the workload index may serve as a useful indicator, or predictor, of aircraft flying qualities, as well as a measure of the demand on the pilot.

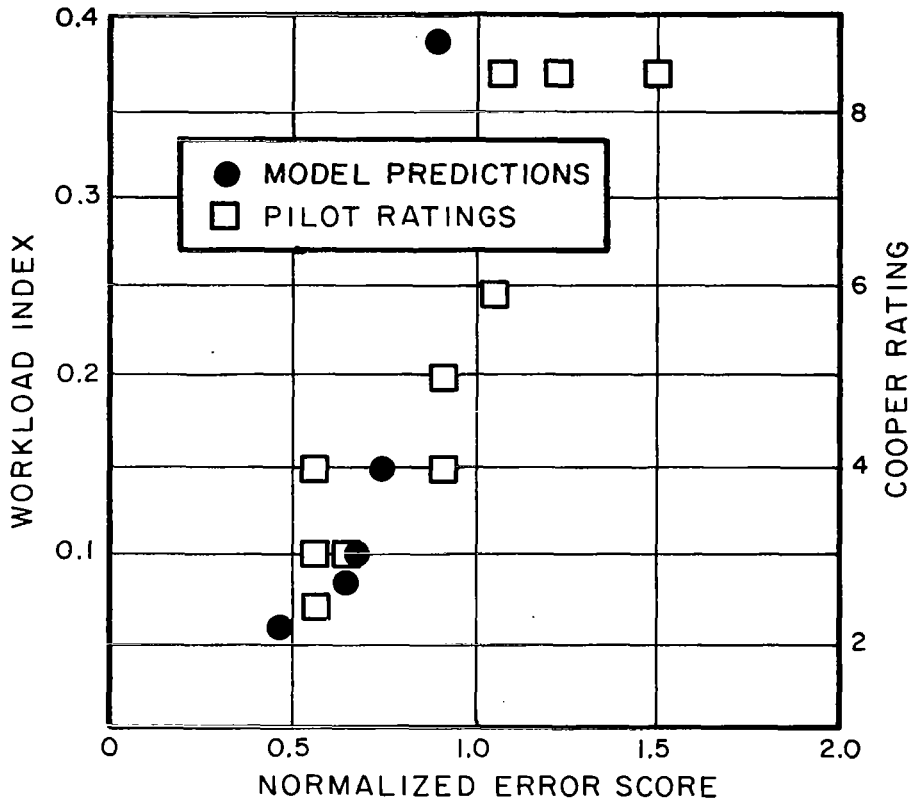


FIG.25 WORKLOAD INDEX AND PILOT RATING VERSUS SYSTEM PERFORMANCE

The workload index, as presently defined, probably will serve best as a measure of *relative*, rather than *absolute*, indication of the demand on the pilot (or of the vehicle handling quality). In order to derive an absolute metric, we would have to know the minimum noise ratio that can be achieved by a pilot under all circumstances. We have seen that this noise ratio depends in part on the way the subject is trained. Furthermore, we suspect that the pilot never applies literally all his mental capacity to a tracking task.

The workload index can be used to predict "trade-offs" among system parameters and between different tasks that leave the workload index unchanged. We have observed such a parameter trade-off in a recent experiment (see Appendix C) in which the subjects, on the average found a K/s^2 and a $K/s(s-1)$ task of about equal difficulty because of the way in which we manipulated the input power. Predictions of this type are possible because the workload index is affected by all of the parameters of the control system. In future studies we expect to investigate the relative workloads imposed by different kinds of tracking tasks and by other types of tasks (such as decision-making) as well.

6. SUMMARY AND CONCLUSIONS

The principal results of this study are the models for task interference and workload that we have developed from our optimal control model of the human controller. The model for task interference has been tested sucessfully against a variety of experimental results described in this report. This model is based on the assumption that multiple tasks are performed in parallel and that a fixed amount of central-processing capacity must be shared among the tasks. The most consistent results are obtained if we associate a separate "task" with each display *indicator* that the pilot uses (not necessarily with each *variable* observed, since we find that there appears to be no interference between estimations of position and rate of a single indicator).

The model for interference can treat display indicators that are linearly related (coupled) as well as uncoupled indicators; i.e., the same rules apply for multi-axis control situations as for single-axis, multi-variable tasks. Central-processing sources of interference do not appear to depend on the number of controls used in the system. There are, however, motor-related sources of interference that may be important when a multi-axis manipulator is controlled by a single limb. (See Appendix B.)

Pilot "capacity" is related to a minimum equivalent observation noise ratio attainable by the subject. This noise ratio apparently depends on the nature of the subject's training experience, as well as on his inherent abilities. Interference can be accounted for by allowing the equivalent observation noise ratio associated with a given indicator to increase in inverse proportion to the fraction of the pilot's capacity allocated to that display. Incorporating the model for task interference into an optimal-control

framework yields a composite model from which reliable predictions of overall system performance can be obtained.

A metric for pilot workload, the "workload index," is defined as the fraction of the controller's capacity that is required to perform a given task to some specified, or criterion, level of performance. A theoretical study of the relation between workload index and system parameters and comparisons of predicted workload index with pilot-opinion data reported in the literature indicate that this is a reasonable metric for workload. Because of the difficulty of obtaining a reliable numeric to associate with "full capacity," we suspect the workload index will serve more reliably as a means for comparing the relative demands on the pilot imposed by various tasks, rather than as a measure of absolute workload.

In addition to these theoretical results, a number of interesting multi-task experiments were performed. These are discussed in detail in appendices and the principal results are summarized below.

Single-axis compensatory tracking was studied under four conditions of foveal and peripheral viewing. Both the location of the display and the orientation of the display indicator were varied. Vehicle dynamics were K/s, and a simulated first-order noise process was applied as a disturbance to vehicle velocity. Tracking performance was degraded as the display was moved from the center of the visual field into the periphery. In addition, peripheral tracking ability was significantly improved when the display indicator was oriented so that the subject could extrapolate a zero reference from his fixation point to the display. Both error variance scores and normalized observation noise spectra increase with increasingly unfavorable viewing conditions. Changes in the noise spectra occurred primarily at low frequencies, which suggests that

perception of indicator position was degraded more than perception of velocity. On the average, perception in the periphery appeared to be better when the displays were located in the lower half of the visual field. Whereas the level of observation noise was roughly proportional to the error variance when the display was viewed foveally, it was almost invariant to the error variance when the display was viewed peripherally. Hence, the relative quality of peripheral perception can be expected to improve as the amplitude of the displayed signal increases.

Experiments were performed in which two, three, and four independent axes were tracked simultaneously. A separate display was provided for each axis of control. Minimum display separation was 16° visual arc. Visual scanning was not permitted: the subjects maintained fixation on one of the displays for the entire experimental trial. All multi-axis tasks showed evidence of task interference (i.e., the error variance score on a given task was greater when multiple tasks were performed than when that task was performed singly). In general, increased error scores were accompanied by increased normalized observation noise spectra and lower human controller gain. The subjects showed no appreciable tendency to favor the foveal task when provided with foveal and peripheral tasks of approximately equal difficulty. Interference was greater when two axes were controlled by one hand operating a two-axis manipulator than when each hand operated a single-axis manipulator.

The subjects were also allowed to scan while controlling the four independent axes. Four separate displays were provided, located at the four corners of an imaginary square. Two experimental conditions were investigated; (a) homogeneous inputs, in which the spectral characteristics and total power levels of the four inputs were identical, and (b) nonhomogeneous inputs, in which the mean-squared input on the axis corresponding to the lower left display

was four times as great as the mean-squared inputs applied to each of the remaining three axes. The subjects fixated the upper two displays most of the time even when the inputs were homogeneous; apparently because of the superior visual information obtained in the lower portion of the peripheral visual field. Scanning behavior was changed negligibly when the input on the lower left axis was increased. The failure of the subjects to fixate the lower left display for a substantially greater fraction of time in this circumstance stemmed, we assume, from the relative improvement of peripheral perception that accompanied a larger error signal on that display.

APPENDIX A

APPENDIX A. APPARATUS AND PROCEDURES

Two experimental programs are described in this report: an extensive study of multi-axis control behavior, and a subsequent (and less extensive) study of interference effects within a single axis of control. The apparatus and procedures described in this appendix were, with some modifications, used throughout the primary experimental program. To some extent, these same procedures were employed in the second program. Table A-1 lists the set of experimental conditions common to most of the experiments. Deviations from these conditions are mentioned specifically in the presentation of experimental results in Appendices B and C.

In the primary experimental program, the subjects were presented with four compensatory tracking displays, each of which contained a stationary reference bar and a moving bar to indicate system error. The error bars were controlled by compatible movements of two 2-axis control sticks. The four axes were linearly independent, and simultaneous control was required for various combinations of 1,2,3, and 4 axes during the course of the experimental program. Figure A-1 shows a linear flow diagram of one of the four axes. Throughout this experimental program we used a state-regulation configuration, as opposed to the command-input configuration previously used by us and other workers (Refs. 1-4) in order to provide compatability with recent theoretical developments (Ref. 6). The forcing function was injected in parallel with the pilot's control action and thus appeared as a disturbance on vehicle velocity.

TABLE A-1

Experimental Conditions Common to Most Experiments

Experimental Variable	Standard Condition
Displays	Four displays arranged on the corners of a square whose sides were 16° visual arc in length. See Figure A-2.
Controls	Two 2-axis controls, with each axis of control affecting a single display.
Controlled-element Dynamics	K/s in all axes. Control gain was such that 1 newton of force produced a display deflection rate of 2 degrees visual arc per second.
Forcing Functions	Sums of sinusoids to simulate a first-order noise process with a break frequency at 2 rad/sec. Similar statistical properties, but different waveforms, on each axis. See Tables A-2 to A-4.

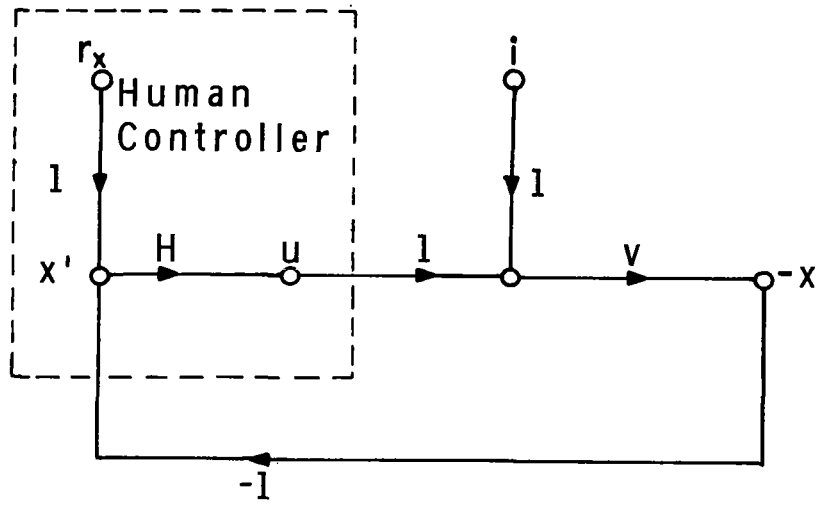


FIG.A-1 LINEAR FLOW DIAGRAM OF A SINGLE-AXIS COMPENSATORY MANUAL CONTROL SYSTEM

System forcing function, system error, and control movement are represented by i,x, and u, respectively. The controller remnant is represented by an equivalent observation noise r_x, and the "perceived error" is shown as x'. H and V represent the human controller's describing function and the vehicle dynamics.

The vehicle output is shown here as -x so that we may adopt the standard practice of indicating negative feedback.

PRINCIPAL EXPERIMENTAL HARDWARE

Computing Machinery

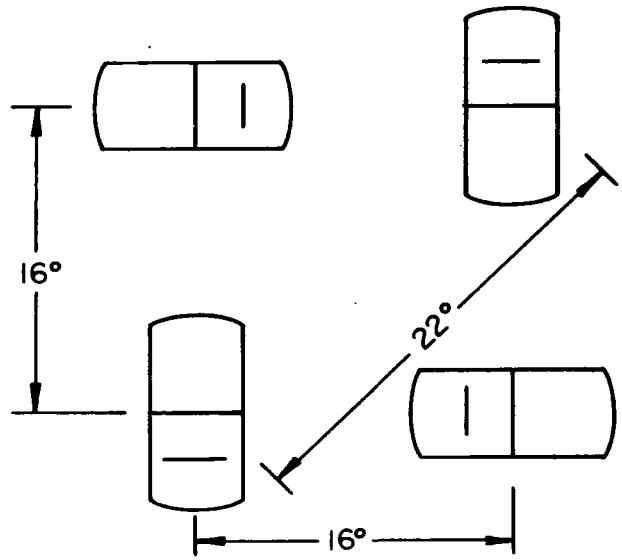
An Applied Dynamics AD/4 Analog Computer was used to simulate vehicle dynamics, drive the displays, and compute mean-squared system errors. An SDS-940 time-shared digital facility was used to generate forcing functions, convert analog data to digital format for storage on magnetic tape, and aid in data analysis.

Subject Booth

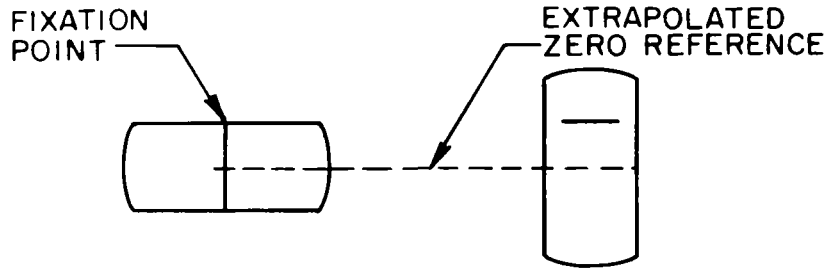
The displays and controls were located in a subject booth that was isolated both acoustically and visually. A chin rest was provided to control the subject's point of regard and to minimize rotational head motions.

Displays

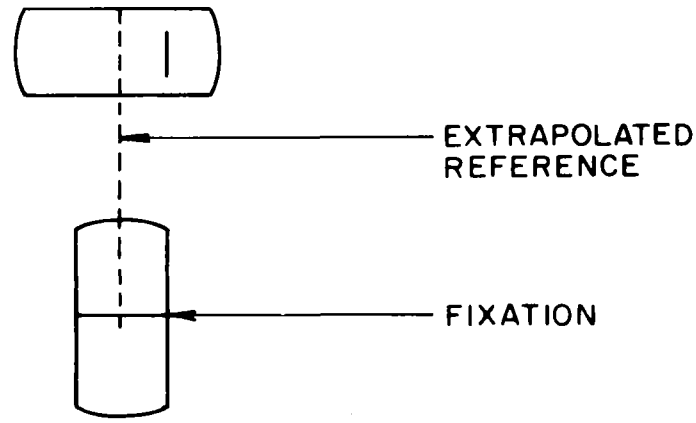
The subject was provided with four oscilloscopic displays arranged on a plane surface located 72 cm in front of the subject's point of regard. The displays were located at the corners of an imaginary square, as shown in the scaled drawing of Figure A-2a. The centers of the displays were separated by about 16 degrees of visual arc along the sides of the square and about 22 degrees along the diagonal. Each scope face was masked with black paper to produce a rectangular background of 5 by 10 cm. The phosphor of the display tube was type P-11, which gave a bluish cast to the reference and error indicators. An overlaid reticle provided a rectangular array of grid lines separated by about 1/2 cm. Intensity levels for the display and background were adjusted to be the same for all displays and were kept the same throughout the experimental program. A constant low level of room lighting was maintained. A typical display presentation (minus the grid lines) is shown in



a. Four-axis Display Configuration



b. Horizontal Reference Extrapolation



c. Vertical Reference Extrapolation

FIG.A-2 DISPLAY CONFIGURATION USED IN THE EXPERIMENTS
Dimensions Shown in Degrees of Visual Arc

Figure A-3; the dimensions shown indicate degrees of visual arc with reference to the subject's point of regard.

Preliminary investigation revealed that peripheral tracking performance was aided by the ability of the subject to extrapolate a zero reference from his fixation point to the peripheral display. This was true apparently because perception of the stationary baseline presented on the peripheral display faded after a few seconds. In order to provide similar viewing conditions for each fixation point, the error indicators of the upper left (UL) and lower right (LR) displays were presented as vertical bars which moved in the horizontal dimension, and the lower left (LL) and upper right (UR) were horizontal bars which moved in the vertical dimension. This arrangement allowed the subject to extrapolate a zero reference only to the peripheral display located in the nearest clockwise position to his fixation point, no matter which of the four displays was designeated as the fixation point. Examples of horizontal and vertical reference extrapolation are shown in Figures A-2b and A-2c, respectively.

Controls

The subject manipulated two aluminum sticks, each of which was attached to a force-sensitive hand control (Measurement Systems Hand Control, Model 435). The stick-control combination provided an omnidirectional spring restraint with a restoring force of about 8×10^5 dynes per centimeter deflection of the tip of the stick. The subject used wrist and finger motions to manipulate the sticks and was provided with arm rests to support his forearms.

The transducer of each hand control provided two independent electrical outputs, one proportional to the horizontal and the other proportional to the vertical component of deflection.

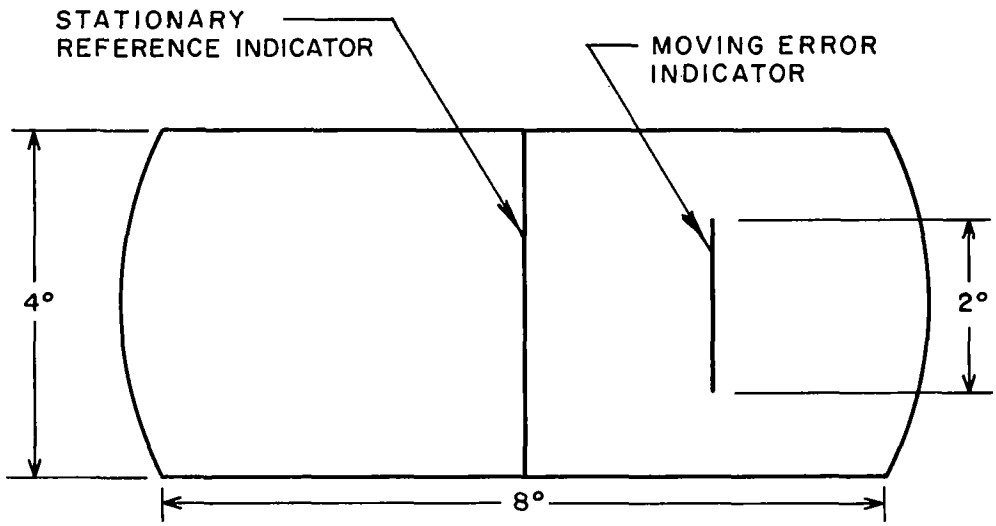


FIG.A-3 TYPICAL DISPLAY PRESENTATION

Dimensions Shown in Degrees of Visual Arc

The sticks were allowed to move freely in both axes in all experiments. The error indicators in the inactive axes were clamped electronically at zero displacement.

In order to provide a high degree of control-display compatibility, each control was oriented so that the stick was horizontal and could be moved in a plane parallel to the plane of the displays. Each display was controlled by a component of stick movement along the same axis as the motion of the error indicator. Thus, the UL display was controlled by x-axis (i.e., horizontally-directed) deflections of the left control stick, the LL display by y-axis motions of the left stick, the UR display by y-axis motions of the right stick, and the LR display by x-axis motions of the right stick. The response of an error indicator was in the same direction as the corresponding component of control deflection.

Oculographic Recording Apparatus

Eye movements were monitored via electro-oculographic techniques (Ref. 17). Voltages proportional to eye position (relative to head position) were detected via Beckman biopotential skin electrodes and preamplified by Electro Instruments Model A20B DC amplifiers. Separate recording systems were used to monitor horizontal and vertical eye movements.

CONTROL SYSTEM PARAMETERS

Controlled-Element Dynamics

The controlled-element dynamics were K/s in all axes for all experiments. In all but one experiment, the control gains were the same on all axes. For most of the experiments, the control gain was such that 1 newton of force produced an error rate of 2 degrees visual arc per second.

Forcing Functions

Forcing-functions were provided via a multichannel FM magnetic tape system during training and were generated by the 940 digital system during data-taking sessions. Up to 13 sinusoids were summed to provide signals that were random-appearing and whose spectra approximated white noise processed by a first-order filter with a pole at 2 rad/sec. In order to assure orthogonality among the component sinusoids, an integral number of cycles of each component was contained in the measurement interval (about 200 seconds). Thus, each component was a harmonic of the fundamental frequency

$$\omega_o = 2\pi/200 = .031 \text{ rad/sec} .$$

For most of the experiments the mean-squared input was about 2.2 $(\text{arc-degrees/sec})^2$ on each axis, where the units refer to the angular velocity of the error indicator that would result if no control action were taken by the subject.

Three sets of forcing functions were used for data taking. Tables A-2 through A-4 show for each sinusoidal component of each forcing function the radian frequency, the number of wavelengths contained in the measurement interval, and the relative amplitude. (The component amplitudes have been normalized to yield a mean-squared signal level of unity.) The initial phase shifts associated with each component of a forcing function were selected from a random process having a uniform distribution between 0 and 2π. The four forcing functions contained in set B (Table A-3) were identical in order to provide tasks on each axis that were as nearly alike as possible. This input set was used only for experiments in which the axes were tracked one at a time. For

TABLE A-2

Parameters of Forcing-Function Set A

AXIS 1			AXIS 2		
No. of Wave-Lengths in Measurement Interval	Frequency (rad/sec)	Relative Amplitude	No. of Wave-Lengths in Measurement Interval	Frequency (rad/sec)	Relative Amplitude
3	.092	.289	4	.12	.354
8	.25	.406	10	.31	.378
18	.55	.463	19	.58	.480
30	.92	.475	33	1.0	.485
45	1.4	.517	47	1.4	.506
68	2.1	.486	72	2.2	.448
90	2.8	.462	86	2.6	.430
134	4.1	.434	126	3.9	.465
178	5.5	.384	190	5.8	.393
266	8.2	.336	258	7.9	.326
358	11.0	.281	370	11.4	.290
518	15.9	.262	530	16.3	.251

TABLE A-3

Parameters of Forcing-Function Set B
(Parameters identical on all axes)

No. of Wavelengths in Measurement Interval	Frequency (rad/sec)	Relative Amplitude
6	.18	.449
17	.52	.497
32	.98	.514
48	1.5	.474
66	2.0	.466
94	2.9	.461
130	4.0	.424
186	5.7	.381
262	8.0	.325
366	11.2	.281
522	16.0	.238
733	22.5	.202
1052	32.3	.162

TABLE A-4

Parameters of Forcing-Function Set C

No. of Wave-Lengths in Measurement Interval	Frequency (rad/sec)	Relative Amplitude	No. of Wave-Lengths in Measurement Interval	Frequency (rad/sec)	Relative Amplitude
AXIS 2			AXIS 1		
6	.18	.449	8	.25	.491
17	.52	.497	18	.55	.456
32	.98	.514	30	.92	.467
48	1.5	.474	45	1.4	.509
66	2.0	.466	68	2.1	.479
94	2.9	.461	90	2.8	.455
130	4.0	.424	134	4.1	.427
186	5.7	.381	178	5.5	.378
262	8.0	.325	266	8.2	.331
366	11.2	.281	358	11.0	.276
522	16.0	.238	518	15.9	.244
733	22.5	.202	737	22.6	.202
1052	32.3	.162	1048	32.2	.162
AXIS 3			AXIS 4		
7	.22	.445	9	.28	.510
16	.49	.494	19	.58	.473
35	1.1	.524	33	1.0	.478
46	1.4	.415	47	1.4	.498
62	1.9	.467	72	2.2	.441
82	2.5	.499	86	2.6	.424
138	4.2	.456	126	3.9	.458
182	5.6	.368	190	5.8	.387
270	8.3	.324	258	7.9	.321
362	11.1	.273	370	11.4	.286
526	16.1	.240	530	16.3	.237
745	22.9	.197	741	22.7	.197
1044	32.0	.159	1040	31.9	.162

the experiments in which two or more axes were tracked simultaneously, it was necessary to provide different inputs on each axis so that the tasks would appear to be linearly uncorrelated. Input set C was designed for this purpose. Although the four forcing functions in this set are composed of the same approximate frequency components, corresponding components do vary slightly so that no two forcing functions contain exactly the same frequencies. This design allows us to test for linear cross-couplings introduced by the subject.

TRAINING AND EXPERIMENTAL PROCEDURES

Subjects

Four subjects, all of them instrument-rated pilots, participated in this experimental program. Three subjects were active in the Air National Guard, and the remaining subject was active as a commercial pilot. Two of the subjects (JF,DM) had participated in previous experimental programs; the remaining subjects were novices with regard to experimental manual control situations. All of the subjects, by virtue of their occupations, were assumed to posess both good eyesight and a native ability for proficient manual control.

Instructions

The subjects were instructed to minimize mean-squared tracking error. When tracking more than one axis simultaneously, they were instructed to minimize a total score given as the sum of the mean-squared error scores obtained from each axis. (The scores were weighted equally in this computation.) The subjects were informed of their scores after each session, and histories of the performance of all subjects were posted and shown to each subject in an attempt to foster a spirit of competition.

Run Length

All training and experimental trials lasted four minutes and were generally presented in sessions of three or four trials each with a minimum rest period of 10 minutes between sessions. Minimum rest periods of 1 minute were provided between successive trials within a session.

Inputs

A number of forcing functions were used during training under a given condition to minimize learning of the input. These forcing functions were of the type shown in Table A-4, except that the highest frequency component was absent. In order to minimize the effects of input differences on the experimental results, however, a single set of forcing functions was used in a given experiment.

Training

The subjects were trained for each experimental condition until their performance levels appeared to be stable. An average of nearly sixty 4-minute trials of training was provided to each subject prior to the first experiment; substantially less training was required for subsequent experiments. In general, the subjects were trained in an equal mixture of the conditions to be investigated in a given experiment.

DATA RECORDING

All experimental data were recorded onto digital magnetic tape via the SDS-940 system and its associated peripheral hardware. Control of the experiment was effected through the STOREDATA system, a program written in a 940-compatible version of FORTRAN II which (1) generated the forcing functions, (2) provided a signal for controlling the analog computer, (3) performed on-line computations of

the incoming data, and (4) converted the data to digital format for storage. Data were sampled at the rate of 20 samples/second. All experimental trials were 4800 samples (4 minutes) in length. All of the analyses described in the following section (except for some of the eye-movement analysis) were performed on 4096 samples (about 3 minutes, 20 seconds) beginning about 20 seconds after the onset of the trial.*

DESCRIPTIVE MEASURES

Mean-Squared Errors

Mean-squared error (MSE) scores were computed for each axis in a given experimental trial, and a total performance measure was computed as the unweighted sum of these scores. Analyses of variance were performed on selected sets of MSE scores to test the significance of differences in scores that accompanied changes in experimental conditions.

Eye-Movement Statistics

The following eye-movement statistics were computed from the multiaxis tracking data.

(a) Fraction of attention: Given as the fraction of time that a given display was fixated.

(b) Fixation frequency: Defined as the number of observations of a given instrument divided by the total run length in seconds.

(c) Mean observation time: Computed as the cumulative fixation time on a given display divided by the number of observations of that display.

*Our implementation of the fast-Fourier transform technique used in obtaining power spectra and describing function requires 2^N data points.

An "observation" was defined as a succession of samples corresponding to foveal fixation of a single display, bounded by samples indicating fixations of other displays. Successive observations of the same display were not defined as such and were treated as a single observation. Measured observation times were necessarily multiples of the analog-to-digital conversion interval of 0.05 second.

The subject's fixation point was determined from recordings of the vertical and horizontal components of eye position by means of two zero-level detectors. The subject was assumed always to fixate one of the four displays. In order to discount the spurious fixations that were often indicated when the subject scanned along a nearly-diagonal path, all measured observations of less than 0.25 second were disregarded. Since the cumulative time accounted for by these short-term observations was only about 1 percent of the total run time, the error introduced by our analysis procedure was well within the expected normal run-to-run variation.

Power Spectra

Power spectra were obtained using Fourier analysis techniques based on the Cooley-Tukey method of computing transforms (Ref. 18). In order to enhance the interpretability of the results, the fundamental frequency component of the Fourier analysis was the same as the base frequency about which the forcing functions were constructed. Each spectrum, therefore, consisted of a set of lines spaced by approximately 0.031 rad/sec and extending from 0.031 to about 64 rad/sec. Measurements beyond 37 rad/sec were disregarded.

It was convenient for analytical purposes to consider each power spectrum as the sum of two component spectra: (a) the "input-correlated" spectrum, consisting only of those measurements coincident with the forcing-function frequencies, and (b) the "remnant" spectrum, consisting of the remainder of the total power spectrum. This interpretation of the measurements was based on the assumption that the remnant was a broadband continuous function of frequency having a relatively low power density level, as compared to the input-correlated portion of the signal which contained a relatively high power density level at a few selected frequencies. Thus, measurements at input frequencies were assumed to represent only the linear response of the system, uncorrupted by the small amount of remnant in the measurement "window". (This assumption was tested for each spectral measurement.)

Computation of the power spectrum allowed the partitioning of the signal variance into the portion of signal power correlated with the input and the portion due to remnant. These component scores were obtained by summing the spectral measurements obtained at input frequencies and at all frequencies excepting input frequencies, respectively. (The power measurement at zero frequency, representing the square of the mean, was not included in the latter summation.)

Estimates of the remnant component of the spectrum at input frequencies were needed for the computation of observation noise (discussed below) and also to provide an estimate of the signal-to-noise ratio at these frequencies. This measurement could not

be obtained directly, since there was no way to subdivide a single measurement into input-related and remnant-related components. Instead, estimates were provided by averages of the power spectral measurements obtained on either side of (but not including) an input frequency. The assumption was made that the remnant spectrum was a continuous function of frequency in the vicinity of the input frequencies. This has been shown to be a reasonably good assumption (Ref. 1). The averaging windows extended roughly 1/8 octave on either side of each input frequency. Since the spectral measurements yielded by the Fourier analysis were spaced linearly with frequency, the number of measurements included in the average increased with increasing frequency – ranging from 1 for the estimate at the lowest frequency to around 180 at the highest frequency.

Describing Functions

Human controller describing functions were obtained using the Fourier analysis techniques described above. Samples of the controller describing function – at input frequencies only – were obtained by dividing the transform of the control signal by the transform of the error signal. This technique is similar to those employed by Tustin (Ref. 19), McRuer, et al (Ref. 1), and Taylor (Ref. 20). Estimates of the signal-to-noise ratio at each input frequency were obtained from a comparison of the power measured at that frequency to the corresponding estimate of remnant power. In order to prevent the expected error in the amplitude-ratio estimate from exceeding 2 dB, estimates of the controller's describing function were disregarded at frequencies for which either the error or control power measurement failed to exceed the corresponding estimate of remnant power by 4 dB.

The describing functions presented in this report relate control activity to the display error. We have chosen to represent control activity in terms of its equivalent effect on error rate when presenting the results of the multi-axis experiments. The amplitude ratios derived from these results thus have units of (deg/sec)/deg or, simply, sec^{-1}. On the other hand, we found that control activity was best represented as newtons of force when presenting the results of the second experimental program. Accordingly, the amplitude ratios for this program are given in units of newtons/degree visual arc. Phase shift is given in degrees throughout this report.

Observation Noise

Although our model of controller remnant is based upon a *vector* observation noise process, our measurement techniques do not readily permit us to extract, from the single spectrum Φ_{uu_r}, the two noise components associated with estimation of display position and display rate. The best we can do, in terms of data manipulation, is to reflect remnant back to an equivalent *scalar* noise injection process.

Consider an observation noise process $r_x(t)$ injected onto system error. If this process is linearly uncorrelated with the input signal i(t), the closed-loop error and control spectra may be separated into the following independent input-related and remnant-related components (refer to Figure A-1):

$$\Phi_{xx} = \Phi_{xx_i} + \Phi_{xx_r}$$

$$\Phi_{uu} = \Phi_{uu_i} + \Phi_{uu_r} \qquad \text{(A-1)}$$

where $$\Phi_{uu_i} = \left| \frac{HV}{1+HV} \right|^2 \Phi_{ii} \qquad \text{(A-2a)}$$

$$\Phi_{uu_r} = \left| \frac{H}{1+HV} \right|^2 \Phi_{rr_x} \qquad \text{(A-2b)}$$

$$\Phi_{xx_i} = \left| \frac{V}{1+HV} \right|^2 \Phi_{ii} \qquad \text{(A-2c)}$$

$$\Phi_{xx_r} = \left| \frac{HV}{1+HV} \right|^2 \Phi_{rr_x} \qquad \text{(A-2d)}$$

Solution of Eqs. (A-2a) and (A-2b) yields for the observation noise spectrum:

$$\Phi_{rr_x} = |V|^2 \frac{\Phi_{uu_r}}{\Phi_{uu_i}} \cdot \Phi_{ii} \qquad \text{(A-3a)}$$

If the observation noise is injected onto error rate instead of error, we obtain

$$\Phi_{rr_x^{\cdot}} = \omega^2 |V|^2 \frac{\Phi_{uu_r}}{\Phi_{uu_r}} \cdot \Phi_{ii} \qquad \text{(A-3b)}$$

Since measurements of Φ_{ii} and Φ_{uu_i} can be obtained only at input frequencies, the observation noise spectrum obtained in this manner can be specified only at those frequencies. At frequencies sufficiently below gain-crossover (i.e., where $|HV| \gg 1$) the observation noise spectrum is approximately identical to the error spectrum at noninput frequencies. Thus, at low frequencies,

$$\Phi_{rr_x} \doteq \Phi_{xx_r} \qquad \text{(A-4a)}$$

or

$$\Phi_{rr_x^{\cdot}} \doteq \omega^2 \Phi_{xx_r} \qquad \text{(A-4b)}$$

Equations (A-3) and (A-4) were both used to compute the observation noise spectrum for a given run. Equation (A-3) was used for estimates at frequencies of 1 rad/sec and higher. Use of this formula at lower frequencies led to anomalous results, apparently because of our inability to obtain reliable estimates of remnant control power at low frequencies. Since remnant error power was relatively large at low frequencies, Eq. (A-4) was used to compute the observation noise spectrum at frequencies below 1 rad/sec. Since the gain-crossover frequency was around 4 rad/sec under most experimental conditions, errors introduced by use of Eq. (A-4) were expected to be on the order of 0.25 dB at 1 rad/sec and less at lower frequencies. In addition, use of Eq. (A-4) provided more low-frequency estimates of the observation noise spectrum than did the formula of Eq. (A-3), since the estimates were not required to correspond to input frequencies.

CALIBRATION OF THE ANALYSIS PROCEDURE

The power spectra and the human controller describing function which are presented in this section in order to illustrate our measurement capabilities were all obtained during data-taking sessions and are representative of the kinds of measures obtained throughout the experimental program. All spectral measurements are presented in dimensions of signal power per measurement window. Although the remnant and noise spectra are more properly considered as samples of spectral density functions that are continuous with frequency, they are presented in the same units as the input-correlated spectra to facilitate evaluation of the measurement techniques.*

*The measurement "window" has the form of $\{[\sin(\pi\omega/\omega_o)]/[\pi\omega/\omega_o]\}^2$ where ω_o is approximately 0.031 rad/sec. The remnant and noise spectra may therefore be converted to units of power per rad/sec by dividing by 0.031 rad/sec, equivalent to adding 15.1 dB.

The ability of the analysis procedures to obtain a correct power spectrum is illustrated by the comparison of theoretical and measured input spectra shown in Figure A-4. The "desired" spectrum is the spectrum that would have been obtained if there were no sources of error. The spectrum labelled "correlated power" contains the measurements obtained at the nominal input frequencies; the spectrum labelled "residual power" indicates samples of measurements obtained at other frequencies. The correlated power spectrum is within 0.1 dB of the theoretical spectrum at frequencies up to 8 rad/sec and begins to fall off relative to the theoretical spectrum at higher frequencies. The difference between the measured and theoretical spectra at high frequencies reflects primarily the effects of a lowpass filter located at the input of the A/D converter. All spectral measurements presented in succeeding chapters of this report have been corrected to account for the effects of this filter. All potential sources of random error, such as (a) error in generating the input sinusoids, (b) analog noise, and (c) noise in the A/D converters are reflected in the residual power spectrum. The large separation between the correlated and residual spectra (about 60 dB at mid-frequencies) shows that we were highly successful in generating input signals which contained significant power at specified frequencies only. (The fraction of total input power that was contained at frequencies other than the specified input frequencies was generally less than 2×10^{-4}.)

In Figure A-5 we compare the remnant and input-correlated portions of the system error spectrum, (Fig. A-5a) and the control effort spectrum (Fig. A-5b) obtained during an experimental trial in which the upper left (UL) display was viewed foveally. The remnant portions of the error and control spectra were at least 10 dB below their corresponding input-correlated spectra at the high- and low-frequency

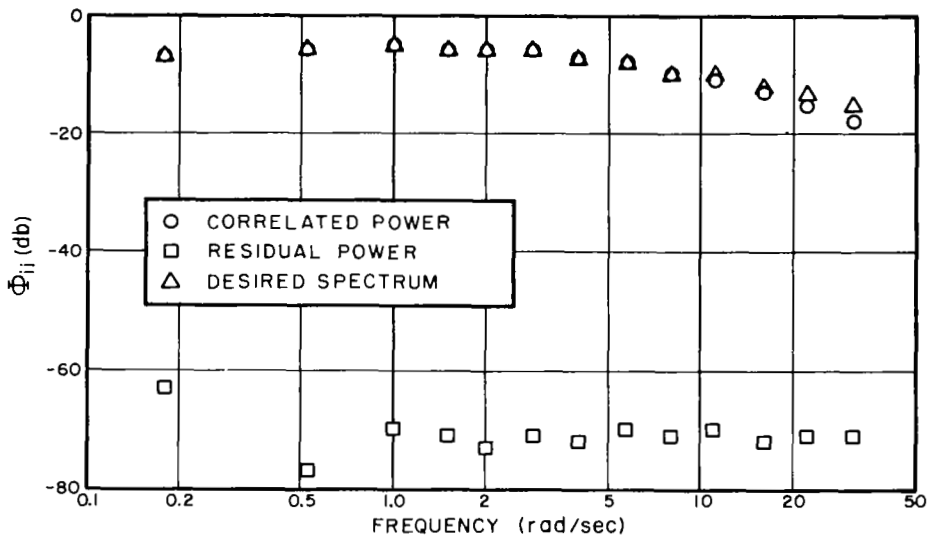


FIG.A-4 COMPARISON OF THEORETICAL AND MEASURED INPUT SPECTRA

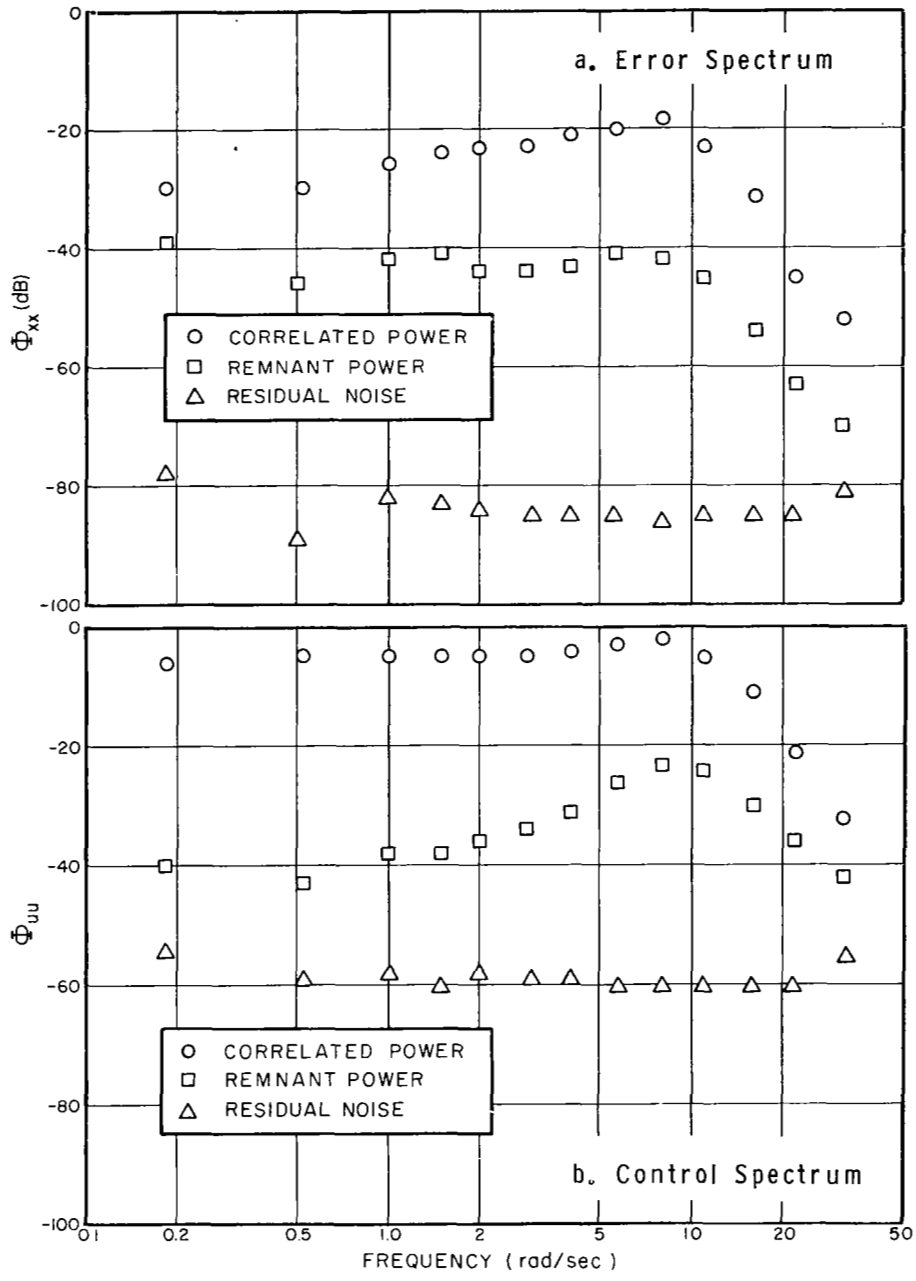


FIG.A-5 COMPARISON OF INPUT-CORRELATED AND REMNANT-RELATED COMPONENTS OF ERROR AND CONTROL SPECTRA

ends of the measurement range, and the separation was significantly greater at mid-frequencies. The relatively low levels of the remnant spectra allowed a valid estimation of the human controller describing function over the entire measurement range for this particular run, as was the case with most of the single-axis foveal measurements. The range of frequency over which valid describing functions could be made generally contracted as the task became more difficult because of the relative increase in controller remnant.

Also shown in Figure A-5 are the residual error and stick spectra measured in the absence of tracking activity on the UL display. The remnant spectra are seen to be at least 10 dB greater than the corresponding residual spectra at all measurement frequencies, which fact would seem to indicate that our remnant measurements are reliable over the entire measurement range. It should be noted, however, that the residual spectra were measured with essentially zero analog input and accordingly may not account for all the measurement noise (say, due to amplitude quantization effects) that occurs when the controller is tracking that axis.

A more critical test of the reliability of our remnant measurements is to test their self-consistency. In the absence of system noise, the remnant error signal will be equal to the remnant control signal filtered by the vehicle dynamics (i.e., we assume that remnant is generated entirely by the human controller). We should then observe that

$$\Phi_{xx_r}/\Phi_{uu_r} = |V|^2 \qquad \text{(A-5)}$$

A comparison of Φ_{xx_r}/Φ_{uu_r} with $|V|^2$ in Figure A-6 shows that the two spectra coincided at frequencies above 1 rad/sec, but that the error/control ratio was consistently less than $|V|^2$ at lower frequencies. These results suggest that, at low frequencies, our measurements either consistently under-estimated the amount of remnant error power or over-estimated the amount of control power. If we attribute measurement errors to additional noise sources that are linearly uncorrelated with the remnant, we must conclude that measurement error must have occurred predominantly on the control channel, since uncorrelated noise powers will add, rather than subtract, on the average. Figure A-5 confirms our conclusion that the remnant error measurements were more reliable than the remnant control measurements at low frequencies, since the dB spread between remnant and residual power was twice as great on the error channel than on the control channel.

In order to show that the human controller describing functions obtained by us are compatible with those that have been obtained by other investigators under similar tracking conditions, a typical describing function of ours is compared in Figure A-7 with a describing function published by System Technology, Inc. in Figure 33 of Ref. 1. Both describing functions were obtained during foveal tracking of a single display with vehicle dynamics of K/s. (The STI data have been scaled to correspond to a K of unity.) The forcing functions used in the two experiments were different, however. STI used a command input whose spectrum was essentially rectangular with a cutoff at 2.5 rad/sec, whereas we used the vehicle disturbance described in Chapter III of this report. Figure A-7 shows that the two describing functions are similar. The most significant difference is the somewhat higher gain achieved in the BBN experiments, resulting in a higher gain-crossover frequency for the BBN data (about 6.5 rad/sec) than

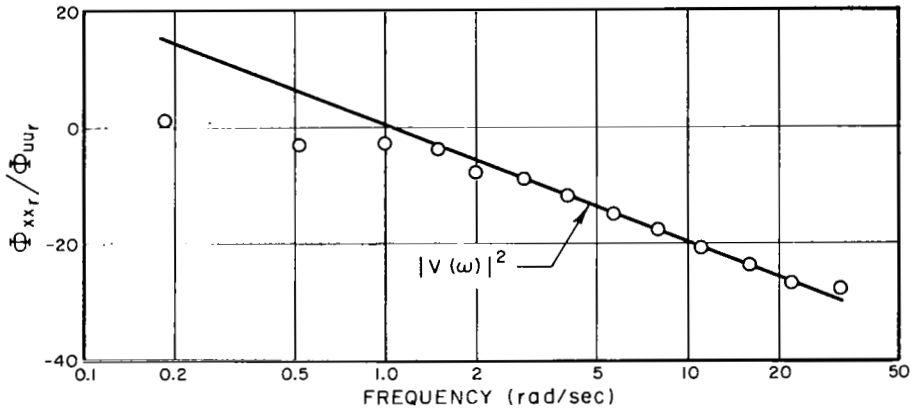


FIG.A-6 COMPARISON OF REMNANT ERROR TO REMNANT CONTROL WITH THE MAGNITUDE OF THE VEHICLE DYNAMICS

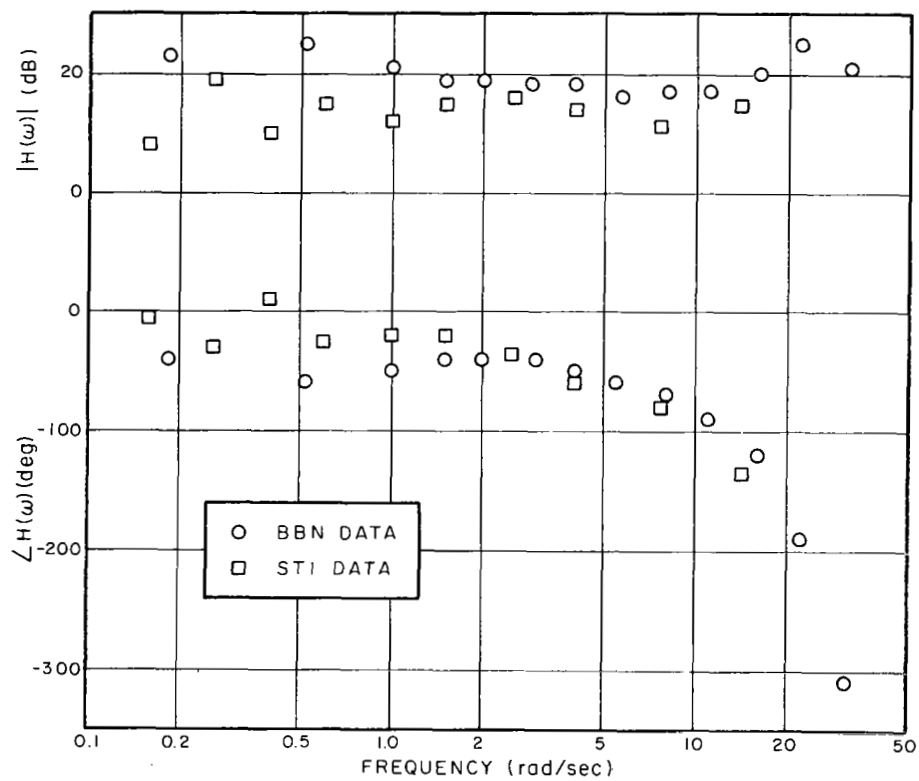


FIG.A-7 COMPARISON OF HUMAN CONTROLLER DESCRIBING FUNCTIONS OBTAINED BY DIFFERENT INVESTIGATORS

that exhibited by the STI data (about 4.5 rad/sec). We suspect that the difference in controller gain is due in part to differences in (a) the input signals, (b) physical characteristics of the manipulators, and (c) training procedures adopted by STI and BBN.

APPENDIX B

APPENDIX B. EXPERIMENTAL RESULTS: MULTI-AXIS CONTROL SYSTEMS

The results of the experimental program on multi-axis control systems are presented in this appendix. Most of the results are presented in the form of averages taken across subjects. Average measures for each subject for each experimental condition are given in Appendix D. Except where otherwise stated, the set of experimental conditions outlined in Table A-1 applies to each of the experiments.

EXPERIMENTAL PLAN

The primary goal of this experimental program was to provide a substantial body of data which would enable us to build and validate models for predicting human control and monitoring behavior in multi-axis control situations. Two variations of a four-axis manual tracking task were performed in which the controller was allowed to scan among the displays. Additional two-, three-, and four-axis tasks were investigated in which visual scanning was prohibited. The object of these experiments was to provide data against which models for task interference could be tested most readily. (A model for task interference is developed in Section 3 of this report, and experimental validation is provided in Sections 4 and 5.)

In addition to the multi-axis experiments, two single-axis experiments were also performed. One of these was performed to evaluate alternative display formats for use in the multi-axis experiments. The other was a set of calibration experiments to determine the base levels of performance under each viewing condition employed in the multi-axis experiments.

COMPARISON OF TWO DISPLAY FORMATS

Experimental Conditions

The following two display formats were compared in terms of their effectiveness as tracking displays: (1) a moving error dot and stationary reference circle, and (2) a moving error bar and stationary reference bar. Each type of display was tracked when viewed foveally and when located 30 degrees into the periphery along the horizontal axis. Motion of the error indicators was in the vertical direction, thereby allowing the subject to extrapolate a zero reference to the peripheral display. Only one display was tracked during a single experimental run. Every subject tracked each display twice under each viewing condition – once with the left hand and once with the right. Forcing-function Set A (Table A-2) was used, and the control gain was such that 1 newton of stick force produced a rate of error deflection of 4 deg/sec.

MSE Scores

Table B-1 shows that the subjects achieved lower mean-squared error scores for both the foveal and peripheral viewing conditions when presented with the bar display. The differences were about 10% when viewing foveally and about 25% for peripheral viewing. An analysis of variance of these scores indicated that both differences were significant at the 0.05 level. Because of the superior performance achieved with the bar display, the bar display was used for the remainder of the experimental program. Scores for individual subjects are given in Table D-1 of Appendix D.

SINGLE-AXIS CALIBRATION EXPERIMENTS

Experimental Conditions

In this experiment we obtained the complete set of single-axis base level performance measures required by our model of the multivariable control situation. Each subject tracked each of four axes

TABLE B-1

Effect of Display Format on Mean-Squared Error Scores

Display Format	Mode of Viewing	
	Foveal	Peripheral
Circle, Dot	.11	.80
Bar, Bar	.10	.63

MS error scores in deg^2 visual arc.
Average of 4 subjects, 2 trials per condition.

for each of the four fixation points to make a total of sixteen single-axis trials. There were four viewing conditions associated with each axis: (a) foveal, (b) 16 degrees into the periphery with the ability to extrapolate the zero reference, (c) 16 degrees into the periphery with no reference extrapolation possible, and (d) 22 degrees into the periphery (along the diagonal), also with no reference extrapolation. The forcing functions were made identical on all axes so that interaxis differences in tracking behavior would not be induced by differences among forcing functions.

MSE Scores

The effects of relative fixation point on average mean-squared error scores are shown in Table B-2. Inter-axis standard deviations, based on the scores (averaged across subjects) corresponding to each viewing condition, are shown in Table B-2a. Inter-subject standard deviations are given in Table B-2b. The latter measures were computed from the subject means obtained by averaging the four MSE scores corresponding to each relative fixation point. System

TABLE B-2

Effect of Relative Fixation Point
on Single-Axis Mean-Squared Error Scores

a. Average Scores for Each Axis Tracked

Axis Tracked	Relative Fixation Point			
	Foveal	16° Periph. Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
UL	.12	.55	.73	1.7
LL	.14	.32	.63	1.4
LR	.13	.45	1.3	1.5
UR	.13	.39	1.5	2.0
Mean	.13	.43	1.0	1.6
SD	.0098	.094	.43	.25
SD/Mean	0.08	.22	.42	.16

Average of 4 subjects, 1 run per subject.

b. Average Scores for Each Subject

Subject	Relative Fixation Point			
	Foveal	16° Periph. Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
JF	.081	.23	.52	1.1
DM	.12	.32	.58	1.4
GP	.18	.64	1.8	2.1
HS	.14	.52	1.3	1.9
Mean	.13	.43	1.0	1.6
SD	.043	.19	.60	.50
SD/Mean	.33	.45	.58	.30

Average of 4 runs per subject. MS error scores in deg^2 of visual arc.

error and control effort scores for each subject for each viewing condition are shown in Tables D-3 through D-7 of Appendix D. An explanation of the symbols appearing in these tables is given in Table D-2.

The rank ordering of the tasks in terms of mean-squared error scores was, from easiest to most difficult: foveal, 16° separation with reference extrapolation, 16° separation with no reference extrapolation, and 22° separation. The fractional differences in scores between pairs of tasks in the above ordering decreased with increasing task difficulty. Similarly, the statistical significance of the differences (as given by an analysis of variance of the error scores) decreased with increasingly difficult pairs of conditions. The easiest of the peripheral tasks yielded a score that was about 3.3 times the foveal score (significant at the 0.001 level). Removal of the facility to extrapolate the baseline increased the score by an additional factor of 2.4, a difference significant at the 0.01 level. The score obtained from viewing at 22° was about 60% greater than the 16°, no reference, score. The latter difference was minimally significant at the 0.05 level.

The relatively small inter-axis standard deviation (less than 10% of the mean score) associated with foveal tracking indicates that the subjects were essentially equally proficient on the four axes tasks under this viewing condition. The larger normalized inter-axis SD's associated with the peripheral tasks may reflect effects such as the up-down differences in peripheral viewing capabilities which we found to exist. Except for the diagonal viewing condition, the inter-axis and inter-subject standard deviations increased, relative to their corresponding mean scores, with increasing task difficulty. The inter-subject SD was consistently greater than the inter-axis SD for each viewing condition, reflecting differences in overall tracking abilities among subjects.

TABLE B-3

Effect of Display Location on
Single-Axis Mean-Squared Error Scores

a. Horizontal Viewing, 16° Display Separation

	Location of Axis Tracked	
	Left	Right
Reference Extension	.32	.39
No Reference Extension	.73	1.3

b. Vertical Viewing, 16° Display Separation

	Location of Axis Tracked	
	Up	Down
Reference Extension	.54	.45
No Reference Extension	1.5	.63

c. Diagonal Viewing, 22° Display Separation

	Location of Axis Tracked		
	Left	Right	Average
Up	1.7	2.0	1.8
Down	1.4	1.5	1.4
Average	1.6	1.7	---

Average of 4 subjects. MS error scores in deg^2 visual arc.

The effects of display location on average peripheral MSE scores are shown in Table B-3. Left-right and up-down differences are shown respectively in Tables B-3a and B-3b for the 16° viewing angle with and without reference extrapolation. Up-down and left-right differences for 22° diagonal viewing are shown in Table B-3c. The greatest left-right difference occurred for the 16° viewing condition with no reference extrapolation. The score obtained when fixating left and tracking the display on the right was about 75% greater than the score obtained in the reverse condition (Table B-3a, bottom row). Why this

is so is not clear, especially since three of the four subjects were right-handed. The largest up-down difference occurred again for the 16° viewing condition with no reference extrapolation. The score obtained when fixating down and tracking up was almost 2.5 times the score obtained when fixating up and tracking down (Table B-3b, bottom row).

The smaller left-right and up-down differences in MSE scores which were observed for the diagonal viewing conditions were consistent in direction with the above results. For example, Table B-3c shows that the largest MSE score was obtained when the subject fixated the lower left display and tracked the upper right, whereas the lowest score (for diagonal viewing) was obtained when fixating upper right and tracking lower left.

Because of the relatively large variances in the scores, analysis of variance tests failed to attribute statistical significance to the left-right and up-down differences. Nevertheless, because of the nonhomogeneous scanning behavior exhibited by the subjects in the multiaxis control situations (described later in this appendix), we suspect that the up-down differences reflect important differences in the subjects' viewing capabilities.

In order to provide an overall-performance measure that will be useful for interpreting the multi-axis results discussed later in this appendix, we show the average total performance score in Table B-4. This measure is defined here as the sum of the four scores (one foveal and three peripheral) corresponding to fixation of a single display. Scores for individual subjects are given in Table D-8 of Appendix D. The largest total score on the average was achieved from fixation of the lower left display. This score

TABLE B-4

Effect of Fixation Point on Total Mean-Squared Error Score

Display Fixated	Display Fixated: Left	Right	Average
Up	2.6	2.7	2.6
Down	3.9	3.5	3.7
Average	3.2	3.1	3.2

MS error scores in deg^2 visual arc. Average of 4 subjects, one score per subject.

TABLE B-5

Partitioning of Average Mean-Squared Error and Control Scores

	Mode of Viewing: Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
MS error, $\overline{X^2}$ (deg^2)	.13	.43	1.0	1.6
Error variance, σ_x^2 (deg^2)	.13	.41	.81	1.2
$\sigma_x^2/\overline{X^2}$	.99	.97	.86	.78
Correlated error, σ_{xi}^2 (deg^2)	.095	.20	.22	.27
Remnant error, σ_{xr}^2 (deg^2)	.034	.22	.59	.93
σ_{xr}^2/σ_x^2	.24	.50	.70	.77
Control variance, σ_u^2 $(deg/sec)^2$	4.2	5.5	6.1	7.2
Correlated Control, σ_{ui}^2 $(deg/sec)^2$	3.1	2.6	2.6	2.6
Remnant Control, σ_{ur}^2 $(deg/sec)^2$	1.1	2.9	3.5	4.7
σ_{ur}^2/σ_u^2	.21	.47	.54	.63

Data averaged over four subjects, four runs per subject.

was about 1.6 times as great as the lowest score (corresponding to fixation of the upper left display). Left-right performance differences were on the average only about 5 percent and were not statistically significant. On the other hand, the average score corresponding to fixation of the lower displays was about 1.4 times the average score when fixation was directed at the upper displays. An analysis of variance showed this difference to be significant at the 0.05 level.

Analysis of the total performance measure thus reveals that significantly better visual information is obtained overall when the upper displays are fixated than when the lower displays are fixated. On this basis we would predict that the subject would tend to fixate the upper displays more than the lower displays when tracking the four displays simultaneously, other conditions being equal.

Comparison of the mean-squared error score with the error variance (mean-squared error minus the square of the mean error), along with a partitioning of the error and control variances into input-correlated and remnant components, is given in Table B-5. The error variance was nearly identical to mean-squared error for foveal viewing and for peripheral viewing with reference extrapolation; thus, the mean error was essentially zero for these conditions. When reference extension into the periphery was not possible, however, the error variance was significantly less than the mean-squared error, accounting for only 78% of the latter for the 22° viewing condition. We suspect that the occurrence of a significant nonzero mean error was a direct consequence of the subject's inability to estimate precisely the location of the zero reference.

There was no appreciable difference between mean-squared control movement and the variance of the control movement. Therefore, we have omitted the mean-squared control score from Table B-5. Note that when the vehicle dynamics are K/s and the input has zero mean, it is necessary for the subject to apply a control signal with zero mean in order to maintain absolute stability of the system.

The fractions of error and control variance accounted for by controller remnant increased monotonically with the difficulty of the basic task, primarily because the remnant powers increased. In addition, the input-correlated portion of the control variance was greatest for the foveal task and was about 15% less for all of the peripheral tasks. This trend in the correlated control power suggests that the controller generated his highest gain for foveal viewing and a somewhat lower gain for the three peripheral conditions. The monotonically increasing remnant power, on the other hand, suggests a monotonically increasing observation noise process with increasing task difficulty. These predictions are substantially borne out by the describing function and observation noise measurements reported below.

Observation Noise Spectra

Average power spectral density functions of normalized observation noise are given for the four viewing conditions in Figure B-1. Averages were obtained across subjects and across axes tracked; thus, each measurement represents the average of sixteen spectral density functions (hereinafter referred to simply as "spectra"). Normalization is with respect to error variance. The spectra as shown include power contained at negative as well as positive frequencies and have units of normalized power per rad/sec. The data shown in Figure B-1 are presented in tabular form in Table D-9 of Appendix D.

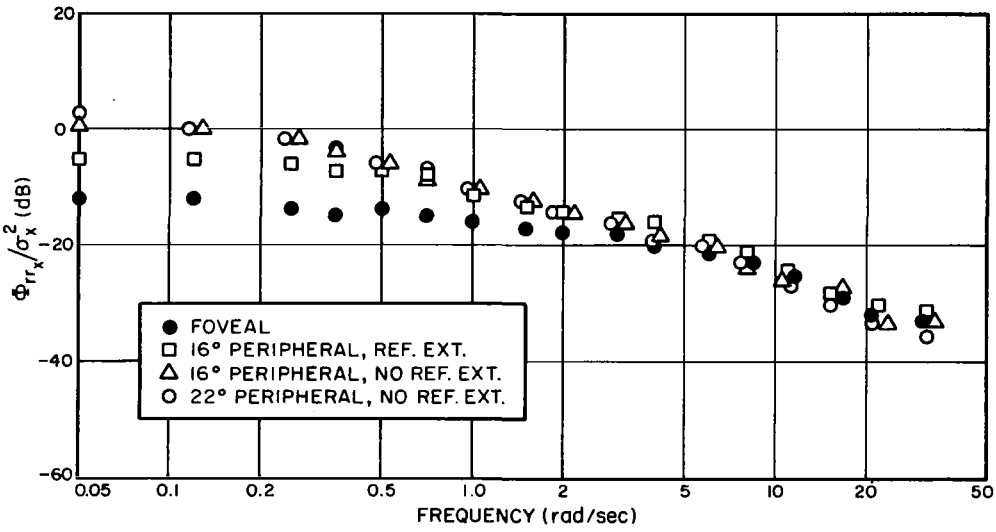


FIG. B-1 EFFECT OF VIEWING CONDITIONS ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Average of 4 subjects, 4 trials/subject

Although there were no consistent differences among the four average spectra at high frequencies, the low-frequency normalized observation noise increased appreciably as viewing conditions were made less favorable. There were no appreciable differences between the normalized noise processes associated with the two peripheral viewing conditions that did not permit reference extrapolation. If we were to plot the absolute (i.e., unnormalized) observation noise spectra, however, the differences would be accentuated and we would see that the relationship of observation noise to viewing conditions was consistent with the relationship of the remnant error and stick scores to viewing conditions.

Table B-6 shows that the inter-axis standard deviation of the normalized observation noise was relatively small – generally on the order of 1.0 to 1.5 dB at a given measurement frequency.* The inter-subject standard deivations, shown in Table B-7, were about twice as great – around 2.5 to 3.0 dB – as might be expected from differences in tracking ability among subjects.

Describing Functions

Average human controller describing functions are presented graphically for the four viewing conditions in Figure B-2 and are tabulated in Table D-10 of Appendix D. As we predicted from the analysis of the input-correlated portion of the control power, the primary difference among the describing functions was the higher gain at low and mid frequencies accompanying foveal viewing. The approximate gain-crossover frequencies for foveal and peripheral

*Standard deviations were computed on the same basis as the standard deviations of the MSE scores given in Table B-2, except that the inter-axis SD was based on observation noise measures obtained from the lower right axis only.

TABLE B-6

Effect of Relative Fixation Point on Inter-Axis Standard Deviations of Single-Axis Normalized Observation Noise Spectra

Computed from average spectra of 4 subjects, 4 runs/subject

Frequency (rad/sec)	Relative Fixation Point			
	Foveal	16° Sep. Ref Ext	16° Sep. No Ref Ext	22° Sep. No Ref Ext
.05	1.4	3.1	2.5	1.8
.12	1.9	1.7	1.3	1.2
.25	1.2	2.0	1.6	0.6
.35	1.7	2.0	1.3	1.1
.50	0.9	1.6	1.1	2.1
.71	0.5	1.3	1.5	1.1
1.0	2.4	1.9	0.9	0.8
1.5	0.9	1.1	0.6	1.3
2.0	0.8	1.3	0.7	0.5
2.9	0.3	1.3	1.2	0.7
4.0	0.4	1.2	1.3	0.7
5.7	0.4	0.8	1.9	0.5
8.0	0.5	0.9	0.7	0.8
11.0	0.3	1.2	0.8	1.6
16.0	0.7	1.8	1.8	2.0
22.0	0.2	1.7	1.8	0.6
32.0	0.7	1.2	0.8	2.2
Average	0.89	1.5	1.3	1.2

Standard deviations in dB.

TABLE B-7

Effect of Relative Fixation Point on Inter-Subject Standard Deviations of Single-Axis Normalized Observation Noise Spectra

Frequency (rad/sec)	Relative Fixation Point			
	Foveal	16° Sep Ref Ext	16° Sep No Ref Ext	22° Sep No Ref Ext
.05	3.9	3.9	3.0	4.5
.12	4.0	2.5	2.4	5.0
.25	1.5	5.4	4.2	0.7
.35	1.3	3.3	1.4	4.6
.50	2.4	2.9	1.6	1.8
.71	3.0	2.1	1.0	2.5
1.0	2.6	4.0	5.4	1.8
1.5	2.2	3.2	1.9	2.0
2.0	2.5	1.7	3.1	1.9
2.9	2.9	2.5	2.5	2.0
4.0	2.4	2.5	1.6	3.7
5.7	2.0	2.5	4.7	2.7
8.0	2.2	0.5	3.6	2.5
11.0	2.4	3.6	1.7	2.6
16.0	2.1	2.7	1.0	1.8
22.0	2.0	4.7	2.6	2.6
32.0	2.3			
Average	2.5	3.0	2.6	2.7

Axis Tracked: Lower Right
Standard deviations in dB.

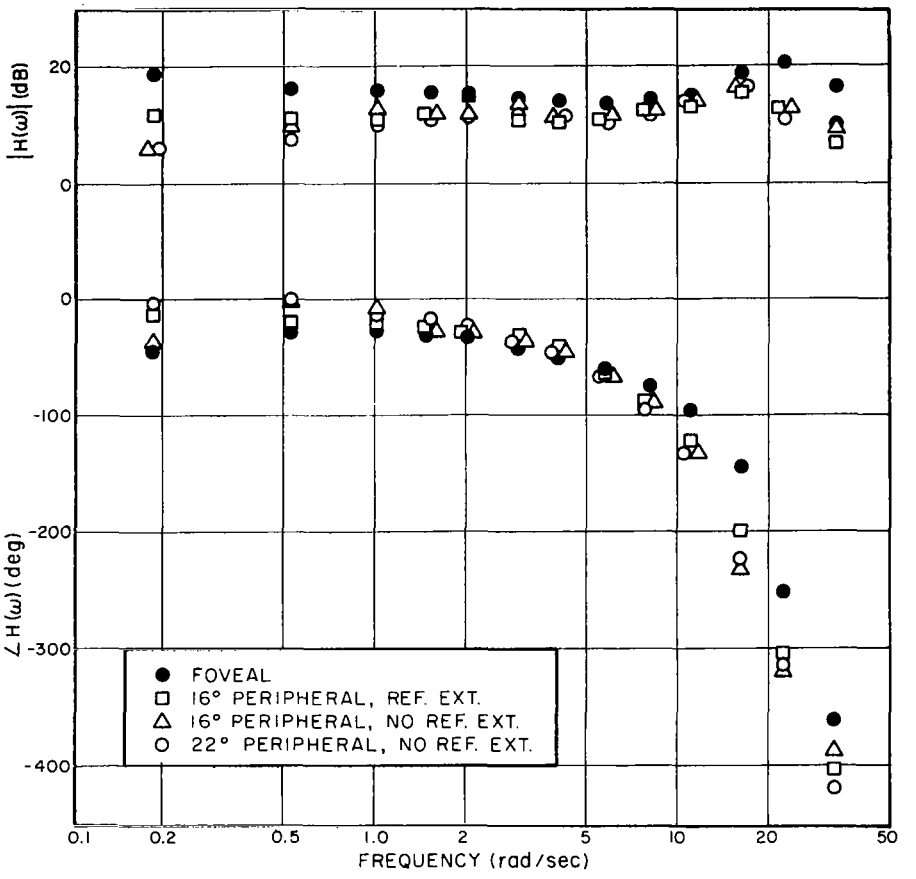


FIG. B-2 EFFECT OF VIEWING CONDITIONS ON THE HUMAN CONTROLLER'S DESCRIBING FUNCTION

Average of 4 subjects, 4 trials/subject

viewing were, respectively, 5 and 4 rad/sec. In addition, a somewhat larger phase lag was observed at high frequencies for peripheral viewing, and the three peripheral amplitude-ratio curves appeared to exhibit a resonance peak around 16 rad/sec, whereas the peak for foveal tracking was closer to 22 rad/sec.

Inter-axis standard deviations of the gain and phase-shift measurements are presented in Table B-8,and Table B-9 shows inter-subject standard deviations. The inter-axis gain and phase SD's were both relatively small, about 1-2 dB on the average for the gain measurements and around 10 degrees on the average for phase shift. The average inter-subject standard deviation was somewhat greater for the gain measurements – about 3 dB. The standard deviation in phase shift was generally under 10 degrees, although the variability was considerably greater for some of the high frequency phase-shift measurements obtained during peripheral viewing with no reference extension.

MULTIAXIS TRACKING PERFORMANCE WITH NO VISUAL SCANNING

Experimental Conditions

This experiment was designed to determine whether or not there is interference among multiple tasks performed in parallel when the displays are separated, and to determine the nature of such interactions as might exist. We consider "interference" to exist whenever a single-axis performance measure observed on a given axis differs significantly from the same measure obtained on that axis when simultaneous tracking of two or more axes is required. In order to isolate such interference from the effects of visual scanning, the subjects were required to fixate a single display while tracking two or more axes.

TABLE B-8

Effect of Relative Fixation Point on Inter-axis Standard Deviations of Single-axis Human Controller Describing Functions

Computed from average describing functions of four subjects, 4 runs per subject.

Frequency (rad/sec)	Relative Fixation Point							
	Foveal		16° Sep, Ref Ext		16° Sep, No ref ext		22° Sep, No Ref Ext	
	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)
.18	0.2	13	4.0	25	2.4	16	—	—
.52	0.9	7	1.3	3	1.3	7	2.1	12
1.0	0.6	6	1.0	5	0.7	3	1.4	15
1.5	0.6	3	2.3	2	1.7	8	1.0	5
2.0	1.0	2	1.3	3	1.1	4	0.5	10
2.9	0.5	2	1.3	2	5.0	3	1.0	3
4.0	0.6	3	1.0	1	0.6	4	0,3	4
5.7	0.4	2	1.0	4	0.9	4	0.3	3
8.0	0.5	2	0.5	2	0.5	4	0.3	4
11.	0.5	2	1.5	3	0.9	15	1.6	14
16.	0.8	6	3.4	15	3.8	25	1.5	4
22.	1.6	17	0.8	33	1.9	13	2.1	12
32.	1.2	25	4.5	6	2.2	15	2.6	36
Average	0.7	7	2.0	8	1.8	9	1.2	10

TABLE B-9

Effect of Relative Fixation Point on Inter-subject Standard Deviations of Single-axis Human Controller Describing Functions

Axis Tracked: Lower Right, One run per subject

Frequency (rad/sec)	Relative Fixation Point							
	Foveal		16° Sep, Ref Ext		16° Sep, No ref ext		22° Sep, No Ref Ext	
	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)	Gain(dB)	Phase(deg)
.18	2.8	17	.4	1	5.1	19	—	—
.52	2.1	14	4.0	7	5.2	18	1.3	6
1.0	2.5	6	3.0	8	4.2	17	3.8	20
1.5	2.2	7	3.5	5	3.7	13	4.0	11
2.0	2.0	5	2.7	4	3.1	9	2.7	9
2.9	1.8	2	3.2	9	1.8	8	3.4	5
4.0	1.4	5	2.3	10	1.8	8	1.0	9
5.7	1.0	1	2.5	8	2.6	12	1.9	7
8.0	1.2	4	2.8	1	1.9	13	1.9	7
11.	1.1	6	4.3	15	1.9	41	2.4	17
16.	0.9	15	6.2	13	3.7	61	2.0	63
22.	3.6	24	1.4	6	4.0	83	5.5	51
32.	3.9	14	—	—	—	—	—	—
Average	2.0	9	3.0	7	3.3	25	2.7	19

This experiment can be divided logically into three distinct phases. In the first phase, the subject tracked two displays: one foveally and one peripherally. This condition was investigated partly to complement and verify the earlier set of experiments (Ref. 4) and partly to provide the most straightforward procedure for investigating a simple foveal-peripheral interference. The second phase of this experiment was performed to investigate interference among peripheral tasks only. Three peripheral displays were tracked individually and simultaneously. One foveal and three peripheral axes were tracked simultaneously in the third phase of this experiment in order to provide an indication of the interference that would arise in the four-axis tracking task. The four-display configuration shown in Fig. A-4 was used throughout this experiment. Input Set C (Table A-4) was used, and the experimental conditions were otherwise as shown in Table A-1.

In order to provide a two-axis situation that was to a large extent compatible with the two-display conditions investigated previously (Ref. 4), only the configurations allowing peripheral extrapolation of the zero reference were investigated. All such configurations were investigated. Each display was fixated in turn by each subject, and for each fixation point the subject was required to track: (a) the display fixated, (b) the peripheral display located in the nearest clockwise position, and (c) the foveal and peripheral displays simultaneously.

One objective of the two-axis no-scan experiment was to determine whether the subjects would "attend" preferentially to the foveal task, as one might expect from habit, or whether they would distribute their efforts between the two axes in a manner appropriate to the scoring situation. A second objective of this experiment was to determine the relation of the magnitude of the

interference to the number of hands required for control. Note that two hands were required for 2-axis control when the two active displays were aligned horizontally, whereas only one hand operating a single controller in two dimensions was needed when the displays were aligned vertically.

In previous experimentation with a two-axis integrated display and control situation (Ref. 3), we found a small interference which we ascribed to "visual-motor" effects. If that interference had been due primarily to cross-coupling effects at the motor end, we would expect to find a greater interference when one hand was used for control than when two hands were used. On the other hand, if the interference found in the previous experiment had arisen primarily from mechanisms residing in the visual or central pathways, then we would expect the results obtained in the two-axis experiments with separated displays to be independent of the number of hands used for control.

In order to reveal any fundamental differences that might exist between interference on the foveal and peripheral tasks, it was important that the subject assign equal importance to the two tasks. The mean-squared input was therefore increased to 5.8 $(\text{deg/sec})^2$ for the foveal task (as opposed to 2.2 $(\text{deg/sec})^2$ for the peripheral input) in an attempt to force the single-axis mean-squared error to be approximately the same for both foveal and peripheral tracking. Thus, if there were no interference, or if the interference were basically the same for both the foveal and peripheral tasks, the two tasks would contribute equally to the total performance score when the two axes were tracked simultaneously. It was assumed that the subjects would learn to assign equal subjective weightings to the two tasks in this hypothesized situation when given sufficient training.

The three-axis peripheral and four-axis no-scan experiments were performed in order to reveal the kinds of interference that might be expected in the scanning experiment. Consequently, the mean-squared inputs were kept the same at 2.2 $(deg/sec)^2$ on all axes to provide compatibility with the first scanning experiment that was conducted. In order to economize experimental time, only the UL display was fixated throughout the experiment. Each subject tracked each display individually twice, the four displays simultaneously twice, and the three peripheral displays simultaneously four times (twice each for the second and third phases of this experiment).

MSE Scores

Average mean-squared error scores are shown for all phases of this experiment in Table B-10. Individual subject averages are given in Tables D-11 and D-12 of Appendix D. The column labeled "Total Score" indicates the sum of the scores of the component tasks. In the case of the four-axis experiment summarized in part (c) of this table, the three-axis peripheral task is treated as a single, complex task, and the "1+3-axis" score indicates the sum of the 1-axis foveal and three-axis peripheral scores.

Statistically significant interference was found in all three variations of this experiment. ("Interference" is defined here as the difference between the single-axis and multi-axis MSE scores.) Table B-10a shows that the 2-axis total score was about 1.7 times the 1-axis score. Interference on the foveal and peripheral axes differed slightly: the peripheral score increased by about a factor of 2.0, whereas the foveal score increased by a factor of about 1.6. We suspect that the greater degradation in performance on the peripheral task was due primarily to the fact that the peripheral task

TABLE B-10

Effect of Number of Axes Tracked on Average Mean-squared Error Scores

a. One Foveal and One Peripheral Task

Foveal		Peripheral		Total Score	
1-axis	2-axis	1-axis	2-axis	1-axis	2-axis
.32	.50	.23	.45	.55	.95

b. Three Peripheral Tasks

16°, ref ext		16°, no ref ext		22°, no ref ext		Total Score	
1-axis	3-axis	1-axis	3-axis	1-axis	3-axis	1-axis	3-axis
.27	.90	.51	1.3	1.2	1.9	2.0	4.2

c. One Foveal and Three Peripheral Tasks

Foveal		Combined Periph.		Total Score	
1-axis	4-axis	3-axis	4-axis	1+3-axis	4-axis
.11	.30	4.2	4.8	4.1	5.1

MS error scores in deg^2 visual arc
Average of 4 subjects

was the easier task – as indicated by the 1-axis MSE score – despite our efforts to provide foveal and peripheral tasks of equal difficulty. Interference on the foveal, peripheral, and total scores was significant at the 0.001 levels.

Table D-11 of Appendix D shows that the amount of interference between axes was dependent on the number of hands required for control. The ratio of the average 2-axis total score to the average 1-axis total score was about 1.5 for tasks requiring two hands and about 2.0 for the one-handed control configuration. This difference, which was significant at the 0.01 level, suggests that interference at the motor level could be reduced somewhat by assigning control of the component tasks to separate limbs. We conclude from these results that the small (10-15%) interference observed in the earlier experiments with integrated controls and displays were caused largely by interference effects in the motor, rather than visual, pathways.

Table B-10b shows that the total 3-axis peripheral score was about twice the sum of the three individual 1-axis peripheral scores. This difference was significant at the 0.001 level. The ratios of 3-axis score to 1-axis score were quite different for the three axes individually. The largest such ratio (about 3.6) was observed for the easiest of the peripheral tasks (16° viewing angle with reference extrapolation). A 3-axis, 1-axis MSE ratio of about 2.5 was observed for the 16° peripheral task with no reference extrapolation, and the smallest ratio (about 1.5) was observed for the 22° peripheral task. The tendency for the largest proportional increase in score to occur on the easiest task has been observed in previous experimental work (Ref. 3) and is consistent with the subject's instructed task of minimizing the total mean-squared error score. That is, if the subject must effectively apportion his information processing capability among the three tasks, it would seem that the optimal strategy is to devote the largest portion of his capacity to

the most important task and allow a relatively large increase in score to occur on the axis that contributes least to the total performance measure. The model validation studies presented in Section 4 support this hypothesis.

Table B-10c shows that the total score increased by only about 25% when the simple foveal and complex peripheral tasks were performed simultaneously. This difference was statistically significant at the 0.05 level. The largest relative performance degradation was observed for the foveal task, for which the 4-axis score was about 2.7 times the 1-axis score. The score for the combined peripheral task increased by only about 20% upon addition of the foveal task. The difference between the relative magnitudes of the foveal and peripheral interactions is consistent with the notions of optimality discussed above.

The MSE scores obtained from the second and third phases of this experiment are shown again in Table B-11 to facilitate comparison. This table shows at a glance that the score on each axis increased as the number of axes that were tracked simultaneously increased. In addition, the differences between the 3-axis and 1-axis scores were consistently greater than the differences between the 4-axis and 3-axis scores. Table B-11 shows also that the ratios of error variance to mean-squared error observed under 4-axis tracking conditions varied with viewing conditions in about the same way as they did for single-axis tracking (Table B-5), except for the somewhat lower ratio associated with foveal tracking. Thus, the interference among axes implicit in the MSE scores cannot be attributed to any significant extent to increased manifestations of the disappearing-reference phenomenon.

TABLE B-11

Effect of Number of Axes Tracked on Average Peripheral Mean-squared Error Scores

Mode of Viewing	Mean-squared Error			$\sigma_x^2/\overline{x^2}$
	Number of Axes Tracked			for 4-axis
	1 Axis	3 Axes	4 Axes	Tracking
Foveal	.11	—	.30	.93
16 periph, ref ext	.27	.90	1.0	.97
16 periph, no ref ext	.51	1.3	1.6	.88
22 periph, no ref ext	1.2	1.9	2.2	.80
Total Score	2.1	—	5.1	—

MS error scores in deg^2 visual arc
Average of 4 subjects

Observation Noise Spectra and Describing Functions

Normalized observation noise spectra and human controller describing functions for the first phase of this experiment (simple foveal-peripheral interaction) are shown in Figs. B-3 and B-4 for foveal and peripheral tracking, respectively. The graphs presented in these figures represent the average performance of four subjects when fixating the UR display and tracking either the UR display foveally, the LR display peripherally, or both displays simultaneously without scanning. Figure B-3 shows that the 2-axis foveal observation noise spectrum was about 3-4 dB greater than the corresponding 1-axis noise spectrum. The shapes of corresponding 1- and 2-axis spectra did not appear to differ in a consistent manner. The 2-axis foveal describing functions differed from the corresponding 1-axis DF primarily by an overall decrease in amplitude ratio of about 2-3 dB. The two-axis describing function also appeared to have a slightly greater high-frequency phase lag and lower resonant frequency than the 1-axis DF. Figure B-4 shows that the same kinds of 1-axis, 2-axis differences occurred on the peripheral axis.

Observation noises and describing functions obtained from a single subject performing the 3-axis and 4-axis tasks are compared with the corresponding 1-axis performance measures in Fig. B-5 through Fig. B-8. An increase in the number of axes to be tracked simultaneously had the following effects on all axes: (a) the overall level of the normalized observation noise spectrum increased, (b) the controller's amplitude ratio decreased, and (c) the high-frequency phase lag increased slightly. For the most part, the shape of the observation noise spectrum is unaffected by the number of axes tracked. The most noticeable exception to this generalization is seen in Fig. B-5a, which shows that the 1-axis, 4-axis differences on the foveal axis are greater at frequencies below gain-crossover (around 4 rad/sec) than at higher frequencies.

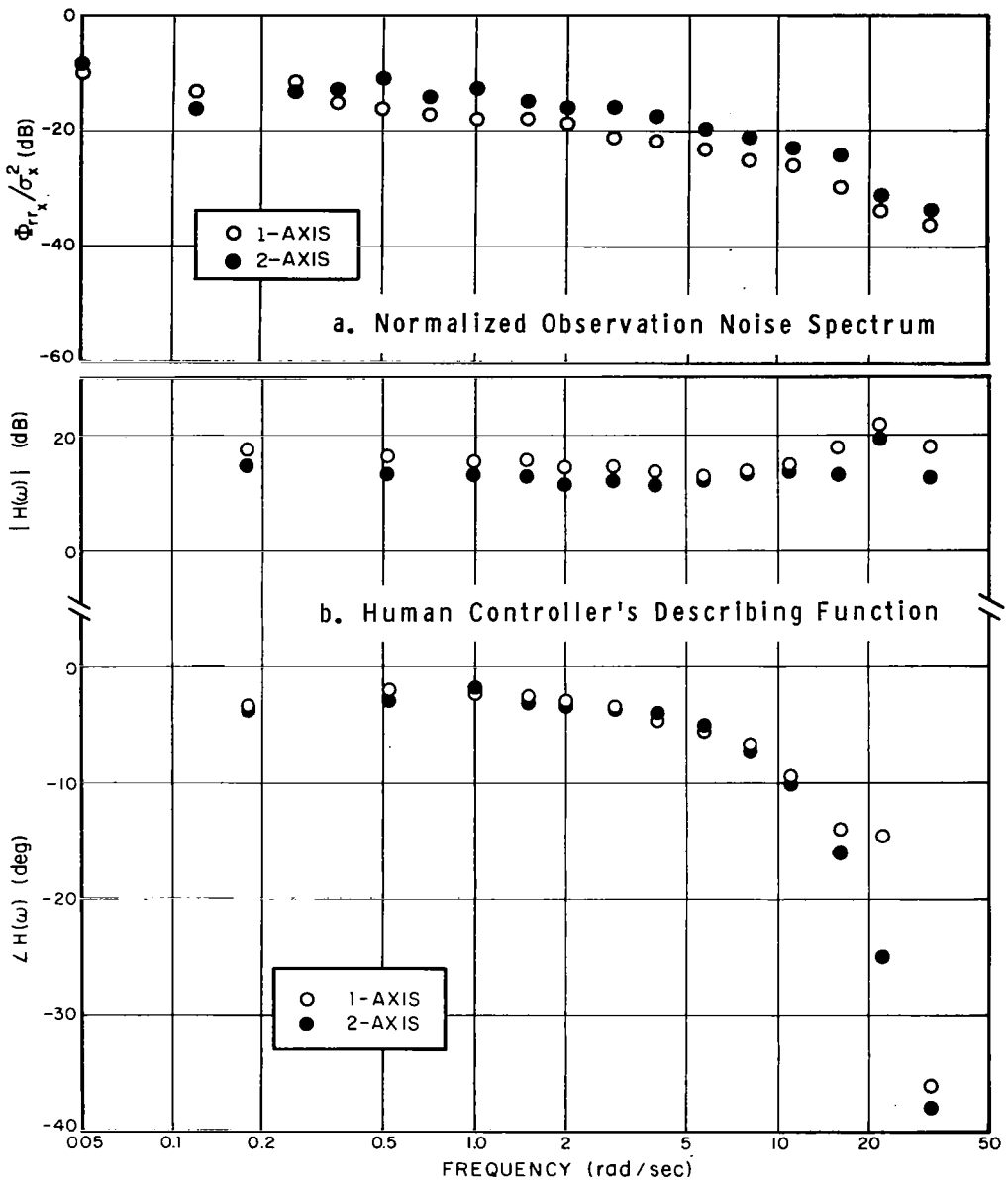


FIG.B-3 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: FOVEAL VIEWING

Average of 4 subjects, 1 trial/subject

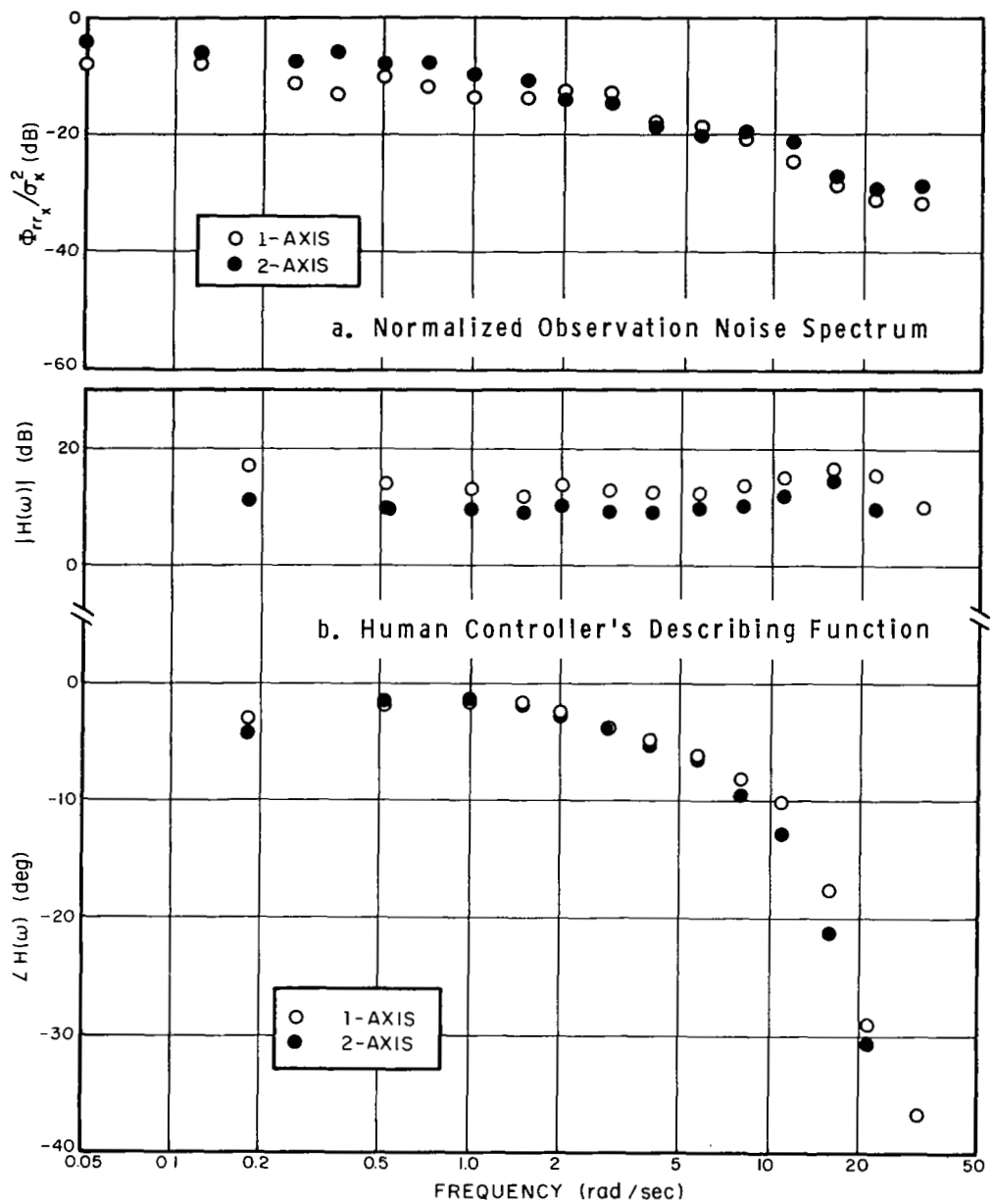


FIG.B-4 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: 16° PERIPHERAL VIEWING WITH REFERENCE EXTRAPOLATION

Average of 4 subjects, 1 trial/subject

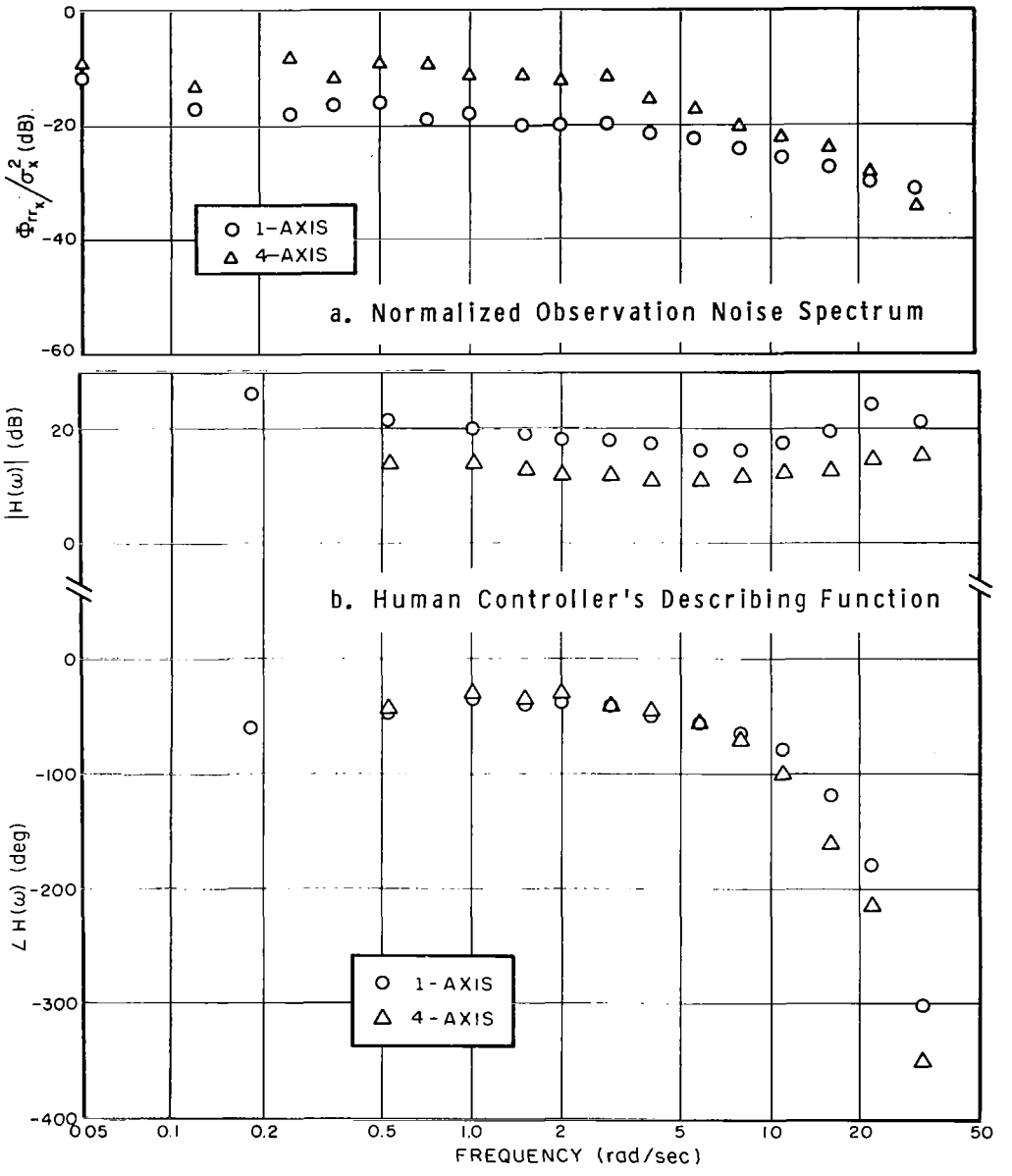


FIG.B-5 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: FOVEAL VIEWING

Subject JF, average of 2 trials

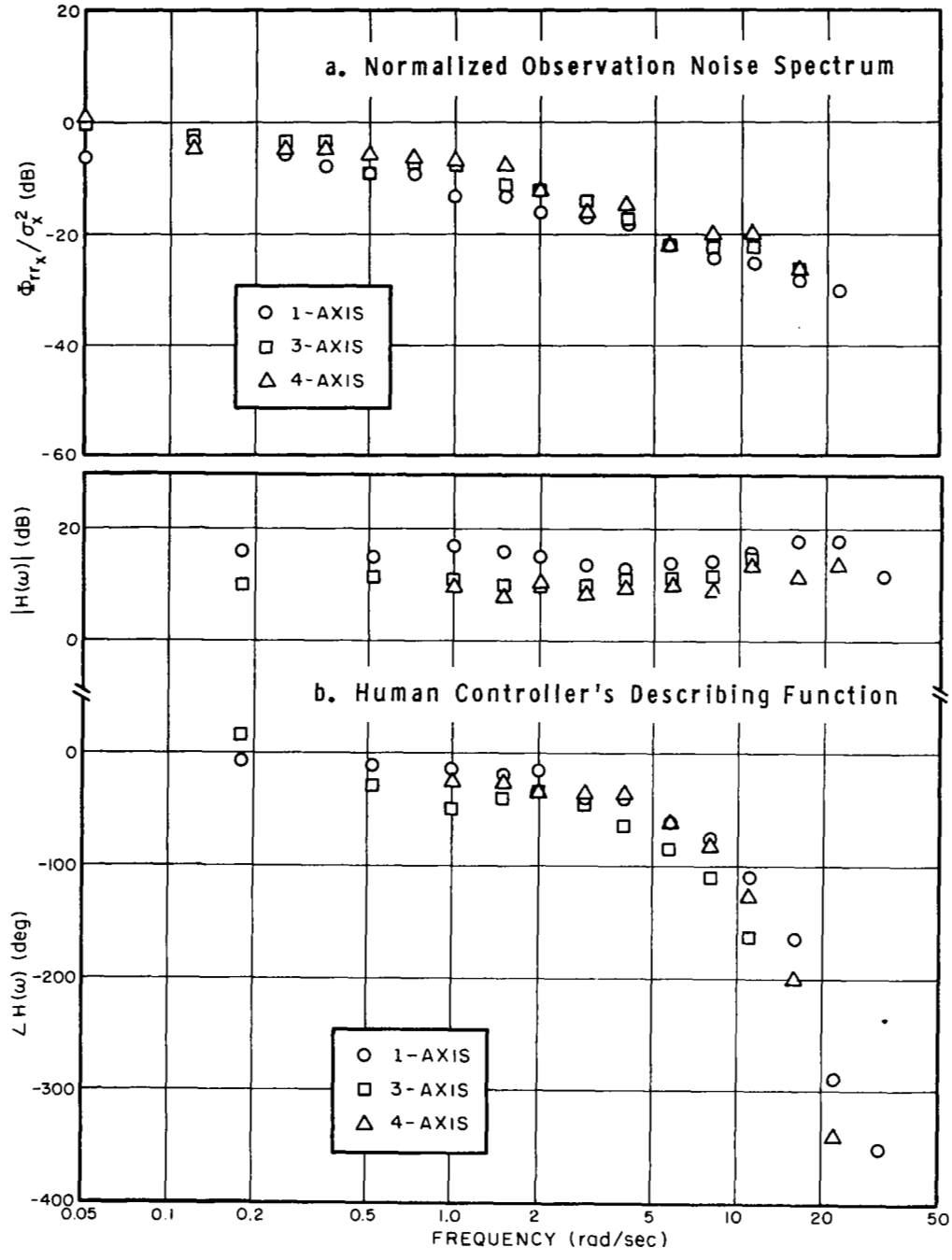


FIG.B-6 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: 16° PERIPHERAL VIEWING WITH REFERENCE EXTRAPOLATION
Subject JF, average of 2 trials

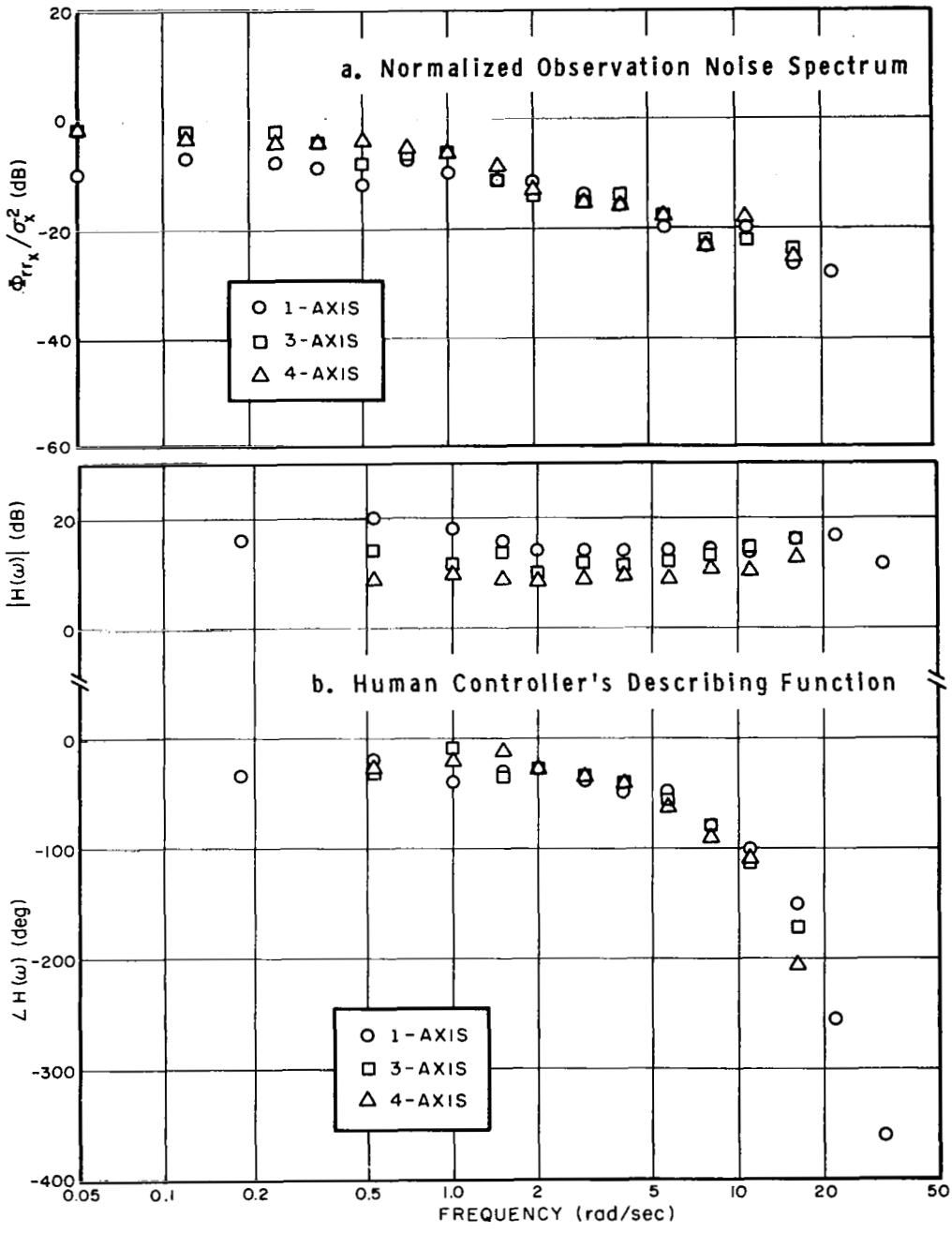


FIG.B-7 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: 16° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION
Subject JF, average of 2 trials

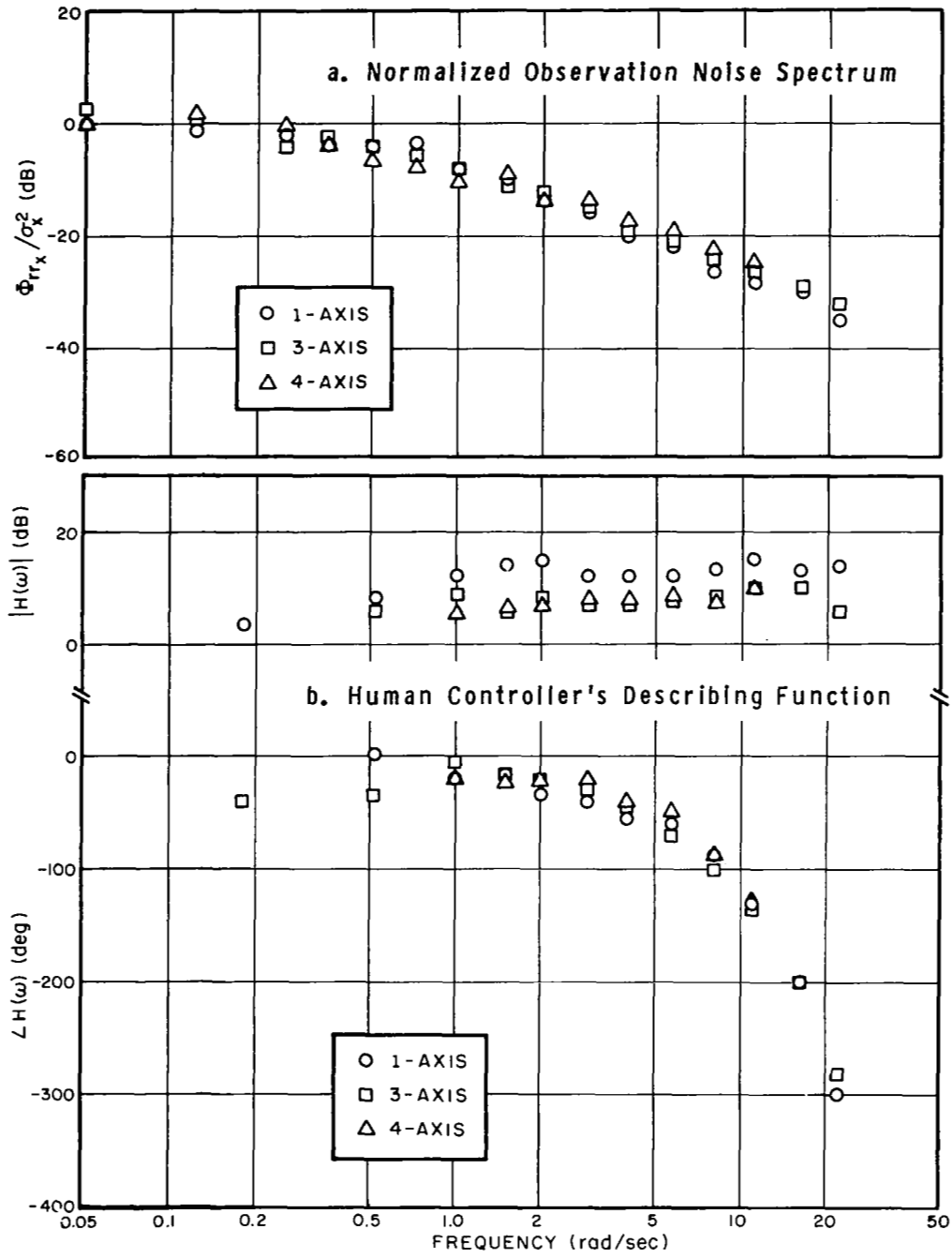


FIG.B-8 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: 22° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION
Subject JF, average of 2 trials

MULTIAXIS TRACKING PERFORMANCE WITH VISUAL SCANNING

Experimental Conditions

The object of this experiment was to provide manual control data against which to test models of human scanning behavior. The subjects were required to track all four axes and were allowed to choose their own visual scanning pattern. Two variations of this experiment were performed. The mean-squared inputs were adjusted equally during the first phase in order to provide a control environment as homogeneous as possible. The MS input was boosted by a factor of 4 on the lower left (LL) axis during the second phase in an attempt to force the subjects to change their scanning pattern. The subjects were instructed throughout this experiment to minimize the total MSE score (given as the sum of the four component MSE scores).

In addition to the scanning experiments, a set of single-axis experiments was performed to provide the complete set of single-axis observation noise measurements required to test models of multi-axis control and scanning behavior. As a secondary benefit, these single-axis trials provided measurements which allowed us to determine the extent to which observation noise scales with input power. The high-input axis and one of the low-input axes (the LR axis) were each tracked individually under each of the four viewing conditions.

The vehicle dynamics were K/s on all axes for the single- and multi-axis experiments, and all input spectra simulated a first-order noise process having a break frequency at 2 rad/sec.

Error and Control Scores

a. Single-Axis Measures.--Average error variance scores obtained from the single-axis experiments are shown in Table B-12. The right column of this table shows the ratio of the high-input score to the low-input score. If the observation noise is basically multiplicative in nature, and if the controller is otherwise linear, the error score should scale with mean-squared input: in this case, by a factor of 4.

Table B-12 shows that only the foveal score scaled approximately with input power. This score increased by a factor of 3.4, as opposed to a factor of 4 for the input. The score corresponding to foveal viewing with reference extrapolation increased by only a factor of about 2, and the remaining two scores corresponding to peripheral viewing with no reference extrapolation actually decreased slightly with an increase in input power. (The decrease in the latter two scores was not statistically significant, whereas the foveal differences were significant at the 0.01 level and the peripheral-with-reference-extrapolation differences were significant at the 0.05 level.) These results suggest that the observation noise processes corresponding to peripheral viewing conditions do not scale with input and therefore cannot be ascribed solely to a multiplicative process acting on the displayed signal. This conclusion is verified further on in this section through examination of the observation noise spectra.

b. Multi-Axis Measures.--Mean-squared error and control scores obtained from the 4-axis experiments are presented in Table B-13. Scores are given for each of the four axes and for the "total score," defined as the sum of the scores across axes. Because of inter-subject differences, particularly with regard to scanning behavior, average performance levels for each subject are shown along with the average for all performance measures computed from the multi-axis data

TABLE B-12

Effect of Mean-squared Input on Error Variance: Single-axis Tracking

Viewing Conditions	MS Input $(deg/sec)^2$		HI/LO Ratio
	0.22	0.87	
Foveal	.13	.44	3.4
16° periph, ref ext	.38	.71	1.9
16° periph, no ref ext	1.0	.88	.88
22° periph, no ref ext	1.3	1.2	.90

MS error variance scores in deg^2 visual arc
Average of 3 subjects, 1 run/subject

TABLE B-13

Effect of Mean-squared Input on Mean-squared Error and Mean-squared Control Scores

Four Axes Tracked Simultaneously with Visual Scanning Allowed

	Homogeneous MSI					Non-Homogeneous MSI				
Subject	UL	LL	LR	UR	Total	UL	LL	LR	UR	Total

a. Mean-squared Error Scores (degrees^2)

JF	.54	.63	.51	.65	2.3	.65	1.5	.61	1.0	3.8
DM	.57	.75	.60	.43	2.3	.60	1.4	.62	.54	3.2
GP	.68	.77	.78	1.1	3.4	.79	1.5	.89	1.1	4.3
HS	.64	.61	.49	.62	2.4	.56	1.2	.57	.56	2.9
Average	.61	.69	.59	.70	2.6	.65	1.4	.67	.81	3.5

b. Mean-squared Control Scores [(deg/sec)^2]

JF	6.2	7.8	5.3	6.6	26	7.6	22	5.8	7.5	43
DM	2.8	4.5	3.5	3.4	14	3.5	22	3.5	3.5	33
GP	6.7	7.8	6.0	13	33	6.5	23	5.4	7.4	42
HS	3.4	5.8	4.4	7.1	21	2.7	18	3.9	5.4	30
Average	4.8	6.5	4.8	7.2	24	5.1	22	4.6	5.9	37

Error scores is degrees^2 visual arc
Average of 4 runs/subject

On the average, the subjects achieved about the same MSE scores on each of the four axes when the inputs were homogeneous. The largest spread in scores was between the UR and LR axes – a difference of about 17 percent. An analysis of variance of the scores, which is summarized in Table D-13 of Appendix D, showed that this difference was not statistically significant.* Inter-axis differences among control scores was proportionally greater, with the UR score about 1.5 times the LR score. This difference was significant at the .001 level.

Table D-13 shows that there were significant axis-by-subject interactions in these scores. We can observe this interaction from inspection of the data tabulated in Table 18. Subject DM, for example, achieved the lowest MSE score on the UR axis and the highest score on the LL axis, whereas subject JM achieved the highest score on the UR axis and the lowest on the LR axis.

Inter-axis differences were much more pronounced in the non-homogeneous-input experiment. As expected, the largest score was obtained on the LL axis (the axis with the higher input power). The score on this axis was about twice the score that was achieved on this axis when the inputs were homogeneous. Table D-13 shows that the LL and UR MS error scores were both significantly different from the LR score, whereas the UL and LR scores did not differ significantly. The error scores revealed no significant axis-by-subject interactions.

*We adopted the 0.05 significance level as the criterion level of "statistical significance"; i.e., we reject the null hypothesis if the significance level is numerically less than 0.05.

A large and statistically significant difference was found between the control scores on the LL and LR axes, with the LL score about 5 times the LR score. Because of significant axis-by-subject interactions manifested by the control scores, the remaining inter-axis differences failed the test of significance.

The increase in the input power on the LL axis caused all MS error scores to increase on the average. The smallest increase was observed on the UL axis (a factor of about 1.07), and the largest increase occurred on the LL axis (a factor of about 2). The total score increased by about 1.3. Analysis of variance tests, summarized in Table D-14, show that the increases on all axes except the UL were statistically significant. The lack of significant axis-by-subject interactions indicates that these trends were consistent across subjects.

Raising the input power had different effects on the MS control scores. The only significant increases were seen on the LL axis and on the total score. The LL score increased by a factor of about 3.4 (compared to a factor of 2 increase for the corresponding error score). The scores on the UL and LR axes remained about the same, on the average, and the UR score actually decreased to about 0.8 times the homogeneous-input score. Significant axis-by-subject interactions were found on all but the LR axis. Thus, the changes in control scores caused by the increase of input power were more subject-dependent than the corresponding changes in error scores.

The mean-squared error scores indicate that the subjects received relatively better information on the LL axis when the mean-squared input on that axis was increased from 2.2 to 8.7 $(deg/sec)^2$. Had the subjects tracked with the same efficiency on that axis for both input conditions, the MS error score would have increased by

a factor of 4, rather than the factor of about 2 that was observed experimentally. The relative improvement in tracking performance on the LL axis could result from increased foveal attention to that axis, from relatively improved observational characteristics, or from a combination of the two. Examination of the eye-movement data and observation noise spectra shows that improved observational characteristics was the dominant factor.

Visual Scanning Behavior

Fractional distributions of fixation time are shown in Table B-14 along with the fractional distribution of mean-squared error (the latter defined as the mean-squared error on a given axis divided by the sum of the error scores on all four axes). We note first of all that the distribution of fixation time was markedly non-homogeneous even when the inputs were homogeneous. The UR display was fixated the most (43% of the time) and the LL display the least (12% of the time). The upper two displays as a group were fixated 71%, whereas the lower two received only 29% foveal attention.

If the control task were truly homogeneous, then an uneven distribution of fixation times would result in an uneven distribution of mean-squared error. Specifically, since the subjects fixated the upper displays 71% of the time, we would expect most of the error to occur on the lower two displays. This was true only for subject DM. On the average, the total MSE score was distributed nearly evenly among the four axes. On this basis we infer that the subjects must have received better visual information when fixating the upper displays than when fixating the lower ones. This conclusion is supported by the single-axis results, summarized in Table B-4, which show that the total score corresponding to fixation of the upper displays was significantly lower than the score corresponding to fixation of the lower displays.

TABLE B-14

Effect of Mean-squared Input on Fractional Distributions of Fixation Time and Mean-squared Error

	Homogeneous MSI				Non-Homogeneous MSI			
Subject	UL	LL	LR	UR	UL	LL	LR	UR

a. Fractional Distribution of Fixation Time

JF	.36	.13	.15	.37	.31	.22	.13	.34
DM	.20	.13	.26	.41	.18	.18	.21	.44
GP	.34	.09	.15	.42	.39	.11	.06	.44
HS	.23	.12	.13	.52	.38	.08	.09	.45
Average	.28	.12	.17	.43	.32	.15	.12	.42

b. Fractional Distribution of Mean-squared Error

JF	.23	.27	.22	.28	.17	.40	.16	.26
DM	.24	.32	.26	.18	.19	.45	.20	.16
GP	.21	.23	.23	.34	.18	.35	.21	.26
HS	.27	.26	.21	.26	.19	.42	.20	.19
Average	.24	.27	.23	.27	.18	.40	.19	.22

Increasing the input power on the LL display had a negligible effect on the average scanning behavior. Three of the subjects spent slightly more time fixating the LL display, and one subject decreased his fixation on that axis. The net effect was to increase the fraction of fixation time on the LL axis from 0.12 to 0.15. This increased attention came entirely at the expense of the LR display, where fractional fixation time decreased from 0.17 to 0.12. As we showed earlier, the increase in input power produced the expected trend in the distribution of mean-squared error scores. Table B-14b shows that the fraction of total score attributable to the LL axis rose to 0.40, with about 20% of the total error occurring on each of the remaining axes.

The failure of the change in input power to draw more foveal attention to the LL axis suggests that the subject receives relatively better peripheral information about that axis when the input power is increased. The behavior of the single-axis mean-squared error scores summarized in Table B-12 supports this conclusion, as does the analysis of the observation noise spectra presented later in this chapter.

Mean observation times and fixation frequencies for each display are given in Table B-15. Also shown is the overall scan frequency, which is defined as the number of transitions of the fixation point per second and is equal to the sum of fixation frequencies over the four displays. The interaxis differences in mean observation time and fixation frequency followed the same trend, on the average, as the fractional allocation of fixation time for both variations of the experiment. For the homogeneous-input situation the largest value of each of these three measured quantities was seen on the UR axis. The rank ordering of the remaining axes, from largest to smallest values, was UR, UL, LL, LR for all three

TABLE B-15

Effect of Mean-squared Input on Mean Observation Time, Fixation Frequency, and Scanning Frequency

Subject	Homogeneous MSI				Non-Homogeneous MSI			
	UL	LL	LR	UR	UL	LL	LR	UR

a. Mean Observation Time (sec)

JF	.80	.51	.57	1.02	.81	.85	.68	1.03
DM	.47	.39	.46	.69	.41	.60	.48	.70
GP	.78	.65	1.00	.86	.64	.61	.53	.62
HS	.65	.45	.42	1.00	1.58	.73	1.02	1.48
Average	.68	.50	.61	.89	.86	.70	.68	.96

b. Mean Fixation Frequency (looks/sec)

JF	.44	.24	.25	.36	.39	.26	.20	.33
DM	.44	.35	.56	.64	.43	.31	.44	.63
GP	.44	.14	.15	.48	.61	.18	.11	.70
HS	.33	.26	.29	.50	.21	.11	.09	.30
Average	.41	.22	.31	.50	.41	.22	.21	.49

c. Scan Frequency (looks/sec)

JF	1.29	1.18
DM	1.99	1.81
GP	1.21	1.60
HS	1.38	0.74
Average	1.47	1.33

quantities. The overall average scan frequency decreased from 1.47 to 1.33 looks/sec when the LL input was increased. The reader should note that there were wide inter-subject variations in the measurements relating to scanning behavior, as shown in Tables B-14 and B-15.

Increase of the LL input power produced an increase in mean observation time on all axes. The largest relative increase occurred on the LL axis (from 0.50 to 0.70 sec – a factor of 1.4), whereas the lowest relative increases occurred on the UL and LR axes (about a factor of 1.1). The mean fixation frequencies were basically unchanged for three axes and decreased by about 1/3 on the LR axis.

Observation Noise Spectra and Describing Functions

a. Single-Axis Measures.--We inferred from the single-axis mean-squared error scores that observation noise power was not linearly related to input power under peripheral viewing conditions. Examination of the observation noise spectra confirms this conclusion. Figure B-9 compares, for each viewing condition, the observation noise spectra obtained with mean-squared inputs of 2.2 and 8.7 $(deg/sec)^2$. These spectra are un-normalized and have units of error power per rad/sec.

If the observation noise power were to scale with input power, we would expect the spectrum corresponding to the larger input to be 6 dB above the spectrum corresponding to the lower input for each viewing condition. Figure B-9a shows that this relation held for foveal viewing. On the other hand, Figs. B-9b through B-9d, which show the peripheral observation noise spectra, reveal no consistent differences caused by the change in input power. It thus appears that the primary effect of placing a display in the periphery is to introduce an observational noise process that is fundamentally an additive process; i.e., one that does not scale with signal power.

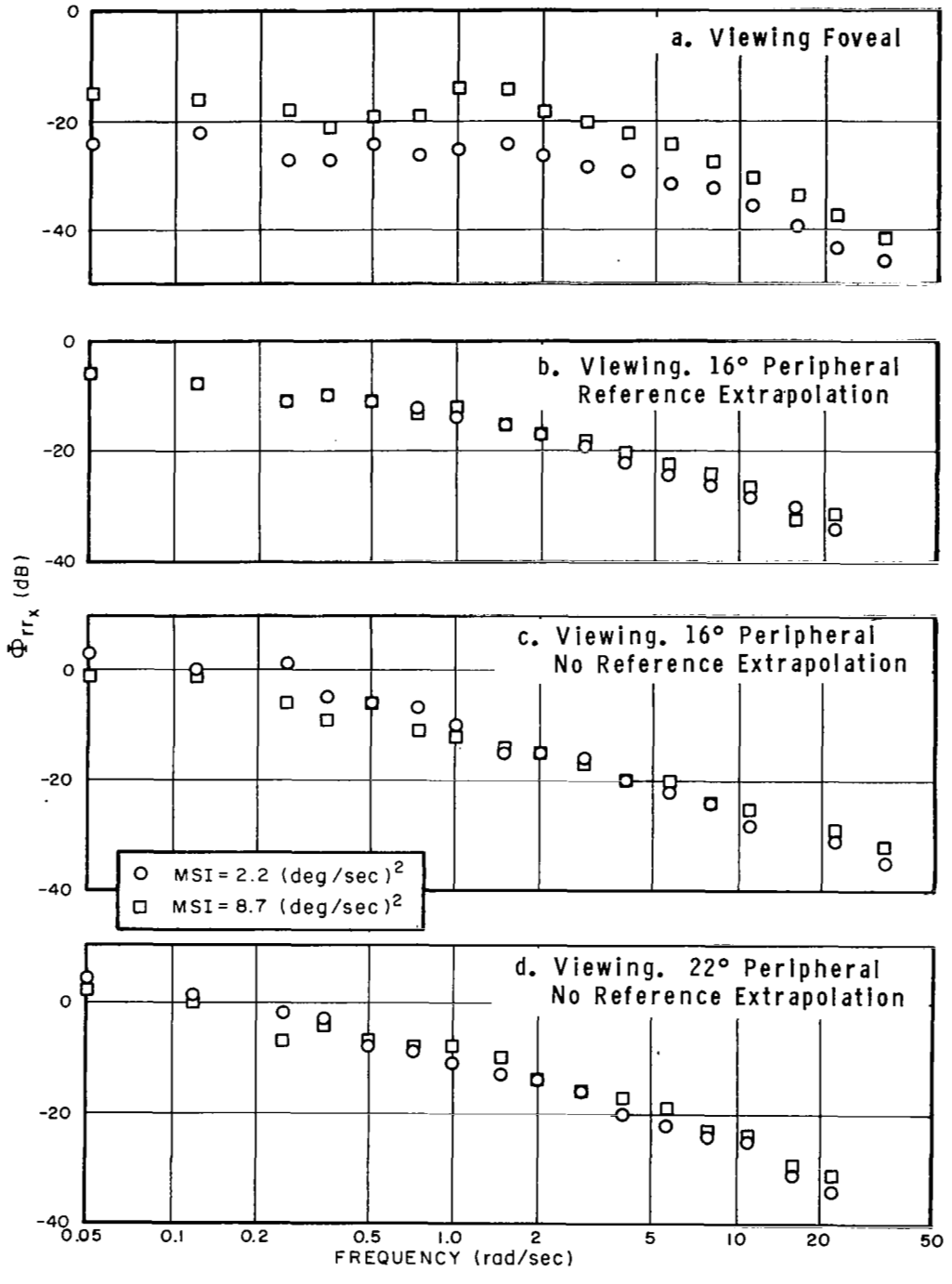


FIG.B-9 EFFECT OF MEAN-SQUARED INPUT ON THE SINGLE-AXIS OBSERVATION NOISE SPECTRUM
Average of 3 subjects, 1 trial/subject

The observation noise spectra are shown again in Fig. B-10 to illustrate directly the differences between foveal and peripheral observation noise spectra. Spectra corresponding to an axis having the smaller mean-squared input are shown in Fig. B-10a; spectra from the axis with the high input level are given in Fig. B-10b. A comparison of these two sets of curves reveals that foveal-peripheral differences, especially at low frequencies, were substantially greater on the axis with the lower input. We interpret this result as indicating that the controller has relatively more to gain by looking at one of the low-input axes foveally than by fixating the high-input axis. (Equivalently, he has more to lose by observing the low-input axis peripherally than by observing the high-input axis peripherally.) The fact that the subjects did not reduce their fractional fixation time on the LL axis when the MS input to that axis was increased indicates that the improved peripheral information on that axis was offset by the increased importance of that axis with respect to its contribution to the total performance score.

The trend in the observation noise measurements was verified by the subjective impressions of the subjects. Without first informing them of the results of our eye-movement measurements, we asked two of the subjects whether or not they spent more time fixating the LL axis when the input was increased. They replied that they did not, because (a) they could see the signal on the LL display well enough peripherally when the input was increased, and (b) when they did fixate the LL display, they had difficulty estimating the signals on the remaining displays.

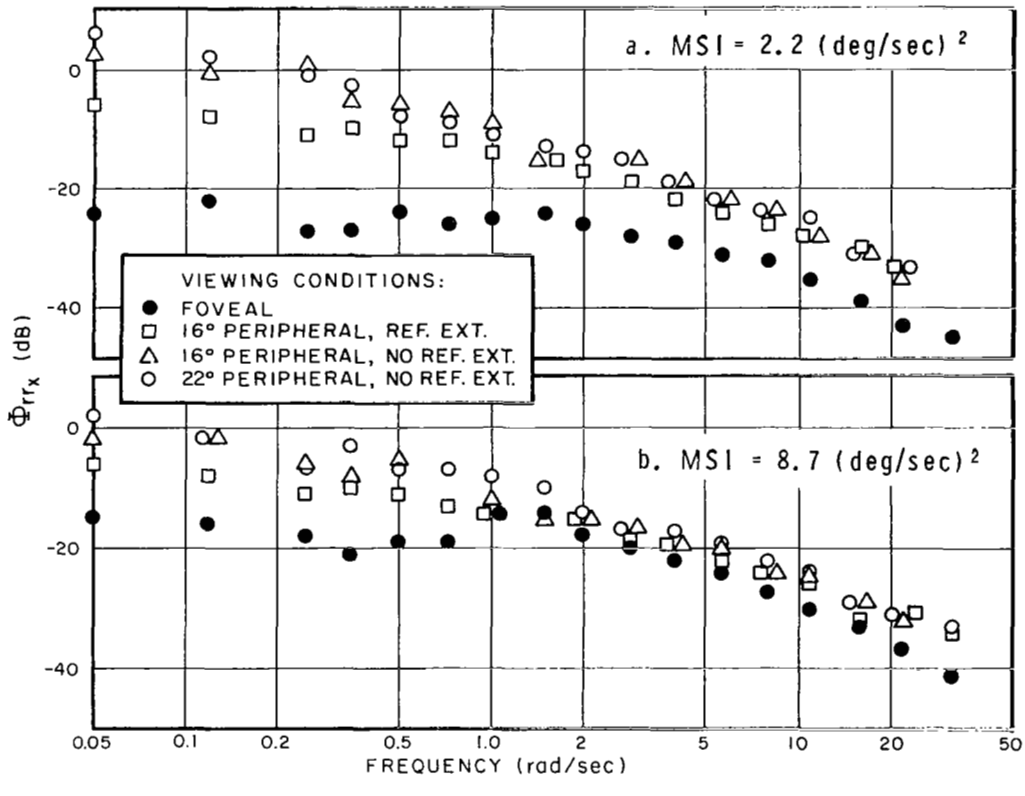


FIG.B-10 EFFECT OF VIEWING CONDITIONS ON THE SINGLE-AXIS OBSERVATION NOISE SPECTRUM

Average of 3 subjects, 1 trial/subject

The effects of viewing conditions on the pilot's describing function are shown in Fig. B-11. Because of the similarity of the describing functions corresponding to peripheral viewing of a given display, only the 22° peripheral describing function is compared with the foveal describing function for each of the two MSI conditions. Foveal-peripheral differences differ somewhat for the two input conditions. As we would expect from the trend in the observation noise spectra, the overall difference between foveal and peripheral gain was slightly less on the high-input axis. On the other hand, foveal-peripheral differences in phase shift appear to be about the same on both axes over most of the measurement band.

b. Multi-Axis Measures.--We have performed only a preliminary analysis of the observation noise spectra and human controller describing functions for the 4-axis scanning experiments. In Fig. B-12a we present a sample normalized observation noise spectrum which represents the average of the spectra measured on the LL and UR axes during the 4-axis scanning experiment with homogeneous inputs. Averages have been taken over subjects JF and HS. Single-axis observation noise spectra corresponding to foveal and 22° peripheral viewing are shown for comparison. A similar single-axis, multi-axis comparison of the controller describing functions is shown in Fig.B-12b. The shape of the multi-axis observation noise spectrum appears to be similar to the single-axis foveal spectrum, with perhaps a somewhat lower break frequency, and has an asymptotic low-frequency level that is about midway between the foveal and peripheral levels. The multi-axis describing function, on the other hand, resembles very closely the single-axis peripheral results. Comparison of the multi-axis results to the single-axis foveal results reveals a consistent trend; that is, an increase in observation noise is accompanied by decrease in average controller gain.

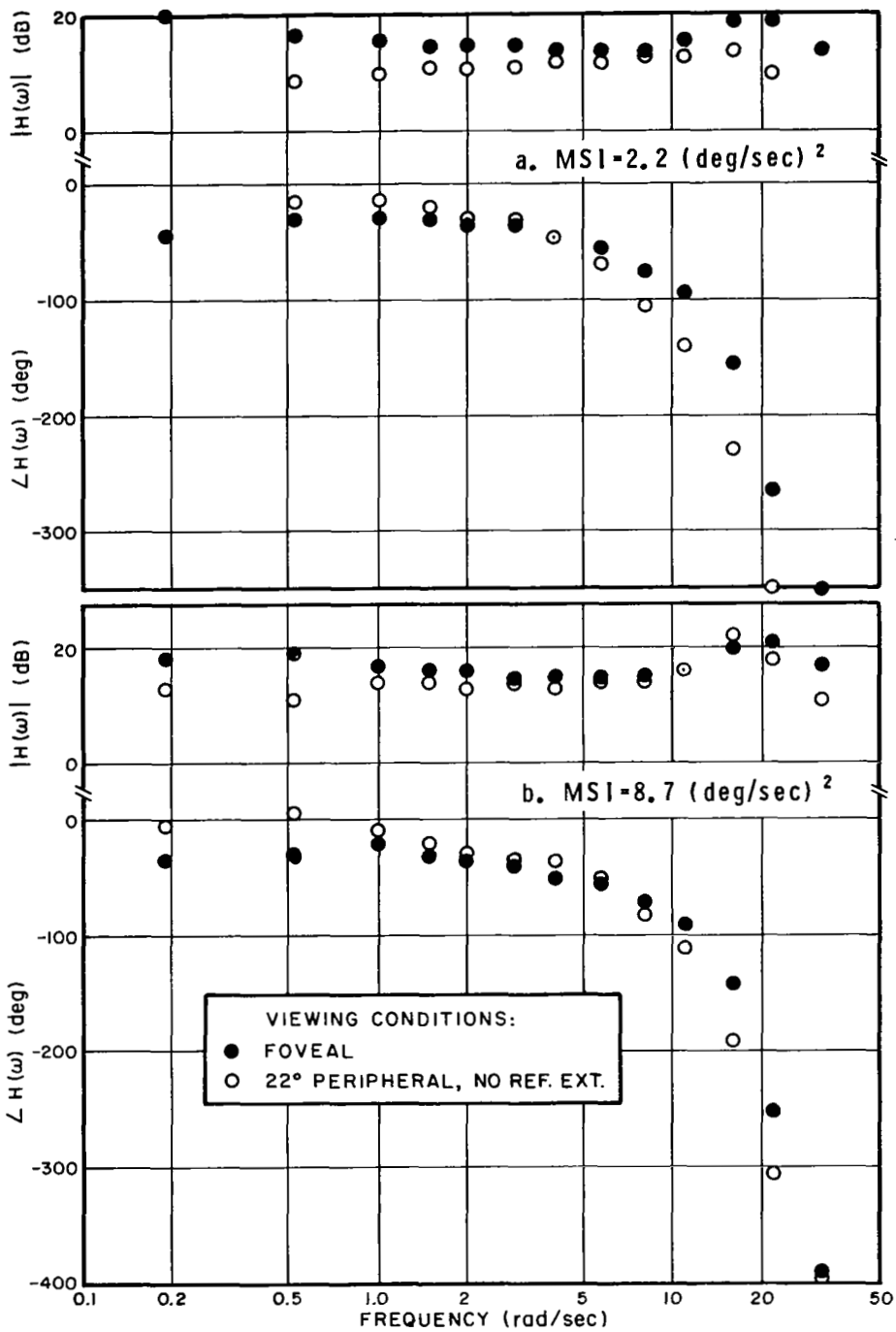


FIG.B-11 EFFECT OF VIEWING CONDITIONS ON THE SINGLE-AXIS HUMAN CONTROLLER'S DESCRIBING FUNCTION
Average of 3 subjects, 1 trial/subject

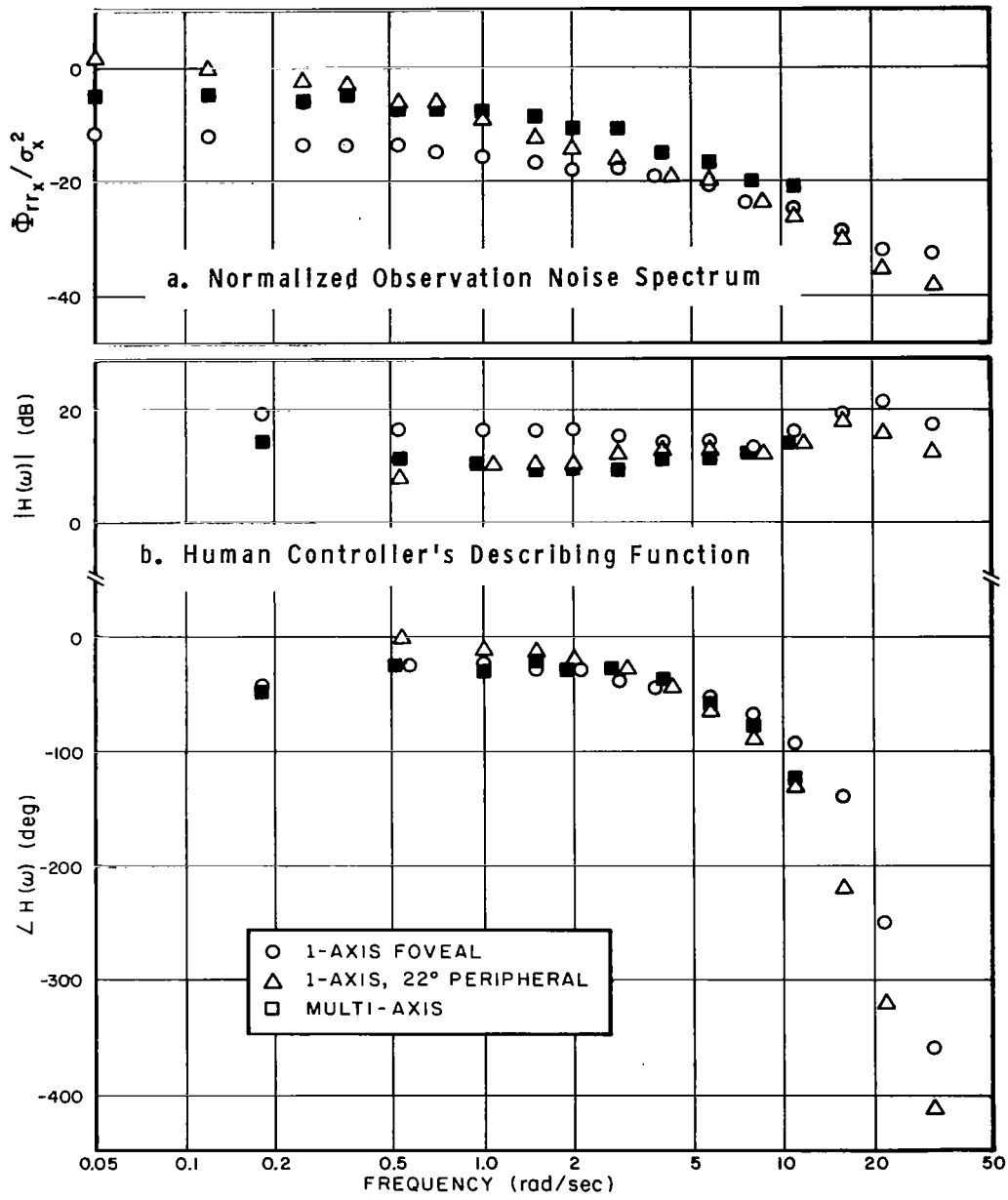


FIG.B-12 COMPARISON OF SINGLE-AXIS AND MULTI-AXIS PERFORMANCE MEASURES

Average of 2 subjects, 2 trials/subject

SUMMARY OF RESULTS

An initial experiment was performed to determine the relative merits of an error bar, reference bar-type of display presentation versus an error dot, reference circle presentation. Mean-squared error scores were found to be significantly less for the bar presentation, especially when the display was viewed peripherally. Consequently, the bar presentation was used throughout the remainder of the experimental program.

A full set of single-axis trials was performed for each subject. The rank ordering of task difficulty, as indicated by the mean-squared error score, was, from easiest to most difficult: (a) the foveal task, (b) 16° peripheral viewing with the ability to extrapolate a zero reference, (c) 16° viewing with no reference extrapolation possible, and (d) 22° viewing with no reference extrapolation. Overall, the performance score was less when the upper displays were fixated than when the lower displays were fixated, indicating that better visual information could be obtained when the subject fixated the upper displays. Appreciable nonzero mean error was measured when peripheral displays were tracked without benefit of reference extrapolation, presumably because the baseline present on the peripheral display tended to disappear shortly after peripheral viewing was initiated.

Changes in the controller's equivalent normalized observation noise spectrum corresponded to changes in the error scores; the low-frequency portion of the spectrum increased as the task difficulty was increased. Controller gain decreased as the display was moved from the fovea to the periphery, and the phase lag at high frequencies increased. There were negligible differences among the describing functions corresponding to the three peripheral viewing conditions.

A set of 2-, 3-, and 4-axis experiments without visual scanning all showed task interference. The two-axis foveal-peripheral task showed an increase in total score of about 1.7 when the two axes were tracked simultaneously. Slightly more than half of this increase occurred on the peripheral axis. Observation noise increased about 3-4 dB on both axes, and both controller gains decreased by about 2-3 dB. Since the input levels were adjusted so that the foveal and peripheral tasks would contribute approximately equally to the total score, we conclude from these results that there is no fundamental difference between interference on a foveal task and interference on a peripheral task. The interference was significantly greater when the two axes had to be controlled by a single hand than when two hands were used.

The three- and four-axis tasks also exhibited interference, with the largest percentage increase in MSE score per axis decreasing with the relative contribution of the component score to the total performance measure. (No attempt was made to equalize the component scores in these experiments.) Similarly, the observation noise level increased, again with the largest increase occurring on the axis which contributed least to the total performance score (i.e., the foveal axis). Controller gain also decreased, but the decrease was approximately the same on all axes. The apparent tendency of the subject to minimize the effects of interference on the more important tasks suggests a phenomenon akin to an optimal allocation of attention among the four axes.

Two sets of four-axis experiments with visual scanning permitted were conducted: one with the mean-squared inputs the same on all axes and one with the lower-left (LL) mean-squared input quadrupled. Nonhomogeneous distribution of fixation time was observed under homogeneous as well as nonhomogeneous input conditions.

Even when the inputs were homogeneous, the subjects fixated the upper two displays collectively about 71% of the time; yet, the mean-squared error scores were about the same on each axis. This scanning behavior is consistent with the hypothesis, based on the single-axis MSE scores, that the subjects obtain better visual information overall when fixating the upper displays than when fixating the lower ones.

Increasing the LL mean-squared input caused the error scores on all axes to increase, with the largest increase occurring on the LL axis, as expected. The average scan frequency decreased slightly, with the result that mean observation times were increased from 10 to 40 percent. The increase in input power did not, however, cause the subject to devote appreciably more time fixating the LL axis.

A set of single-axis experiments was performed, one with the lower mean-squared input and one with the larger, to complete the set of measurements needed for a model of multi-axis control behavior and also to investigate the effects of input power on the observation noise spectrum. The resulting MSE scores, observation noise measurements, and controller describing functions all show that the subjects received better peripheral information when the MSI was increased. The relative improvement in peripheral information was apparently enough to offset the increase in importance of the LL task when the LL input was increased so that there was little change in the controller's scanning behavior. The failure of the peripheral observation noise levels to increase with increasing MSI suggests that peripheral noise represents a process that is fundamentally an additive process (unlike foveal observation noise, which appears to be basically multiplicative).

APPENDIX C

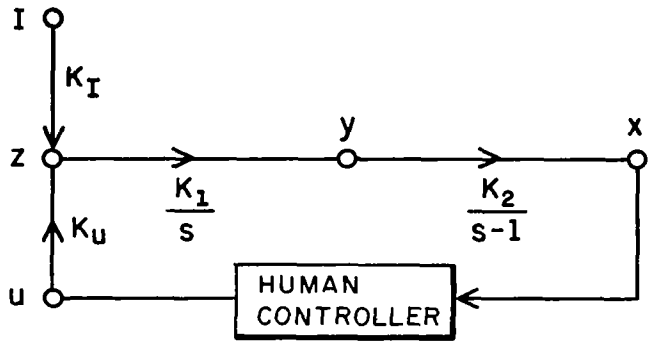
APPENDIX C. EFFECT OF VEHICLE DYNAMICS ON OBSERVATION NOISE RATIO

Our experiments with vehicle dynamics of K, K/s, and K/s^2 consistently have yielded observation noise ratios on the order of -20 dB for foveal viewing of the display (see Refs. 9 and 6 and Section 3 of this report). The observation noise ratio of -26 dB obtained with the unstable dynamics $K/s(s-\lambda)$, (see Section 4) was inconsistent with these previous results. In an attempt to determine why the lower observation noise ratio associated with the unstable dynamics was so much smaller than expected, we conducted an experiment in which the same set of subjects tracked unstable and stable vehicle dynamics in three different experimental trials. The vehicle dynamics were of the form

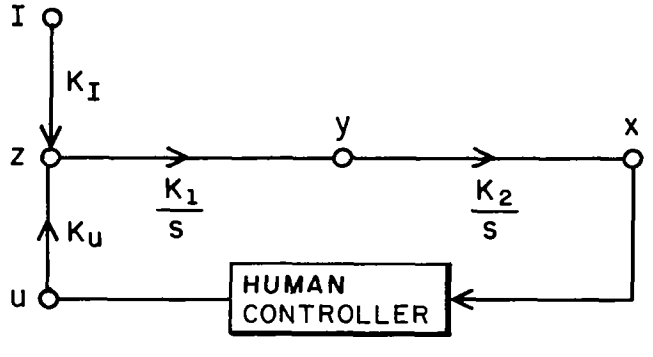
$$V = \frac{K}{s(s-\lambda)} \qquad \text{(C-1)}$$

Experimental variables were: (a) the degree of vehicle instability λ and (b) the point of application of the input forcing function. The subject was displayed a single quantity, the system error x(t), throughout this set of experiments.

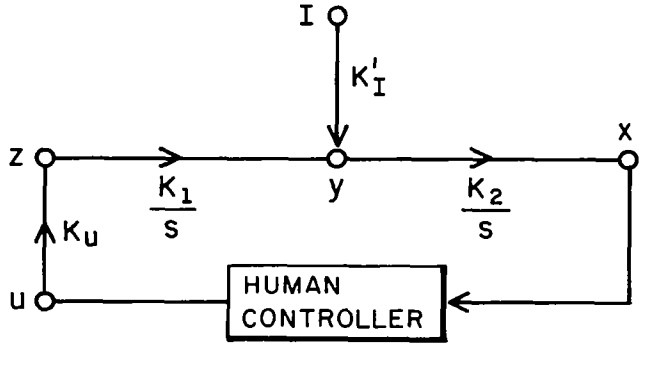
Diagrams of the three experimental tasks are shown in Fig. C-1; experimental parameters are tabulated in Table C-1. Task A was a replication of the 1-indicator display condition discussed in Section 4 except for a change in control sensitivity that is discussed below. Task C replicated a control situation studied previously (Ref. 6) in which the vehicle dynamics were K/s^2 and the input was applied as a disturbance to vehicle velocity, rather than to vehicle acceleration. This task is representative of those that yielded -20 dB observation noise ratios. Since tasks A and C differed two ways (instability and input injection point), we needed an additional



A. UNSTABLE VEHICLE, ACCELERATION DISTURBANCE



B. STABLE VEHICLE, ACCELERATION DISTURBANCE



C. STABLE VEHICLE, VELOCITY DISTURBANCE

FIG.C-1 DIAGRAMS OF THREE EXPERIMENTAL TASKS

TABLE C-1

Parameter Values for the Second Experiment

Parameter	Experimental Condition		
	A	B	C
K_1[(arc-degrees/sec)/volt]	.200	.200	.200
K_2^2[(degrees/sec)/degree]	2.50	2.50	2.50
K_I[dimensionless]	.157	.240	--
K_I'[degrees/volt]	--	--	.0200
K_u[volts/newton]	5.64	5.64	5.64
$K_u \cdot K_1 \cdot K_2$[(deg/sec^2)/newton]	2.82	2.82	2.82
σ_1^2(volts)	8.72	8.72	8.72

Experimental Conditions:

A. V=K/s(s-1), acceleration input

B. V=K/s^2, acceleration input

C. V=K/s^2, velocity input

intermediate task which would differ from each of these two in only one respect so that we could localize the cause of any performance differences that might be observed. Hence, we chose as task B, a neutrally stable vehicle (K/s^2) with the input applied as an acceleration disturbance. (We did not investigate the alternate intermediate condition of unstable dynamics with a velocity disturbance.)

Initially, we attempted to use the same control sensitivity that was used in the prior experiment with unstable control dynamics, but the subjects complained that the system was not sufficiently sensitive for adequate control of condition C. Accordingly, the overall control sensitivity (the product of the parameters K_u, K_1, and K_2) was increased from 1.0 to 2.82 $(deg/sec^2)/newton$.

Input signals having spectral characteristics identical to those used in the prior experiment with unstable dynamics were used throughout this experiment. The effective input power (determined primarily by the input gain parameter K_I or K_I') was varied from one condition to the next in order to maintain a consistent level of apparent task difficulty as measured by the error variance. In order to select appropriate values of input gain, we obtained first the principal model parameters (the relative control-rate weighting, the time delay, and the noise ratios) by matching the results of the 1-indicator experiment described in Section 4. Keeping these parameters constant, we simulated the experimental conditions for tasks B and C and determined the input gains that would yield error variance scores that were constant across the three experimental conditions.

The same four subjects who participated in the prior unstable dynamics experiment participated in this validation experiment as well. Since tasks B and C were new to three of the subjects, most

of the training was concentrated on these tasks. A few sessions with task A allowed the subjects to maintain their skill at tracking the unstable vehicle. An equal number of trials per condition was provided during experimentation (each subject performed each task four times), and the tasks were presented in a balanced order.

EXPERIMENTAL RESULTS

Optimal-control model behavior was matched to the average behavior of the subjects for each experimental condition. Two model parameters were held fixed across conditions: time delay was 0.2 sec, and the motor noise ratio was set at about -29.5 dB. The remaining parameters of control-rate weighting and observation noise ratio were selected for each condition to provide the best match to the measured variance scores and frequency-domain characteristics. The primary purpose of this exercise was to provide an estimate of the controllers' observation noise ratio.

Model parameters are tabulated in Table C-2. The major parameter of interest, the observation noise ratio, was essentially the same for the two acceleration-disturbance tasks. Ratios of -25 dB and -26 dB, respectively, were found for the unstable and stable vehicle tasks. This small difference was within the limits of precision of our model-matching procedure. Although a somewhat higher noise ratio, -23 dB, appeared to provide the best overall match to the velocity-disturbance data, a noise ratio of -26 dB provided almost just as good a match. Accordingly, model results using these two noise ratios are given for the velocity-disturbance task.

TABLE C-2

Effect of Experimental Conditions on Model Parameters

Task	Vehicle Dynamics	Type of Input	Relative $\dot{u}$ weighting	Lag Factor (sec)	Time Delay (sec)	Noise Ratio	
						Motor (dB)	Obs (dB)
A	$k/s(s-1)$	Acc.	1.2×10^{-4}	.073	.20	-29.2	-25.0
B	k/s^2	Acc.	1.2×10^{-4}	.079	.20	-29.5	-26.0
C(a)	k/s^2	Vel.	0.6×10^{-4}	.070	.20	-29.4	-23.0
C(b)	k/s^2	Vel.	0.6×10^{-4}	.070	.20	-29.4	-26.0

Model scores are compared with average measured variance scores in Table C-3. (Average scores for individual subjects are given in Tables D-17 of Appendix D.) Matching of experimental results was very good for the two tasks in which the input was applied as an acceleration disturbance. Tables C-3a and C-3b show that all performance scores were matched to within 5% (except for the control-rate score, for which the predicted score was about 30% too high). The performance scores obtained from the velocity-disturbance experiment were not matched so well. Of the two observation noise ratios tested in the model, a consistently better match to the scores was provided by the -23 dB ratio.

Since our primary goal is to infer the subjects' average observation noise ratios, we should place primary emphasis on the match between measured and predicted equivalent observation noise spectra.* Figures C-2 through C-4 show that the observation noise

*For a given control situation, the predicted observation noise spectrum depends primarily upon the noise parameters and much less strongly upon time delay and relative cost weightings. The variance scores, however, are influenced appreciably by all of the model parameters.

TABLE C-3

Comparison of Measured and Predicted Variance Scores

Data	Observation Noise Ratio (dB)	σ_x^2 $(deg)^2$	$\sigma_{\dot{x}}^2$ $(deg/sec)^2$	σ_y^2 deg^2	σ_u^2 $newt^2$	$\sigma_{\dot{u}}^2$ $(newt/sec)^2$
a. Unstable vehicle, acceleration disturbance						
Exptl	---	.18	2.2	.38	9.7	6.0×10^2
Model	-25.0	.18	2.2	.38	9.8	7.8×10^2
b. Stable vehicle, acceleration disturbance						
Exptl	---	.19	2.1	.34	10	5.9×10^2
Model	-26.0	.18	2.2	.35	10	7.3×10^2
c. Stable vehicle, velocity disturbance						
Exptl	---	.30	5.1	.98	18	1.4×10^3
Model (a)	-23.0	.27	3.9	.70	12	1.2×10^3
Model (b)	-26.0	.22	3.6	.66	11	1.1×10^3

Average of 4 subjects, 4 trials/subject

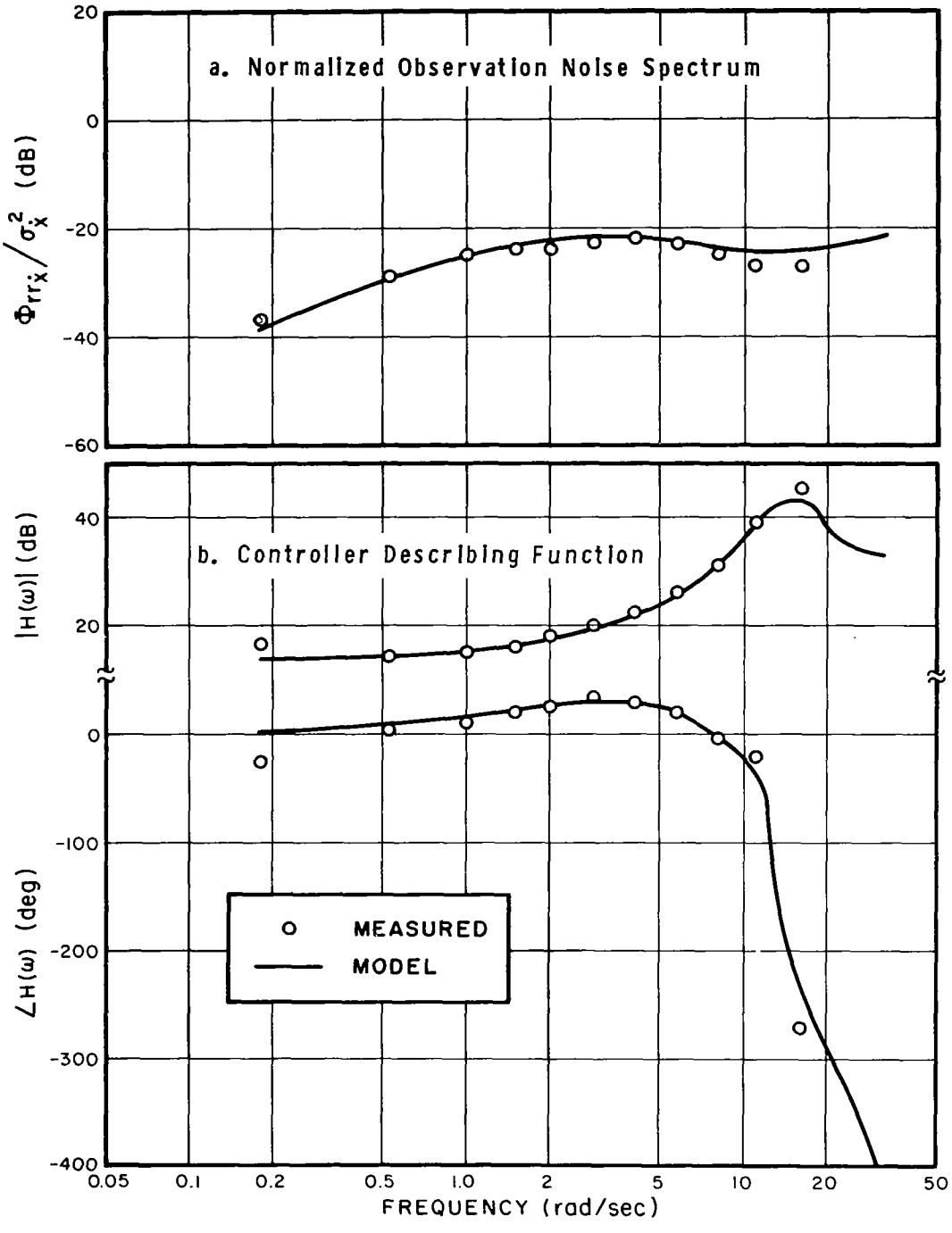


FIG.C-2 COMPARISON OF MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES: UNSTABLE VEHICLE WITH ACCELERATION DISTURBANCE

Average of 4 subjects, 4 trials/subject

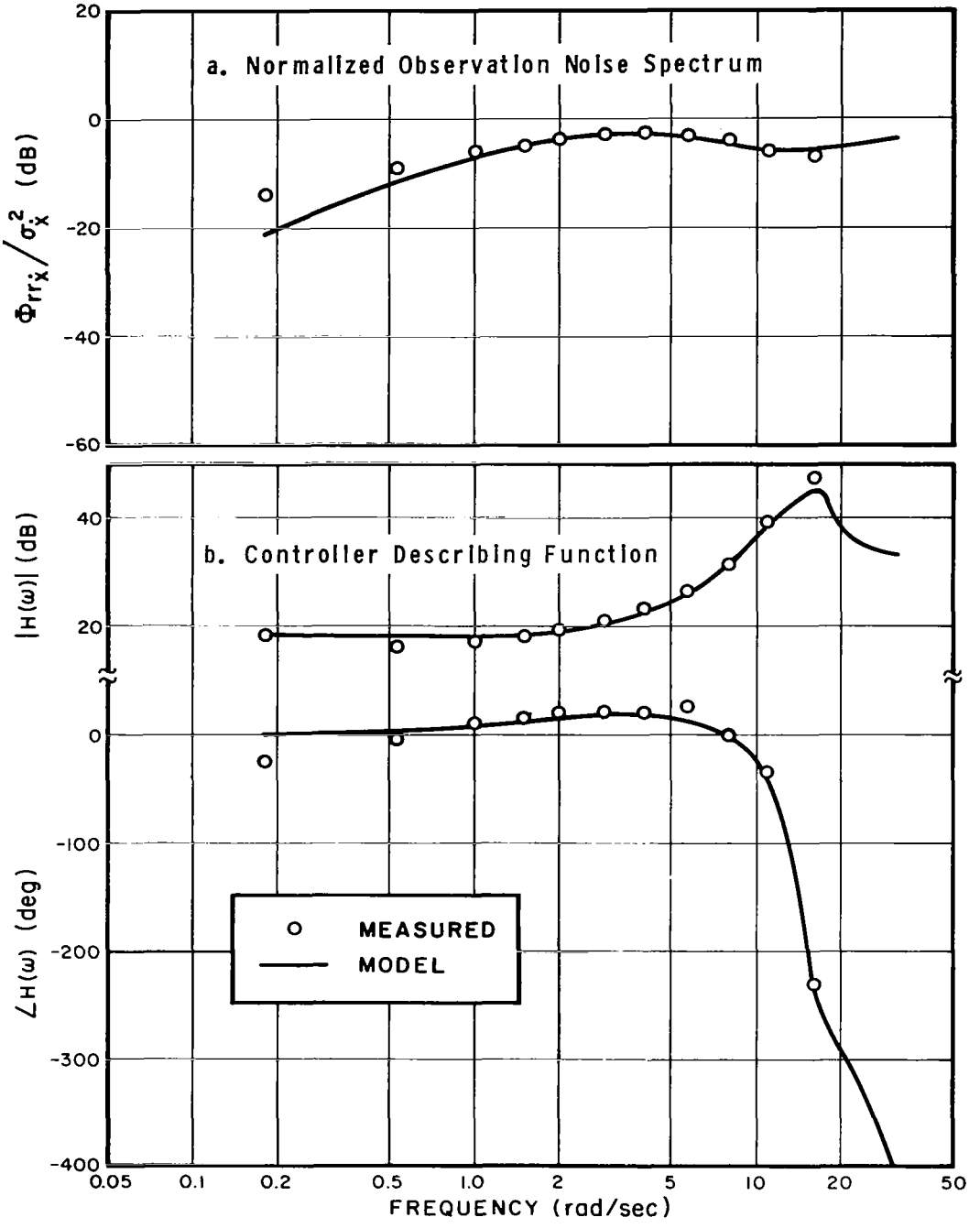


FIG.C-3 COMPARISON OF MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES: STABLE VEHICLE WITH ACCELERATION DISTURBANCE

Average of 4 subjects, 4 trials/subject

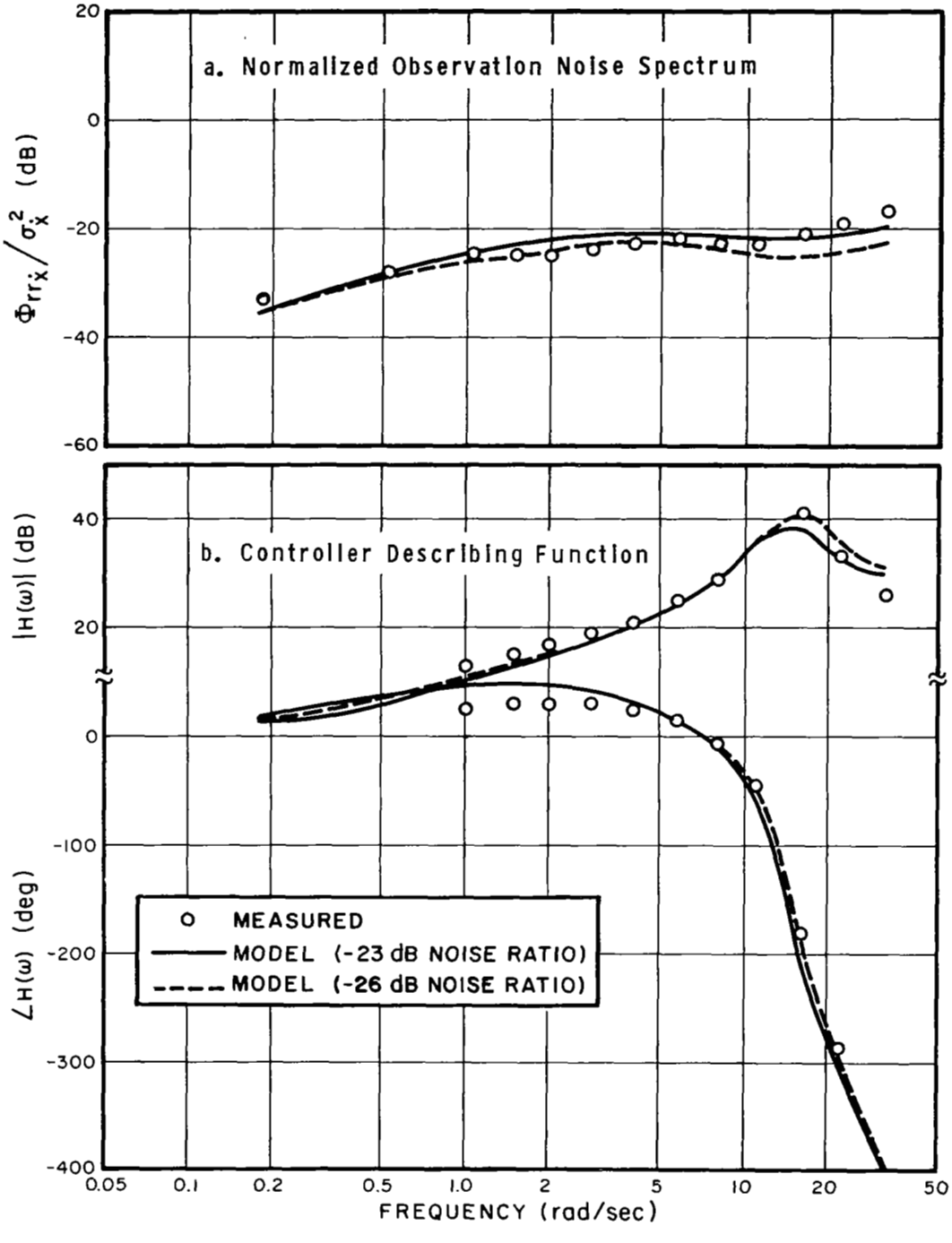


FIG.C-4 COMPARISON OF MEASURED AND PREDICTED FREQUENCY-DOMAIN MEASURES: STABLE VEHICLE WITH VELOCITY DISTURBANCE

Average of 4 subjects, 4 trials/subject

spectra, as well as the controller describing functions, were matched very well for the two experimental tasks in which the input was applied as a disturbance to vehicle acceleration. A reasonably good match was also obtained for the frequency-domain results of the velocity-disturbance experiment. Figure C-4a shows that good matches to the observation noise spectra were provided by the two noise ratios tested in the model; better mid-frequency results were obtained by the lower noise ratio (-26 dB) at the expense of a less good match at high frequencies.

On the basis of these results we conclude that the subjects' equivalent observation noise ratios were substantially the same for the three tasks investigated in this experiment and considerably lower than the nominal values of -20 dB obtained for the observation noise ratios in most of our earlier results. The unexpectedly low noise ratios measured in our prior experiment with unstable dynamics, therefore, did not result from a direct dependency of noise ratio on vehicle dynamics. Moreover, the low noise ratios did not result from a dependence on input characteristics, although there is some indication that the velocity disturbance led to higher noise ratios.

The simplest explanation for the improved performance of the subjects in this experiment is simply that they had learned to become more effective controllers as a result of their exposure to the unstable dynamics that were employed in this and in the prior experiment. With the unstable dynamics, the error variance was considerably more sensitive to observation noise than with the stable dynamics. As a result, the payoff from improving the noise ratio was appreciable and we postulate that the subjects learned to effect such an improvement. Once having managed a reduction in noise ratio, the subjects were apparently able to

transfer the low noise ratio to the other control tasks.

In Figure C-5 is plotted the error variance as a function of observation noise ratio for the three experimental conditions investigated in this experiment. The optimal-control model was used to predict the relationship between error variance and observation noise ratio. Observation noise ratio was varied between -26 dB and -20 dB; the remaining model parameters were selected from Table C-2.

Predicted error variance was much more sensitive to observation noise ratio when the vehicle dynamics were unstable (and the input applied as an acceleration disturbance) than when the vehicle dynamics were K/s^2 and the input was applied as a velocity disturbance. An intermediate sensitivity was observed for the stable-vehicle, acceleration-disturbance task. For the velocity-disturbance task, a reduction of noise ratio from -20 dB to -26 dB reduced the predicted error score by only 2 dB, whereas the same variation of noise ratio represented about a 6.5 dB variation of performance for the unstable vehicle task. In other words, reducing the noise ratio was almost three times as effective when the vehicle was unstable. Hence, the subjects apparently found it worth the extra effort to reduce observation noise in this situation, but not in the control situations using stable vehicle dynamics that we investigated in earlier experiments.

SUMMARY AND DISCUSSION

An experiment was performed to determine why the observation noise ratios observed in the single-axis, multivariable experiment with unstable dynamics were substantially lower than the -20 dB noise ratio found in earlier experiments. Three control tasks were studied: (a) the second-order, unstable-vehicle task $k/s(s-\lambda)$ investigated in the prior experiment, (b) a K/s^2 task with the

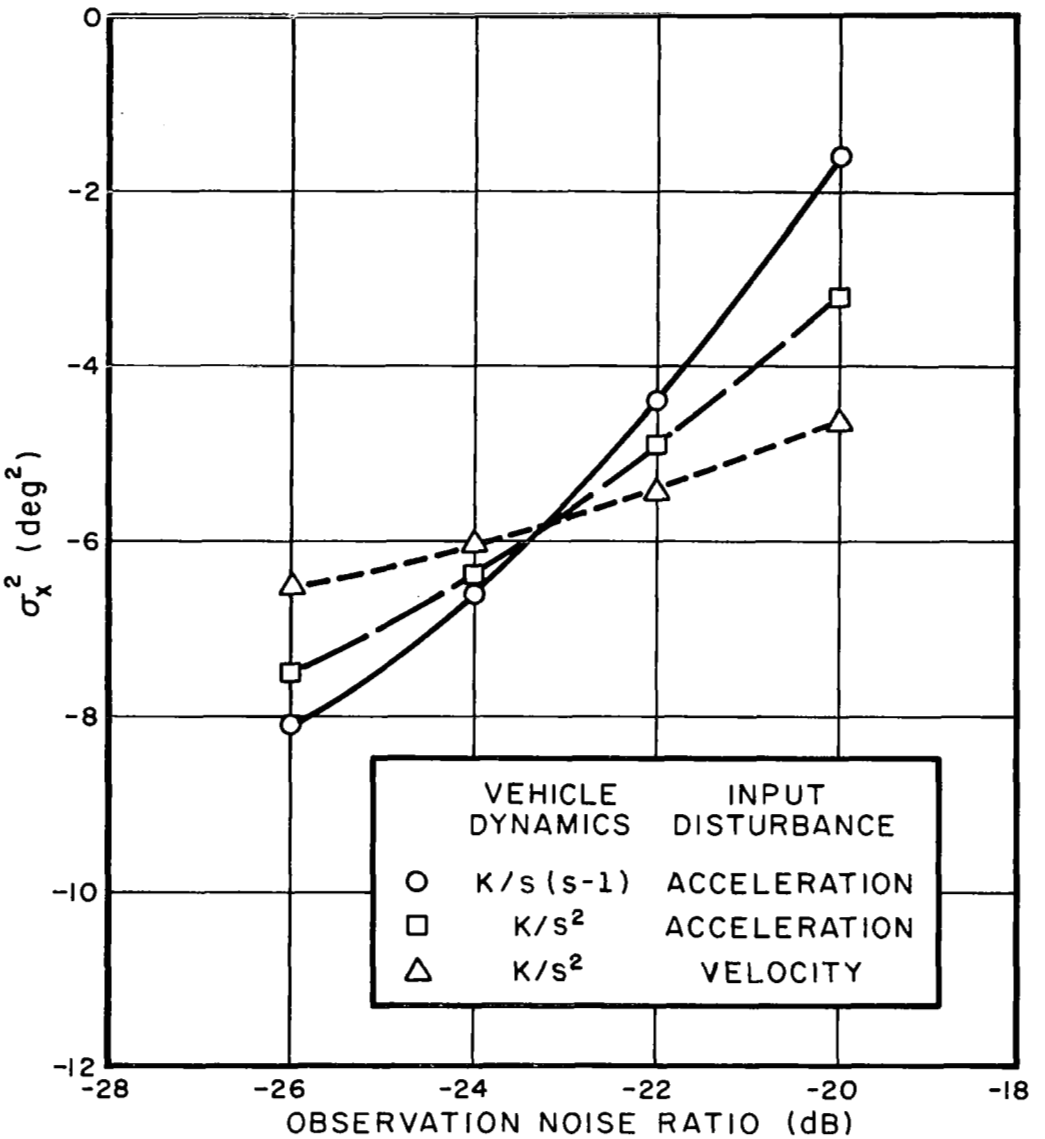


FIG.C-5 EFFECT OF OBSERVATION NOISE RATIO ON PREDICTED ERROR VARIANCE SCORE

input applied as a disturbance to vehicle velocity, a control situation also studied in an earlier experiment, and (c) a K/s^2 task with an acceleration disturbance. By matching the output of the optimal-control model to the experimental data, we concluded that the noise ratios were not substantially different for the three tasks. We therefore ruled out the possibility of a direct functional relationship between noise ratio and control system parameters.

The most likely explanation for the relatively low noise ratio is that training on the unstable dynamics provided very strong motivation for the subjects to reduce their observation noise. An investigation of model behavior showed that predicted error score was three times as sensitive to observation noise ratio for the unstable-vehicle task than for the K/s^2 task with the velocity input. Once trained to achieve a low noise level, the subjects apparently retained this ability when presented with the stable-vehicle tasks.

Since the "minimum" observation noise ratio that the pilot is either willing or able to achieve depends on task parameters and training procedures, we must now face the question of how to select the appropriate noise ratio when we wish to obtain model predictions for a control situation we have not previously investigated. If we wish to study task interference, what ratio should correspond to "full capacity"? We have no simple answer to these questions. Clearly, the equivalent observation noise ratio is related to how well, and in what manner, the subject is trained. We would have trouble, then, predicting the absolute limit of performance attainable by a pilot in a given control situation.

Rather than concern ourselves with the minimum attainable noise ratio, we suspect that a more relevant approach to manual control problems in general is to determine the maximum permissible noise ratio that the pilot can allow himself without exceeding performance specifications. By comparing the observation noise requirements of various tasks, we can determine the relative "demands" they make on the pilot. We have applied this approach to the problem of predicting pilot workload in Section 5 of this report.

APPENDIX D

TABLES OF PERFORMANCE MEASURES

TABLE D-1

Effect of Display Format on Mean-Squared Error Scores

Subject	Mode of Viewing	Display Format	Axis Tracked Left	Right	Average
JF	Foveal	bar	.0778	.0793	.0786
		dot	.0750	.0834	.0792
	Peripheral	bar	.300	.790	.545
		dot	.390	.970	.680
DM	Foveal	bar	.0994	.0997	.0996
		dot	.125	.107	.116
	Peripheral	bar	.386	.378	.382
		dot	.477	.872	.675
GP	Foveal	bar	.140	.130	.135
		dot	.145	.132	.139
	Peripheral	bar	1.02	.968	.994
		dot	.957	1.09	1.02
HS	Foveal	bar	.0869	.0924	.0897
		dot	.103	.107	.105
	Peripheral	bar	.499	.696	.597
		dot	.444	1.17	.807

MS error scores in degree^2 visual arc.

1 run per subject per condition.

Tables D-3 through D-7 contain a breakdown on the error and control scores for the single-axis tracking experiments reviewed in Appendix A. Table D-2 below contains a list of the variables represented in these tables.

TABLE D-2

List of Performance Measures Appearing in Tables D-3 - D-7
(units refer to equivalent display deflection)

Symbol	
$\overline{x^2}$	mean-squared error in degree^2
σ_x^2	error variance in degree^2
σ_{xi}^2	input-correlated portion of σ_x^2 in degree^2
σ_{xr}^2	remnant-related portion of σ_x^2 in degree^2
σ_{xr}^2/σ_x^2	fractional remnant power (dimensionless ratio)
$\sigma_x^2/\overline{x^2}$	fractional variational power (dimensionless ratio)
σ_u^2	variance of control effort in (degree/second)^2
σ_{ui}^2	input-correlated portion of σ_u^2 in (degree/second)^2
σ_{ur}^2	remnant-related portion of σ_u^2 in (degree/second)^2
$\sigma_{ur}^2/\sigma_{ui}^2$	fractional remnant power (dimensionless ratio)

Note: Since mean-squared control effort and the variance of the control effort agreed to within 1% in almost all cases, we have not tabulated mean-squared control effort.

TABLE D-3

Effect of Relative Display Location on Average Performance Measures

Relative Display Location	Axis Tracked	Display Fixated	$\overline{x^2}$	σ_x^2	σ_{xi}^2	σ_{xr}^2	$\frac{\sigma_x^2}{\overline{x^2}}$	$\frac{\sigma_{xr}^2}{\sigma_x^2}$	σ_u^2	σ_{ui}^2	σ_{ur}^2	$\frac{\sigma_{ur}^2}{\sigma_u^2}$
Foveal	UL	UL	.117	.115	.0863	.0282	.981	.245	4.18	3.18	1.00	.230
	LL	LL	.140	.137	.0928	.0430	.974	.267	4.69	3.19	1.50	.270
	LR	LR	.134	.134	.104	.0292	.994	.213	3.89	3.11	0.783	.199
	UR	UR	.133	.132	.0985	.0347	.995	.240	4.07	2.99	1.08	.225
		AVG	.131	.129	.0954	.0339	.986	.241	4.21	3.12	1.09	.211
16° Peripheral Reference Extrapolation	UL	LL	.545	.516	.213	.303	.950	.569	5.63	2.49	3.13	.493
	LL	LR	.321	.317	.168	.149	.986	.469	5.41	2.78	2.63	.429
	LR	UR	.445	.432	.223	.208	.979	.447	5.44	2.53	2.91	.469
	UR	UL	.394	.385	.183	.201	.962	.497	5.59	2.65	2.94	.471
		AVG	.426	.413	.197	.215	.969	.495	5.52	2.61	2.90	.465
16° Peripheral No Reference Extrapolation	UL	UR	.727	.613	.192	.421	.865	.675	5.93	2.70	3.23	.497
	LL	UL	.626	.553	.182	.371	.894	.666	6.40	2.62	3.80	.552
	LR	LL	1.26	1.08	.283	.794	.886	.721	6.22	2.64	3.71	.565
	UR	LR	1.53	.983	.208	.775	.778	.728	6.03	2.61	3.42	.528
		AVG	1.04	.807	.216	.590	.856	.697	6.14	2.64	3.54	.535
22° Peripheral No Reference Extrapolation	UL	LR	1.72	1.31	.221	1.09	.811	.813	6.30	2.29	4.01	.628
	LL	UR	1.40	.784	.205	.580	.647	.741	7.63	2.57	5.06	.625
	LR	UL	1.46	1.33	.363	.965	.903	.739	7.72	2.77	4.94	.630
	UR	LL	1.95	1.35	.277	1.08	.764	.789	7.29	2.62	4.70	.629
		AVG	1.63	1.19	.267	.929	.781	.770	7.23	2.56	4.68	.628

Average of 4 subjects, 1 run per subject

TABLE D-4

Effect of Relative Display Location on Performance Measures: Control of UL Display Only

Relative Display Location	Display Fixated	Subject	$\overline{X^2}$	σ_x^2	σ_{xi}^2	σ_{xr}^2	$\frac{\sigma_x^2}{\overline{X^2}}$	$\frac{\sigma_{xr}^2}{\sigma_x^2}$	σ_u^2	σ_{ui}^2	σ_{ur}^2	$\frac{\sigma_{ur}^2}{\sigma_u^2}$
Foveal	UL	JF	.0804	.0791	.0549	.0242	.984	.306	5.22	3.77	1.45	.277
		DM	.119	.119	.0983	.0207	1.000	.174	3.47	2.90	0.572	.165
		GP	.150	.147	.0978	.0491	.980	.335	4.59	3.12	1.17	.321
		HS	.118	.113	.0942	.0188	.958	.166	3.45	2.91	0.544	.157
		AVG	.117	.115	.0863	.0282	.981	.245	4.18	3.18	1.00	.230
16° Peripheral Reference Extrapolation	LL	JF	.254	.241	.101	.140	.949	.582	5.37	3.05	2.32	.432
		DM	.396	.383	.210	.173	.967	.451	4.13	2.61	1.51	.369
		GP	.847	.800	.271	.529	.945	.661	9.55	2.26	7.29	.763
		HS	.681	.638	.268	.370	.937	.580	3.45	2.05	1.40	.407
		AVG	.545	.516	.213	.303	.950	.569	5.63	2.49	3.13	.493
16° Peripheral No Reference Extrapolation	UR	JF	.505	.500	.109	.391	.990	.782	6.56	3.27	3.30	.502
		DM	.607	.458	.243	.215	.755	.470	3.65	2.36	1.29	.354
		GP	1.07	.766	.238	.527	.716	.689	9.27	2.75	6.52	.703
		HS	.727	.726	.176	.549	.999	.757	4.21	2.41	1.80	.427
		AVG	.727	.613	.192	.421	.865	.675	5.93	2.70	3.23	.497
22° Peripheral No Reference Extrapolation	LR	JF	1.03	1.01	.135	.874	.981	.866	6.89	2.62	4.27	.620
		DM	1.07	.777	.254	.523	.726	.673	4.97	2.47	2.50	.504
		GP	2.83	1.53	.236	1.29	.541	.846	7.50	2.20	5.30	.707
		HS	1.95	1.94	.258	1.69	.995	.867	5.84	1.86	3.98	.681
		AVG	1.72	1.31	.221	1.09	.811	.813	6.30	2.29	4.01	.628

TABLE D-5

Effect of Relative Display Location on Performance Measures: Control of LL Display Only

Relative Display Location	Display Fixated	Subject	$\overline{X^2}$	σ_x^2	σ_{xi}^2	σ_{xr}^2	$\frac{\sigma_x^2}{\overline{X^2}}$	$\frac{\sigma_{xr}^2}{\sigma_x^2}$	σ_u^2	σ_{ui}^2	σ_{ur}^2	$\frac{\sigma_{ur}^2}{\sigma_u^2}$
Foveal	LL	JF	.0770	.0768	.0589	.0179	.997	.233	4.73	3.73	1.00	.212
		DM	.110	.110	.0946	.0154	1.000	.140	3.51	3.04	.472	.135
		GP	.239	.238	.122	.116	.996	.488	6.92	2.98	.394	.569
		HS	.134	.121	.0958	.0251	.903	.208	3.59	3.00	.585	.163
		AVG	.140	.137	.0928	.0436	.974	.267	4.69	3.19	1.50	.270
16° Peripheral Reference Extrapolation	LR	JF	.200	.194	.0927	.101	.970	.522	5.67	3.41	2.26	.398
		DM	.327	.327	.210	.117	1.000	.357	3.77	2.71	1.06	.281
		GP	.432	.429	.194	.235	.993	.548	8.74	2.76	5.98	.684
		HS	.326	.319	.176	.143	.979	.448	3.45	2.24	1.21	.352
		AVG	.321	.317	.168	.149	.986	.469	5.41	2.78	2.63	.429
16° Peripheral No Reference Extrapolation	UL	JF	.378	.368	.123	.244	.974	.664	7.14	3.79	3.35	.469
		DM	.621	.534	.222	.312	.860	.585	3.93	2.21	1.73	.439
		GP	.759	.605	.197	.409	.797	.675	10.0	2.37	7.66	.763
		HS	.746	.704	.184	.519	.944	.738	4.54	2.10	2.44	.537
		AVG	.626	.553	.182	.371	.894	.666	6.40	2.62	3.80	.552
22° Peripheral No Reference Extrapolation	UR	JF	.780	.717	.127	.590	.919	.823	8.30	3.27	5.04	.606
		DM	2.38	.840	.180	.661	.353	.786	7.72	2.63	5.10	.660
		GP	1.10	.770	.227	.543	.700	.705	11.3	2.69	8.59	.762
		HS	1.32	.810	.284	.526	.614	.649	3.20	1.69	1.52	.473
		AVG	1.40	.784	.205	.580	.647	.741	7.63	2.57	5.06	.625

TABLE D-6

Effect of Relative Display Location on Performance Measures: Control of LR Display Only

Relative Display Location	Display Fixated	Subject	$\overline{X^2}$	σ_x^2	σ_{xi}^2	σ_{xr}^2	$\frac{\sigma_x^2}{\overline{X^2}}$	$\frac{\sigma_{xr}^2}{\sigma_x^2}$	σ_u^2	σ_{ui}^2	σ_{ur}^2	$\frac{\sigma_{ur}^2}{\sigma_u^2}$
Foveal	LR	JF	.0920	.0920	.0739	.0182	.994	.197	4.43	3.68	.752	.170
		DM	.131	.131	.113	.0179	1.000	.137	3.27	2.92	.350	.107
		GP	.152	.152	.119	.0331	1.000	.218	3.96	3.04	.919	.232
		HS	.162	.159	.111	.0477	.981	.300	3.90	2.79	1.11	.285
		AVG	.134	.134	.104	.0292	.994	.213	3.89	3.11	.783	.199
16° Peripheral Reference Extrapolation	UR											
		JF	.207	.206	.113	.0932	.995	.453	5.35	3.20	2.11	.395
		DM	.296	.295	.224	.0706	.997	.240	3.30	2.49	.810	.245
		GP	.618	.617	.287	.330	.998	.535	9.32	2.52	6.80	.730
		HS	.658	.608	.268	.340	.924	.559	3.80	1.88	1.92	.505
		AVG	.445	.432	.223	.208	.979	.447	5.44	2.53	2.91	.469
16° Peripheral No Reference Extrapolation	LL											
		JF	.596	.524	.142	.382	.879	.729	5.77	2.91	2.86	.495
		DM	.571	.563	.211	.352	.986	.625	4.87	2.90	2.47	.508
		GP	1.85	1.79	.507	1.28	.968	.717	8.96	2.02	6.94	.775
		HS	2.03	1.44	.272	1.16	.709	.811	5.27	2.73	2.55	.483
		AVG	1.26	1.08	.283	.794	.886	.721	6.22	2.64	3.71	.565
22° Peripheral No Reference Extrapolation	UL											
		JF	1.73	1.37	.106	1.27	.792	.923	7.70	3.01	4.69	.609
		DM	.857	.712	.197	.515	.831	.724	6.47	2.53	3.94	.609
		GP	1.44	1.43	.343	1.09	.993	.760	11.2	3.29	7.85	.706
		HS	1.80	1.79	.806	.986	.994	.550	5.51	2.24	3.27	.594
		AVG	1.46	1.33	.368	.965	.903	.739	7.72	2.77	4.94	.630

TABLE D-7

Effect of Relative Display Location on Performance Measures: Control of UR Display Only

Relative Display Location	Display Fixated	Subject	$\overline{X^2}$	σ_x^2	σ_{xi}^2	σ_{xr}^2	$\frac{\sigma_x^2}{\overline{X^2}}$	$\frac{\sigma_{xr}^2}{\sigma_x^2}$	σ_u^2	σ_{ui}^2	σ_{ur}^2	$\frac{\sigma_{ur}^2}{\sigma_u^2}$
Foveal	UR	JF	.0751	.0739	.0571	.0168	.984	.228	4.76	3.73	1.03	.217
		DM	.102	.102	.0880	.0139	1.000	.137	3.37	2.81	.557	.165
		GP	.191	.190	.113	.0767	.995	.403	5.13	2.84	2.29	.447
		HS	.163	.163	.132	.0314	1.000	.193	2.01	2.56	.451	.150
		AVG	.133	.132	.0985	.0347	.995	.240	4.07	2.99	1.08	.225
16° Peripheral Reference Extrapolation	UL	JF	.240	.213	.0850	.128	.887	.601	5.77	3.17	2.60	.450
		DM	.241	.232	.166	.0657	.963	.283	3.93	2.97	.962	.245
		GP	.661	.660	.270	.390	.998	.591	8.35	1.96	6.40	.767
		HS	.433	.433	.212	.222	1.000	.511	4.29	2.48	1.81	.423
		AVG	.397	.385	.183	.201	.962	.497	5.59	2.65	2.94	.471
16° Peripheral No Reference Extrapolation	LR	JF	.579	.579	.110	.469	1.000	.810	5.91	3.14	2.77	.469
		DM	.524	.461	.237	.224	.880	.487	4.23	2.34	1.88	.446
		GP	3.50	1.82	.173	1.65	.520	.905	9.86	2.85	7.01	.711
		HS	1.50	1.07	.312	.755	.713	.708	4.12	2.12	2.00	.486
		AVG	1.53	.983	.208	.775	.778	.728	6.03	2.61	3.42	.528
22° Peripheral No Reference Extrapolation	LL	JF	.641	.613	.126	.487	.956	.794	6.45	3.09	3.36	.521
		DM	1.29	.966	.228	.738	.749	.764	5.48	2.51	3.08	.561
		GP	3.15	1.20	.259	.938	.381	.784	8.63	2.17	6.45	.748
		HS	2.71	2.63	.494	2.14	.970	.812	8.61	2.72	5.89	.684
		AVG	1.95	1.35	.277	1.08	.767	.789	7.29	2.62	4.70	.629

TABLE D-8

Effect of Fixation Point on Total Mean-Squared Error Score: Single-Axis Tracking

Subject	Display Fixated			
	UL	LL	LR	UR
JF	2.43	1.57	1.89	1.57
DM	1.81	2.37	2.06	3.39
GP	3.01	6.08	6.27	2.97
HS	3.10	5.55	3.92	2.67
Average	2.59	3.89	3.53	2.65

MS Error in degrees^2 visual arc

One score per subject

TABLE D-9

Effect of Viewing Conditions on Average Normalized Observation Noise Spectrum

Measurement Frequency (rad/sec)	Spectral Density Level in dB			
	Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
0.05	-12.2	-5.1	1.1	2.7
0.12	-12.6	-5.4	0.0	0.4
0.25	-14.2	-6.1	-2.3	-2.3
0.35	-14.6	-6.7	-4.0	-3.1
0.50	-13.9	-7.5	-6.3	-6.6
0.71	-15.2	-8.5	-8.8	-6.8
1.0	-16.0	-10.7	-10.3	-9.8
1.5	-16.7	-12.6	-12.0	-12.2
2.0	-17.7	-13.8	-13.8	-13.9
2.9	-18.0	-15.1	-15.9	-16.2
4.0	-19.7	-16.4	-17.6	-19.0
5.7	-20.8	-18.7	-20.2	-20.2
8.0	-23.0	-21.2	-23.7	-23.4
11.2	-25.2	-24.1	-25.9	-26.3
16.0	-29.1	-27.9	-26.9	-29.6
22.0	-32.0	-29.6	-32.7	-32.9
32.0	-33.3	-31.2	-33.2	-35.5

Average of 4 subjects, 4 runs each

0 dB - 1 unit of normalized power per rad/sec

TABLE D-10

Effect of Viewing Conditions on the Average Human Controller Describing Function

Measurement Frequency (rad/sec)	Gain in dB				Phase Shift in Degrees			
	Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext	Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
0.18	18.7	11.7	6.0	6.2	-46	-12	-34	-39
0.52	16.4	11.3	10.0	7.8	-28	-17	-3	2
1.0	16.2	11.4	12.5	10.4	-27	-18	-6	-11
1.5	15.8	12.2	12.1	10.8	-29	-23	-16	-17
2.0	15.6	14.4	12.2	11.4	-32	-28	-29	-21
2.9	14.7	11.6	13.6	12.1	-38	-34	-34	-33
4.0	14.4	11.3	11.7	11.6	-45	-44	-44	-44
5.7	14.0	11.6	11.7	11.5	-56	-60	-63	-63
8.0	14.6	12.8	12.8	12.4	-71	-87	-86	-89
11.0	15.7	14.2	14.7	14.4	-94	-121	-128	-131
16.0	19.4	15.9	16.6	16.8	-141	-196	-226	-222
22.0	21.2	13.0	13.4	15.7	-249	-300	-314	-313
32.0	17.0	7.8	10.1	12.9	-359	-401	-384	-412

Average of 4 subjects, 4 runs each

0 dB = 1 degree/second per degree

TABLE D-11

Effect of Number of Axes Tracked on Mean-Squared Error Scores: One Foveal and One Peripheral Display

Subject	No. Hands	Foveal		Peripheral		Total Score	
		1-axis	2-axis	1-axis	2-axis	1-axis	2-axis
JF	1	.165	.368	.206	.430	.371	.798
	2	.200	.290	.164	.343	.365	.633
	AVG	.183	.329	.185	.386	.368	.716
DM	1	.274	.420	.254	.628	.528	1.048
	2	.313	.393	.233	.412	.546	.805
	AVG	.294	.406	.243	.520	.537	.927
GP	1	.454	.819	.244	.444	.698	1.262
	2	.469	.419	.383	.612	.852	1.032
	AVG	.462	.619	.313	.528	.775	1.147
HS	1	.291	.618	.181	.437	.471	1.060
	2	.402	.670	.171	.316	.574	.989
	AVG	.347	.644	.176	.377	.523	1.024
AVG	1	.296	.556	.221	.485	.517	1.042
	2	.346	.443	.238	.421	.584	.865
	AVG	.321	.500	.230	.453	.551	.953

MS error scores in degrees^2 visual arc

Average of 2 runs/subject

TABLE D-12

Effect of Number of Axes Tracked on Mean-Squared Error Scores:
One Foveal and Three Peripheral Displays

Subject	Foveal			16° Peripheral Ref Ext			16° Peripheral No Ref. Ext			22° Peripheral No Ref. Ext		
	1-axis	3-axis	4-axis	1-axis	3-axis	4-axis	1-axis	3-axis	4-axis	1-axis	3-axis	4-axis
JF	.0691	--	.288	.180	.933	1.10	.221	.525	.739	.837	1.79	1.67
DM	.119	--	.190	.228	.388	.552	.472	1.40	1.47	.956	1.43	1.85
GP	.125	--	.392	.374	1.54	1.63	.549	1.30	2.19	1.98	2.04	2.66
HS	.140	--	.343	.302	.745	.873	.782	1.84	1.86	1.12	2.57	2.51
AVG	.113	--	.303	.271	.902	1.04	.506	1.27	1.57	1.22	1.88	2.17

MS error scores in degrees^2 visual arc

Average of 2 runs/subject: 1-axis and 4-axis

Average of 4 runs/subject: 3-axis

TABLE D-13

Summary of Analysis of Variance of Mean-Squared Error and Mean-Squared Control Scores for 4-axis Scanning Experiments: Inter-Axis Differences

Effect	Homogeneous Input			Non-Homogeneous Input		
	UL	LL	UR	UL	LL	UR

a. Mean-Squared Error

Axis	---	---	---	---	.01	.001
Subject	---	---	---	.001	.001	.001
Axis X Subject	.01	.05	.001	---	---	---

b. Mean-Squared Control

Axis	---	.001	---	---	.001	---
Subject	.05	.001	---	---	---	.05
Axis X Subject	.05	---	.001	.05	.05	.05

Significance level shown for the difference between the score on the LR axis and the score on one of the remaining axes.

TABLE D-14

Summary of Analysis of Variance of Mean-Squared Error and Mean-Squared Control Scores for 4-axis Scanning Experiments

Effect	Inter-Input Differences Axis				
	UL	LL	LR	UR	Total

a. Mean-Squared Error

Axis	---	.001	.05	.01	.001
Subject	.01	.05	.001	.001	.001
Axis X Subject	---	---	---	---	---

b. Mean-Squared Control

Axis	---	.01	---	---	.05
Subject	.05	---	.001	---	---
Axis X Subject	.05	.05	---	.001	.01

Significance level shown for the change in score on a given axis caused by a change in the MS Input on the LL Axis.

TABLE D-15

Effect of Display Conditions on Average Variance Scores: Single-Axis, Multivariable Task

Subject	σ_x^2 deg^2	$\sigma_{\dot{x}}^2$ $(deg/sec)^2$	σ_y^2 $(deg)^2$	σ_u^2 $newton^2$	$\sigma_{\dot{u}}^2$ $(n/sec)^2$
a. Y-Display Off					
KG	.221	1.40	.260	37.0	1.63×10^3
WM	.193	2.02	.353	61.5	2.49×10^3
JM	.255	1.66	.305	42.0	1.66×10^3
WR	.260	2.39	.423	58.2	2.16×10^3
Average	.232	1.87	.335	49.7	1.98×10^3
Geom.Avg.	.231	1.83	.330	48.6	1.95×10^3
b. Y-Display On					
KG	.110	1.00	.180	36.0	1.95×10^3
WM	.159	1.66	.250	51.2	1.97×10^3
JM	.151	1.06	.192	34.2	1.59×10^3
WR	.163	1.48	.263	43.2	1.86×10^3
Average	.146	1.30	.231	41.1	1.84×10^3
Geom.Avg.	.144	1.27	.227	40.6	1.83×10^3
c. Ratio: (Y-Display On)/(Y-Display Off)					
KG	.498	.714	.692	.973	1.20×10^3
WM	.324	.822	.822	.833	$.791 \times 10^3$
JM	.592	.639	.630	.819	$.958 \times 10^3$
WR	.627	.619	.622	.792	$.861 \times 10^3$
Average	.635	.698	.691	.841	$.952 \times 10^3$

TABLE D-16

Summary of Analysis of Variance of System Error Variance Scores: Single-Axis, Multivariable Task

Effect	Significance Level
No. of Indicators	.05* .001**
Subject	--
Subject X Number	.05

*Tested against "interaction" variance

**Tested against "experimental error" variance

TABLE D-17

Effect of Task Parameters on Average Variance Scores:
Single-Axis Task, Second-Order Dynamics

Subject	σ_x^2 deg^2	$\sigma_{\dot{x}}^2$ $(deg/sec)^2$	σ_y^2 deg^2	σ_u^2 $newton^2$	$\sigma_{\dot{u}}^2$ $(n/sec)^2$
a. Unstable Vehicle Dynamics, Acceleration Disturbance					
KG	.162	1.45	.258	6.14	4.29×10^2
WM	.168	2.90	.493	16.1	11.2×10^2
JM	.164	1.88	.328	8.19	4.27×10^2
WR	.253	2.75	.480	10.9	6.19×10^2
Average	.187	2.24	.390	10.3	6.49×10^2
Geom.Avg.	.183	2.16	.376	9.69	5.97×10^2
b. Stable Vehicle Dynamics, Acceleration Disturbance					
KG	.196	1.67	.267	7.62	4.77×10^2
WM	.179	2.41	.385	13.5	8.25×10^2
JM	.169	1.97	.315	9.62	4.69×10^2
WR	.237	2.41	.385	11.2	6.62×10^2
Average	.195	2.11	.338	10.5	6.08×10^2
Geom.Avg.	.194	2.09	.335	10.2	5.91×10^2
c. Stable Vehicle Dynamics, Velocity Disturbance					
KG	.278	3.72	.737	10.0	9.37×10^2
WM	.312	7.15	1.36	38.2	33.9×10^2
JM	.285	5.12	1.00	17.3	9.94×10^2
WR	.330	4.82	.930	15.7	11.9×10^2
Average	.301	5.20	1.01	20.3	16.3×10^2
Geom.Avg.	.300	5.07	.982	17.9	13.9×10^2

REFERENCES

1. McRuer, D.T., Graham, D., Krendel, E.S., and Reisener, W., Jr.: Human Pilot Dynamics in Compensatory Systems Theory, Models and Experiments with Controlled-Element and Forcing Function Variations. AFFDL-TR-65-15, Air Force Flight Dynamics Lab., Wright-Patterson Air Force Base, Ohio, July 1965.

2. McRuer, D.T., and Jex, H.R.: A Review of Quasi-Linear Pilot Models. IEEE, Trans. on Human Factors in Electronics, HFE-8, Sept. 1967, pp. 231-249.

3. Levison, W.H., and Elkind, J.I.: Studies of Multivariable Manual Control Systems: Two-Axis Compensatory Systems with Compatible Integrated Display and Control. NASA CR-554, 1966.

4. Levison, W.H., and Elkind, J.I.: Studies of Multivariable Manual Control Systems: Two-Axis Compensatory Systems with Separated Displays and Controls. NASA CR-875, 1967.

5. Baron, S., and Kleinman, D.L.: The Human as an Optimal Controller and Information Processor. IEEE, Trans. on Man-Machine Systems, MMS-10, Mar. 1969, pp. 9-17.

6. Kleinman, D.L., Baron, S., and Levison, W.H.: An Optimal Control Model of Human Behavior. Paper presented at the Fifth Annual NASA-University Conference on Manual Control, M.I.T. (Cambridge, Mass.), Mar. 1969. Also to be published in Automatica, 1970.

7. Clement, W.: A Theory for the Human Operator's Remnant in Multiloop Display-Control Tasks. Paper presented at the Fifth Annual NASA-University Conference on Manual Control, M.I.T. (Cambridge, Mass.), Mar. 1969.

8. Levison, W.H., Kleinman, D.L., and Baron, S.: A Model for Human Controller Remnant. Report No. 1731 (NAS8-21136), Bolt Beranek and Newman Inc., Cambridge, Mass., Oct. 1968.

9. Levison, W.H., Baron, S., and Kleinman, D.L.: A Model for Human Controller Remnant. Paper presented at the Fifth Annual NASA-University Conference on Manual Control, M.I.T. (Cambridge, Mass.), Mar. 1969. Also in IEEE, Trans. on Man-Machine Systems, Dec. 1969.

10. Broadbent, D.E.: Perception and Communication, Pergammon Press, 1958.

11. Green, D.M., and Swets, J.A.: Signal Detection Theory and Psychophysics, John Wiley and Sons, 1966.

12. Senders, J.W.: The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems. IEEE, Trans. of the Human Factors in Electronics Group, HFE-5, Sept. 1964.

13. Senders, J.W., et al.: An Investigation of the Visual Sampling Behaviour of Human Observers. Report No. 1246 (NAS1-3860), Bolt Beranek and Newman Inc., Cambridge, Mass., May 1965.

14. Jex, H.R.: Two Applications of a Critical Instability Task to Secondary Work Load Research. IEEE, Trans. on Human Factors in Electronics, HFE-8, Dec. 1967, pp. 279-282.

15. McDonnell, J.D.: Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities. AFFDL-TR-68-76, Air Force Flight Dynamics Lab., Wright-Patterson Air Force Base, Ohio, Dec. 1968.

16. McRuer, D.T., and Ashkenas, I.L.: A Theory of Handling Qualities Derived from Pilot-Vehicle System Considerations. Aerospace Engineering, vol. 21, no. 2, Feb. 1962, pp. 60, 61, 83-102.

17. Kris, E.C.: A Technique for Electrically Recording Eye Position. WADC Technical Report 58-660, Dec. 1958.

18. Cooley, J.W. and Tukey, J.W.: An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computation 10, no. 90, Apr. 1965, pp. 297-301.

19. Tustin, A.: The Nature of the Operator's Response in Manual Control and its Implication for Controller Design. Journal of the I.E.E., vol. 94, Part IIA, 1947.

20. Taylor, L.W., Jr.: A Comparison of Human Response Modelling in the Time and Frequency Domains. Proceedings of the Third Annual NASA-University Conference on Manual Control, NASA SP-144, 1967, pp. 137-156.

NASA-Langley, 1971 —— 10