N73-10114

10. Supervisory Sampling and Control: Sources of Suboptimality in a Prediction Task

THOMAS B. SHERIDAN AND WILLIAM B. ROUSE

Massachusetts Institute of Technology

A process supervisor is defined as a person who decides when to sample the process input and what values of a control variable to specify in order to maximize (minimize) a given value function of input sampling period, control setting, and process state. This paper presents experimental data in such a process where the value function is a time-averaged sampling cost plus mean squared difference between input and control variable. The task was unpaced prediction of the output of a second order filter driven by white noise. Experimental results, when compared to the optimal strategy, reveal several consistently suboptimal behaviors. One is a tendency not to choose a long prediction interval even though the optimal strategy dictates that one should. Some results are also interpreted in terms of those input parameters according to which each subjects' behavior would have been nearest optimal. Differences of those parameters from actual input parameters served to quantify how subjects' prediction behavior differed from optimal.

INTRODUCTION

The literature of man-machine systems has long implied that a human operator might best function as a supervisory controller, aloof from instant-by-instant in-the-loop error nulling. The supervisor, instead, programs or adjusts the control variables (or parameters of a lower level open or closed loop process) in order to optimize with respect to some performance function. He does this on a time scale which he tends to make cognitively comfortable and usually different from that of the controlled process itself. Though his actions may be intermittent the controlled process may nevertheless go on continuously.

Several modes of supervisory control may be distinguished. A first (most sophisticated) mode is where the supervisor specifies a subgoal and lets a computer algorithm select the parameters of how and when the process implements its state trajectory to the subgoal. This mode occurred, for example, in controlling the Apollo spacecraft, (ref. 1) and is also characteristic of supervisory control of teleoperators (refs. 2 and 3). This mode elevates the human to the highest level of authority with the least efforton his part in performing the actual control function.

A second mode is where the supervisor manually searches the parameter space of an automatic process to optimize performance, without direct regard for the process state trajectory. Nolan (ref. 4) and Pew and Jagacinski (ref. 5) measured how the human operator adjusted parameters of one system to make its output match that of another system.

A third (least sophisticated) mode is where the supervisor preprograms a process state trajectory as far as he thinks will optimize the performance function (which may include the cost of his own sampling or attention). This is the mode explored in these experiments. This mode has many analogies from everyday human supervisory situations. It requires of the human operator more detailed control functions than either of the above levels.

These experiments build on the previous human operator sampling models of Senders (ref. 6), Smallwood (ref. 7), and Carbonell (ref.

8). The Baron and Kleinman model of the human operator as an optima1 controller also includes an optimal sampling assumption (ref. 9).

THEORY

In a previous paper (ref. 10) it is shown that given

(1) A process input x characterized by a random time series of probability density (%) and by conditional probability density $\{x'|x_0t\}$, where x' is a value of x at time t after a known value of x, x_0 at time $t=0$,

(2) A value function $V(x,y)$, indicating the reward (in a positive sense) or penalty (in a negative sense) acquired per unit time when input is x and control variable is y,

(3) The cost C of sampling the input, the optimal control y at each t after sampling x is

$$y_{\text{opt}}=y[V(x,y),x_0,t]=\max_y \int_{x'} \langle V|x'y\rangle\{x'|x_0t\}. \quad (1)$$

Brackets ( ) indicate an (ensemble) expected value and $\int_{x'}$ indicates integration over all z'.

The optimal sampling interval T was similarly shown to be

$$T_{\text{opt}}=T_{\text{opt}}[V(x,y),\ \{x'|x_0t\},\ \{x\}] \\ =\max_T\left\{\frac{1}{T}\int_{t=0}^{t}\left[\int_{x_0} y_{\text{opt}}[V(x,y),x_0t]\{x_0\}\right]-\frac{C}{T}\right\} \quad (2)$$

While in certain cases analytical expressions for $\{x'|x_0t\}$ are analytically tractable, this function can also be generated montecarlo fashion by computer. Use of the latter is a convenient procedure when computer solution of values of y_{opt} and T_{opt} is desirable.

The above theory applies for any input time series $x(t)$, with its corresponding $\{x\}$ and $\{x'|x_0t\}$, and any value functions of control and sampling, $V(x,y)$ and C. However, it is obvious that arbitrary inputs and value functions will result in tasks which may be unnatural (not be like any familiar real-world experiences) and virtually impossible to learn.

EXPERIMENTS

The task considered in these experiences consisted of asking subjects to make predictions (in a least-squared error sense) $y_j(\Delta t)$, $y_j(2$ At), . . . , $y_j(i\,\Delta t)$, . . . , $y_j(T_j\,\Delta t)$ of the future values of a time series $x_j(\Delta t)$, $x_j(2$ At), . . . , $x_j(i\,\Delta t)$, . . . , $x_j(T_j\,\Delta t)$. These values were outputs of **a** second-order digital filter driven by white noise. T_j* is the prediction interval for the j^{th} trial of an experimental run and was chosen by the subjects. **As** initial conditions, subjects were given $x_{j-1}[(T_{j-1}-1)$ At] and $x_{j-1}[(T_{j-1})$ At] which are also equal to $x_j(-$At) and $x_j(0)$. The end of the j -1^{th} interval serves as initial conditions for the j^{th} interval.

After inputing $y_j(\Delta t)$, $y_j(2$ At), . . . , $y_j(i$ At), . . . , $y_j(T_j\,\Delta t)$, subjects were shown the actual output $x_j(\Delta t)$, $x_j(2$ At), . . . , $x_j(i$ At), . . . , $x_j(T_j$ At) and their score for the j^{th} trial given by

$$V_j=\frac{\sum_{i=1}^{Ti}[x_j(i\text{ At})-y_j(i\text{ At})]^2+C_j}{T_j}, \quad (3)$$

where, C_j = the cost to the subject of updating his information about the actual x's. The subjects' monetary reward for performance on the j^{th} trial monotonically increased with decreasing V_j. Thus, a trade-off developed between the cost of making prediction errors and the cost of information (referred to as sampling cost).

For the $j+1^{\text{th}}$ trial, subjects specified T_{j+1} and used $x_j[(T_j-1)$At] and $x(T_j\,\Delta t)$ as initial conditions. They then input $T_{j+1}\,y$'s and the experiment continued as above.

The task was self paced, not forced pace; subjects had plenty of time to think through decisions before making them.

As previously indicated, the x's were generated by passing zero mean white noise through a second order filter. The transfer function of the filter $G(s)$ is given by

$$G(s)=\frac{w^2}{(s+w)^2} \quad (4)$$

where w =filter bandwidth in radians per At (unity damping ratio). The noise was realized

* The prediction interval can also be called a sampling period because the subjects' information concerning the actual x's is updated after T_j. Thus, prediction interval and sampling period are **used** synonymously in this paper.

digitally using a gaussian random number generator which was proven satisfactory by calculating its autocorrelation function.

The filtered output x was generated by using equations for the statistics of the distribution of the state of knowledge of x (after sampling at $t=0$) given by

$$m_{j+1}(t)=x_j(T, At)(1+wt)e^{-wt}+\dot{x}_j(T_j\,\Delta t)te^{-wt}, \quad (5a)$$

$$v_{j+1}(t)=v_0[1-(1+2wt+2w^2t^2)e^{-2wt}], \quad (5b)$$

where

$m_{j+1}(t)$ = mean of probability density function representing our state of knowledge of x during the $j+1^{th}$ trial
$v_{j+1}(t)$ = variance of x during the $j+1^{th}$ trial
v_0 = steady-state variance of signal
t = zero at beginning of $j+1^{th}$ trial.

Computationally, it was not necessary to use equation **(4)** as the filter output was directly obtained by using equations (5a) and (5b) which were discretely realized by letting $v_0=225$ and using

$$x(T_j\,At)=\frac{x(T_j\,At)-x[(T_j-1)\,At]}{At}. \quad (6)$$

To determine the optimal solution for the j^{th} trial, the optimal prediction interval T_0 (the subscript j is omitted because T_0 was generally constant over the entire j trials of a run) was found by minimizing the expected value of equation *(3)* V_j with respect to T_0. This expected value is given by

$$\langle V_j\rangle=\frac{\sum_{i=1}^{T_0} v_j(i\,At)+C_j}{T_0} \quad (7)$$

Once T_0 is determined, the optimal prediction strategy is to follow the nonstationary mean $m_j(t)$ until the next sample. For this procedure $\{x'|x_0t\}$ is characterized by $m_j(t)$ and $v_j(t)$. These were estimated on-line by simulating the filter output one hundred times from the given initial conditions and averaging across the results. This estimation technique was employed to eliminate problems that arose from using approximation of equation **(6).**

The experiments were performed entirely on a PDP–8 digital computer using a teletype as input and display apparatus. Each subject made his specification of inputs by typing numbers at his own comfortable pace.

The experimental input and display arrangement was identical for both experiments. A sample trial is illustrated in figure *1.* The subject input the bandwidth, sampling cost, and run number at the beginning of each run. These were specified by the experimenter. He was given an initial X and x. He chose T_j and input the appropriate number of y's (integers between 0 and ***100).*** As feedback he received the x's that occurred during the interval and his sampling cost per unit time C/T, squared error per unit time F/T, and V_j which is $(C+E)/T$. At the end of each run, he was told how much he made during that run. During all nine sessions of these experiments, sample signal plots for each of the bandwidths were posted within the subjects' view and subjects were encouraged to use them as necessary. After each session, the subjects' comments were gathered and recorded.

The experiments were performed using four male subjects including one graduate student and three undergraduates. All of the subjects had some familiarity with system dynamics. Each subject was told that the input he was predicting was the output of a second-order filter and that this output had a mean equal to fifty and a variance of v_0.

The optimal strategy was completely explained to the one graduate student hence referred to as

```
 G
NO RUNS 1
W  2  C: 900  R:630
X= 80    XD=  3

C/T=      0  En=      0  V=      0
T:6 2  Y  84  Y  65  Y  85  Y  84  Y  62
X=82  X= 83  X= 84  X= 85  X= 84  X= 81
CR=     158  E/T=      1  V=    151
T:8 81  Y  79  Y  76  Y  76  Y  74  Y  73  Y  73  Y  71

X=80  X= 79  X= 77  X= 72  X= 71  X= 70  X= 69  X= 67
C/T=    113  E/T=      9  V=    121
T:8
Y 6 5  Y  84  Y  62  Y  61  Y  60  Y  60  Y  56  Y  56
X=64  X= 61  X= 57  X= 55  X= 54  X= 54  X= 52  X= 51
UT-     113  En=     26  V=    138
T:8 50  Y  50  Y  51  Y  48  Y  48  Y  47  Y  46  Y  46

X=49  X= 47  X= 48  X= 50  X= 51  X= 51  X= 51  X= 51
C/T=    113  E/T=     12  V=    125
T:6
Y :51  Y  50  Y  49  Y  49  Y  47  Y  45
X=52  X= 54  X= 54  X= 52  X= 51  X= 51
C/T=    158  E/T=     16  V=    166
T:8
Y  51  Y  50  Y  48  Y  46
X=51  X= 53  X= 56  X= 58
C/T=    113  E/T=     52  V=    164
V AVG=    141
PAY THIS RUN $=  2.57
PAY TODAY $= 2.57    THANK YOU
```

Initializing parameters

Subject selects T
Subject inputs y
Computer generates x
Scores over this interval

Above process repeated for each new interval

FIGURE 1.—Input and display arrangement on teletypewriter.

the "trained subject" S_1 before the experiments began to assess the affect **of** such knowledge on performance.

The subjects were rewarded per session according to the following pay structures, where minimizing expected value of criterion function, V_j, (eq. *(3))* is equivalent to maximizing expected pay.

$$P_1 = \$1.00 + \$\frac{1}{N}\sum_{k=1}^{N}(C_k - V_k), \tag{8a}$$

$$P_2 = \$1.00 + \frac{1.5}{N}\sum_{k=1}^{N}\left(\frac{C_k - V_k}{P_k C_k}\right) \tag{8b}$$

$$P_3 = \frac{\$2.5}{N}\sum_{k=1}^{N}\left(\frac{C_k - V_k}{P_k C_k}\right) \tag{8c}$$

where

N = number of runs during session,
P_k = a normalizing percentage for k^{th} run,
V_k = cumulative score for the trials of the k^{th} run $= \sum_{i=1}^{j} V_i$
C_k = sampling cost for k^{th} run.

P_1 was used during the first three sessions for S_1 and S_2, P_2 during the remaining six sessions for S_1 and S_2, and P_3 with S_3 and S_4 for all nine sessions. These modifications in pay structure were made in an effort to improve the sensitivity of reward to performance and to insure a reasonable incentive to do well. However, discussions with S_1 and S_2 showed that they were not particularly conscious of the changes in pay structure.

Experiment I consisted of two parts. Each part included three **1-1/2** hour sessions. Before the experiment began each subject performed three practice trials. The individual runs lasted 40 to **60** time units. Runs were terminated at this fixed time regardless of whether or not a subject was in the middle of a prediction interval. This termination was necessary because subject and optimal predictor must run over the same length of time for a valid comparison.

During the first part of this experiment eight combinations of input bandwidth w and optimal sampling period T_0 were employed as indicated below

w \ T_0	3	5	8
0.2	1	2	3
.3	4	5	6
.4	7	8	...

Subjects were assigned a random sequence of the above eight runs configurations. Each run was forty units of time in length. Three replications of this matrix of assignment of runs constituted the first three sessions of the first part of experiment **I.**

The second part of experiment I used the following experimental matrix:

w \ T_0	3	5	8	12
0.2	1	2	3	4
.3	5	6	7	8
.4	9	10	11	12

A new random sequence of the above **12** runs was used to assign runs to each subject. Each run was 40 time units in length. Three replications of this sequence constituted the three sessions of the second part of experiment **I.**

Experiment II consisted solely of measuring the subjects' prediction ability as they were told what prediction period to use. The experimental matrix was

w \ T_0	5	8	12	17
0.2	1	2	3	4
.4	5	6	7	8

A new random assignment of the above eight runs was used for each subject. Each run was **60** time units in length. Three replications of this sequence constituted the three **1** hour sessions of this experiment.

Summarizing this section, two experiments were performed to study the human operator's ability to pick optimal prediction intervals and to predict between samples.

RESULTS

Experiment I included two variables of interest: prediction or sampling period chosen by the subject and the average criterion function V resulting from tracking. Figures **2, 3,** and **4** illustrate the sampling periods chosen by the subjects plotted versus the optimal sampling periods. These data from the sixth session represent the steady state choices of sampling periods and shows the subjects to have consistently underestimated these periods except on the low *T's.*

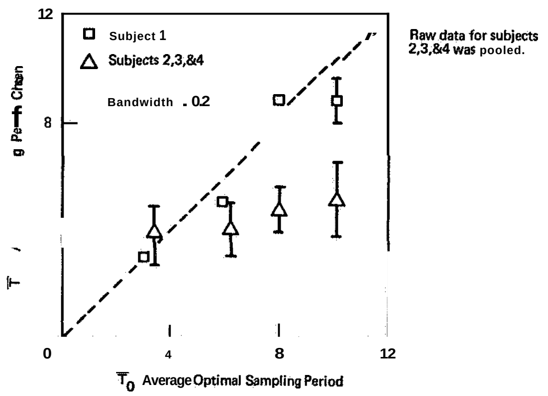

FIGURE 2.—Average sampling periods (at 0.2 bandwidth) chosen **for** the sixth session.

The trained subject did much better since he was aware of the range of *T's* before the experiment began.

Since the subjects chose *T*'s much smaller than they should have, their scores were much different than optimal, due mainly to higher sampling cost per unit time. The subjects' and optimal *V*'s are compared in figures 5, **6,** and 7. On the basis of the ratio V optimal/*V* subject, the subjects were found to be closest to optimal for *T* =*5* with a percentage of **0.864.** This experiment does not provide a fair comparison between *V's* because optimal and subject used different *T's.* However, the data do indicate that a different optimal sampling strategy may exist if a subject's tracking strategy is not optimal. Thus his picking a large *T* (given that he doesn't really predict optimally over that long of an interval) may actually give him a worse score than picking a smaller T. This phenomenon did not occur often enough to permit any more than conjecture on this point.

Experiment II only investigated the subject's ability to predict in the interval between samples, as they were told what T to use. The subjects' and optimal *V*'s are compared in figures 8 and 9.

The long *T* trials show the effect of a suboptimal prediction strategy more than the short

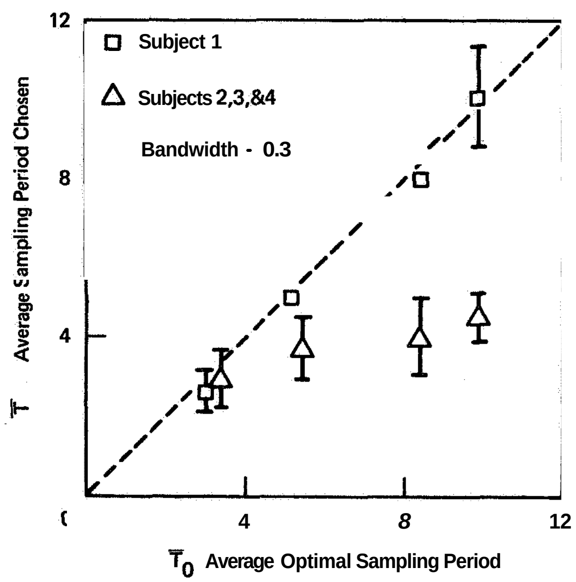

FIGURE 3.—Average sampling periods (at **0.3** bandwidth) chosen for the sixth session.

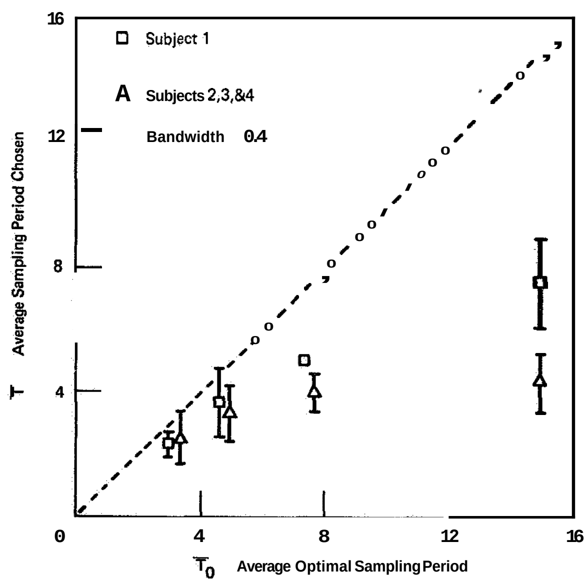

FIGURE 4.—Average sampling periods (at **0.4** bandwidth) chosen **for** the sixth session.

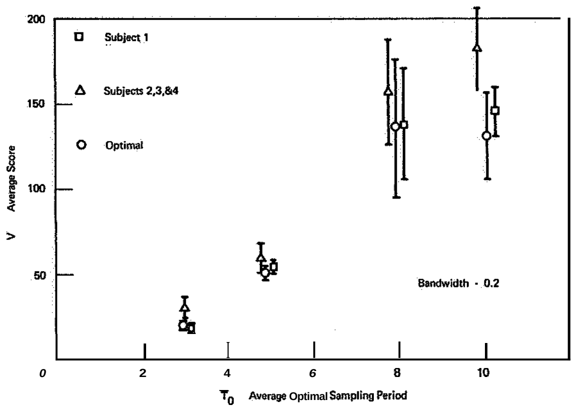

FIGURE 5.—Average scores (at 0.2 bandwidth) for experiment I.

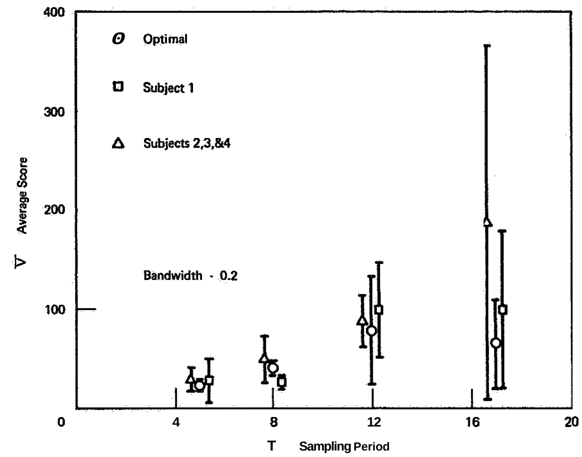

FIGURE 8.—Average scores (at 0.2 bandwidth) for experiment 11.

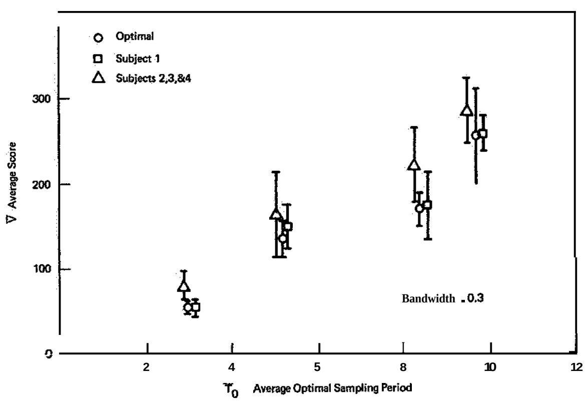

FIGURE 6.—Average scores (at 0.3 bandwidth) for experiment I.

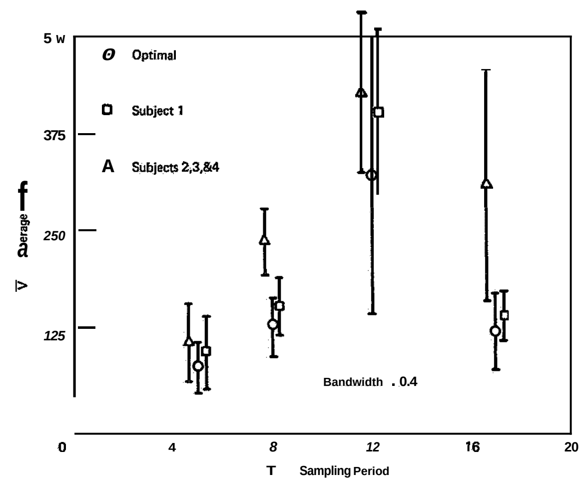

FIGURE 9.—Average scores (at 0.4 bandwidth) for experiment 11.

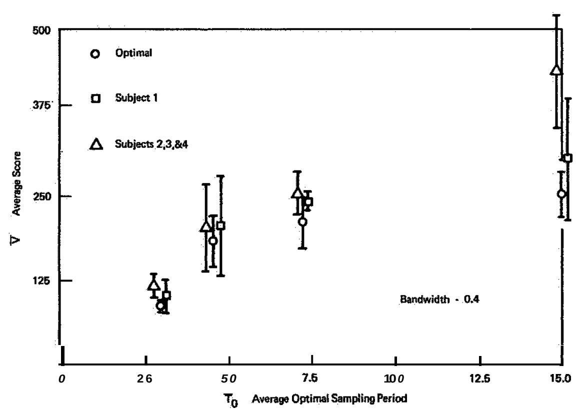

FIGURE 7.—Average scores (at 0.4 bandwidth) for experiment I.

T trials. This is reasonable because approximations to the optimal strategy of predicting the mean diverge more for longer *T*'s.

Plots of subjects' trajectories versus optimal trajectories, figures **10** through **13,** show that certain subjects consistently tend to extrapolate linearly and all subjects return to the mean much more slowly than the optimal. Thus, for *T*'s of **5,**

Notes:

(1) **Bars** on **data of figures 2 through 9 represent** ± 1 **standard deviation.**

(2) **Raw data for subjects 2, 3, and 4 were pooled (figs. 2 through 9).**

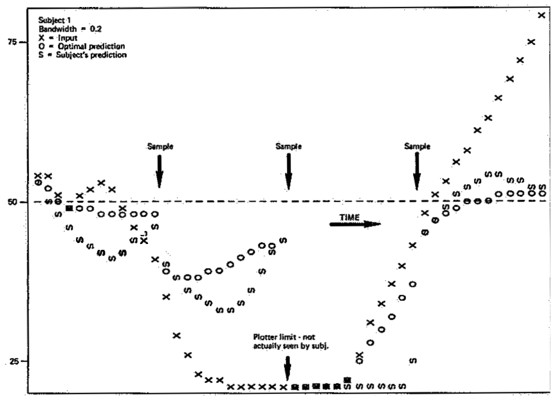

FIGURE 10.—Example of experimental time histories (subject 1).

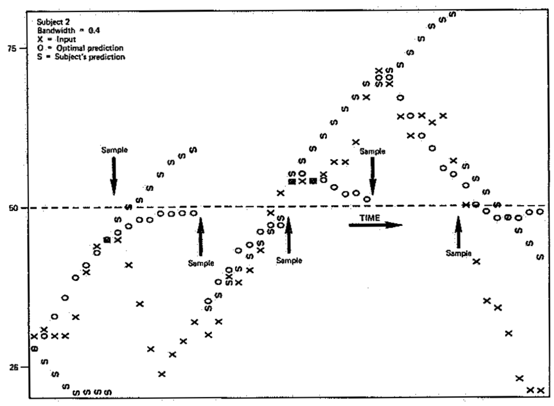

FIGURE 11.—Example of experimental time histories (subject 2).

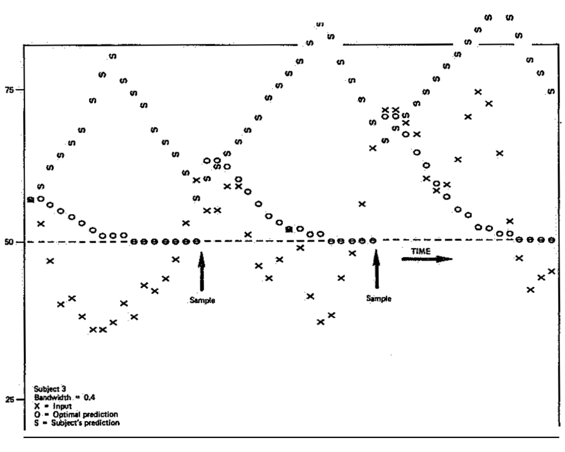

FIGURE 12.—Example of experimental time histories (subject 3).

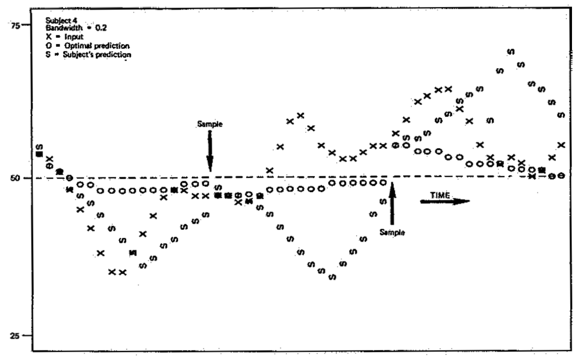

FIGURE 13.—Example of experimental time histories (subject 4).

subject and optimal are not very different while longer *T's* emphasize the difference between a linear and an exponential prediction. The one well-trained subject did fairly well, but of course on the average could not score better than the optimal.

Reasons for suboptimality include the subjects having an erroneous internal model of the process. The model could be erroneous by being of the wrong order or may be of the correct order but with the wrong parameters. Many other reasons are possible but will not be considered at this point.

If we assume that the subjects each used some second-order internal model and performed optimally with respect to that model we then find a least-squared error fit of the following model to the data:

$$y\{[T_j-(T_j-1)]\,At)=K_1\times(T_{j-1}\,At)+K_2\frac{x(T_{j-1}\,At)-[(T_{j-1}-1)\,At]}{At}. \quad (9)$$

From this we can see how the subjects weight position and derivative in their prediction strategy. The above second order model was fit only to the subjects' first choice after sampling. (Fitting the subjects whole trajectory would require assuming a model for the whole trajectory.) Figures **10** through **13** show that it is hard to justify the same model (different parameters as optimal for points further away from the sample than one or two). However, modeling only the first point yields an idea of the sources of subjects' suboptimality. Fitting equation **(9)** to the data from session nine of experiment II results in the following parameters:

W	Control	K_1	K_2	K_1/K_2
0.2	Optimal	0.98	0.91	1.09
	S_1	.99	.95	1.04
	S_2	.99	.68	1.33
	S_3	.98	.14	6.83
	S_4	1.05	.89	1.11
0.4	Optimal	0.94	0.67	1.40
	S_1	.98	.83	1.18
	S_2	.97	.39	2.51
	S_3	1.01	.14	7.07
	Sa	1.05	.62	1.70

Using these constants, we can quantify the performance of the subjects as displayed in figures **10** through **13.** In general, the subjects overestimated K_1 which kept them from returning to the mean as quickly as they should. Some subjects overweighted the derivative while others underweighted it. In either case, this caused a degradation of performance.

DISCUSSION AND CONCLUSIONS

None of the subjects were able to score as low as the optimal (although training helped). This was due to their inability or unwillingness to pick appropriate sampling or prediction periods and to predict optimally.

Inappropriate choice of T is attributable to two sources. First, inability on the part of the subjects to predict over longer periods of time without sampling may have caused them to choose shorter sampling periods over which they felt their skills more closely resembled optimal. Second, the subjects were unwilling to take the chance of a large error that might arise from a longer sampling period even though this also gave them a very small time-averaged sampling cost and thus a chance at a very high reward. This tendency is termed risk aversion and is evidenced by the subjects' comments that were collected.

Inability to predict optimally resulted from various sources. The subjects' internal model of the input process may have been erroneous as previously discussed. Also, subjects may not have fully realized what prediction strategy would minimize the specific value function. In particular, some of the subjects obviously did not know that predicting the estimated mean of the signal distribution was the optimal strategy and consequently they attempted to make their y's look like the x's. This strategy is disastrous if the subject guesses the wrong direction after sampling. Subjects' comments also indicated that they may need more than the last two points to predict the next point. In effect this amounts to their assuming a higher order filter than actually existed.

Summarizing, these experiments have enabled us to study sources of suboptimality in a specific task. These sources are perhaps applicable to many other tasks (i.e., systems, value functions, etc.).

REFERENCES

1. NEVINS, J. L.; JOHNSON, I. S.; AND SHERIDAN, T. B.: Man Machine Allocation in the Apollo Navigation, Guidance and Control System. Draper Lab Memo, M.I.T., **1968.**
2. SHERIDAN, T. B.: Use of Artificial Computation Loops within Human Control Loops for Remote Manipulation. Unpublished memo, Department of Mechanical Engineering, M.I.T., Oct. **28, 1964.**
3. FERRELL, **W.** R.; AND SHERIDAN, T. B.: Supervisory Control **of** Remote Manipulation. IEEE Spectrum, vol. **4,** no. **10,** Oct., **1967.**
4. NOLAN, G.: Human Response in Matching the Parameters of an Operating Dynamic System. SM Thesis, Department **of** Mechanical Engineering, M.I.T., **1959.**
5. PEW, R. W.; AND JAGACINSKI, R. T.: Mapping an Operator's Perception of a Parameter Space. Proceedings **of** the Seventh Annual Conference on Manual Control (Los Angeles), June **1971.**
6. SENDERS, J. W.: The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems. IEEE Trans. Human Factors in Electronics, vol HFE–5, no. **1,** Sept., **1964.**
7. SMALLWOOD, **R.** D.: Internal Models and the Human Instrument Monitor. IEEE Trans. Human Factors in Electronics, vol. HFE–8, pp. **181–187,** Sept., **1967.**
8. CARBONELL, J. R.: A Queuing Model of Many-Instrument Visual Sampling. IEEE Trans. Human Factors in Electronics, vol. HFE–7, pp. **157–164,** Dec., **1966.**
9. BARON, S.; AND KLEINMAN, D. L.: The Human **as** an Optimal Controller and Information Processor. IEEE Trans. Man Machine Systems, vol. MMS–IO, no. **1,** Mar., **1969.**
10. SHERIDAN, T. B.: On How Often the Supervisor Should Sample. IEEE Trans. System Science and Cybernetics, vol. SSC–6, pp. **140–145,** Apr., **1970.**