Report No.8108-134
October 1963

Technical Report

of

INTERRELATIONSHIPS BETWEEN TRACKING PERFORMANCE MEASURES, CONTROL DYNAMICS, and the EFFECTS of INCENTIVES

by

Roland H. Tanck

and

Horace F. Stokes Jr.

ACF ELECTRONICS
DIVISION
• RIVERDALE, MARYLAND •
ACF INDUSTRIES

ACF ELECTRONICS

Report No.8108-134
October 1963

Technical Report

of

INTERRELATIONSHIPS BETWEEN TRACKING PERFORMANCE MEASURES, CONTROL DYNAMICS, and the EFFECTS of INCENTIVES

by

Roland H. Tanck

and

Horace F. Stokes Jr.

ACF ELECTRONICS
DIVISION
• RIVERDALE, MARYLAND •
ACF INDUSTRIES

TABLE OF CONTENTS

FOREWORD

The research described in this report was conducted by the Human Factors Engineering Branch of ACF Electronics under Contract No. NASw-545 with the National Aeronautics and Space Administration.

We wish to acknowledge the contributions of Dr. John C. Townsend, and Dr. Richard A. Wunderlich of Catholic University, Mr. John Alexiou and Mr. Harold Garrett of ACFE.

A special word of gratitude must be given to Mr. Jack R. Huffner and Miss Margaret Mannarino of the Human Factors Engineering Branch of ACF Electronics. Each offered valuable suggestions, comments and criticisms that substantially aided the authors in the preparation of the research report.

The authors wish to express their sincere appreciation to Dr. Rayford Saucer and Mr. James Adams of NASA, Langley Flight Research Center for their assistance as contract monitors. Credit is also due Mr. Robert W. Taylor of NASA, Hdqs. whose insight generated much of the structure of the research.

INTRODUCTION

The machine component of a man-machine system is characteristically more predictable, consistent, and amenable to objective analysis than the behavior of the human operator. However, the magnitude of operator's contributions make it imperative that every effort be made to measure and analyze his output in a system and understand those factors which influence his behavior. The relationship between the physical aspects of the systems environment such as stimulus levels, ambient conditions, and anthropometric relationships have been the subject of extensive investigations. However, many of those conditions which make up the psychological environment have been generally ignored. One major psychological variable which is assumed to have an effect on man's performance is motivation; yet, few reports of research could be found concerning the relation of this variable to performance in man-machine systems. Apparently, control systems have been developed and measures of systems performance have been established without adequate knowledge of how either of these are affected by transient variations in the motivation of the operator. A demonstration of the presence or absence of motivational effects would give systems designers new insights and confidence in making overall systems evaluations.

Presently, the output of an operator in a system is usually defined in terms of one, or more, of seven measures. A need exists to examine these seven measures in terms of their sensitivity to changes in motivation and to determine whether the effects of motivation are general or limited to specific parameters. If the motivational effects are specific, which would be the best measures to use when motivation is likely to be an important factor ? To date, few attempts to analyze performance differences in terms of specific measures have been reported in the literature.

In addition, there is the problem of interrelationships among the seven measures. Few studies of tracking performance have used more than one measure in any given experimental situation, and in those cases where more than one measure of performance was used, inconsistencies in the correlations between the measures make any meaningful predictions impossible. In order to obtain meaningful data on the interrelationship between the seven measures, it would be necessary to collect data on them simultaneously under a well controlled multivariate design.

In an operational system, the operator can be expected to use several types of control dynamics. These control dynamics impose various combinations of rate, position and time requirements on his outputs. Experimental findings concerning the relationship between motivation and performance measures which may be true for one control dynamic would not necessarily apply to others. For this reason, it is desirable in any investigation of performance measures to take into account the effect of control dynamics.

Limited studies exploring specific aspects of tracking behavior have made significant contributions; however, we have reached a state of the art in system design where a more comprehensive research attack is necessary. Studies which permit the operator to perform in his natural work environment while data are collected on numerous machine and operator variables are a logical, and necessary, progression of laboratory research. Advances in research technology have made it possible to deal with the expanded numbers of subjects, measures and variables needed for a dynamic approach to systems evaluation.

In an academic context, research investigating the effects of motivation on learning, memory, perception and the performance of laboratory type motor tasks is extensive and virtually unanimous in claiming improved performance whenever incentives to succeed are introduced. Experiments dealing with motivation and its implications for man-machine systems are limited to a few studies.

Bahrick, Fitts, and Rankin (1) found that by manipulating motivation with monetary rewards they were able to improve time on target scores for a pursuit tracking task. However, the subject's performance on a peripheral task involving arithmetic computation decreased whenever monetary rewards were introduced.

Using a modified Two-Hand Coordination Test to provide time-on target scores, Noble (4) attempted to influence tracking scores with verbal statements comparing the subject's performance with a criterion stated as necessary to pass the test. No significant changes in performance, as measured in this study, resulted from the verbal attempts to manipulate the incentive conditions.

In an attempt to formulate a new theoretical approach to motivation, Eysenck (2) reports a series of studies involving performance on a pursuit motor task. As motivation he introduced rest pauses which he claims reduce reactive inhibitions in that they are followed by improved performance due to reminiscence. In comparing low drive, high drive and schizophrenic subjects, Eysenck suggests that the amount of reminiscence be considered as a measure of motivation.

As can be seen, the studies of motivation and tracking behavior are limited in number and scope. They have dealt with specific problems, usually with a theoretical interest, and have been restricted to a single measure of performance and a single control system dynamic.

<u>Objectives of this Research</u>

This study is designed to:

1. Determine the effects of incentive on each of the seven commonly used measures of tracking performance.

2. Determine the effects of three types of control dynamics on seven measures of tracking performance.

3. Determine the interaction of the effects of the introduction of incentive and use of three types of control dynamics on seven measures of tracking performance.

4. Determine the interrelationships of seven measures of tracking performance before the introduction of an incentive.

5. Determine the interrelationship of seven measures of tracking performance after the introduction of the incentive.

6. Determine the effects of the introduction of an incentive on the interrelationship of the measures.

7. Determine the effect of change in dynamics on the interrelationship of seven performance measures.

METHOD

At the outset, it will be necessary to define the experimental variable of this study. No simple, universal, and unequivocal concept of motivation exists. What may be motivating to one individual may be of no interest to another. Further, the direction and strength of the individual's response to a motivating stimulus is not predictable. As stated by Lazarus, Deese and Osler (3), "Unfortunately there is no way of assessing with any degree of confidence the degree of motivation of subjects..... so, the individual experimenter is left a free choice as to how he will operationally define motivation."

In view of these acknowledged difficulties, the problem of motivation was approached in this study by reducing the range of individual difference through the selection of subjects thought to be homogeneous in their needs, and imposing an arbitrary operational definition of motivation. This definition states simply that "motivation is the psychological state induced by the promise of monetary reward." Strictly speaking, this monetary reward should be considered an "incentive" rather than motivation or a motivating force. However, it is felt that in this study the terms can be used interchangeably.

The compensatory tracking task used required the subject to minimize the distance between a stationary command index and a controlled index driven by the control input from the operator and an independent forcing function.

Performance Measures

Seven measures of performance were selected on the basis that they are the ones most generally used and thus would provide the greatest continuity with other studies, and also they include those having the greatest usefulness for system engineers. The seven measures are identified and defined as follows:

AVERAGE ERROR (A. E.) - the average difference between the command and controlled indices for the specified period of time. $\sum_{i=1}^{N} \frac{e_i}{N} = \text{A. E.}$

AVERAGE ABSOLUTE ERROR (A. A. E.) - the average difference between command and controlled indices without regard to algebraic sign.

$$\sum_{i=1}^{N} \frac{|e_i|}{N} = \text{A. A. E.}$$

ROOT MEAN SQUARE ERROR (R. M. S.) - the square root of the average squared difference between command and controlled indices. $\sqrt{\sum_{i=1}^{N} \frac{e_i^2}{N}}$

This is the most frequently used measure in performance studies. In many studies it is the only measure taken.

TIME ON TARGET (T. O. T.) - the percent of total time that the controlled index is within an arbitrary tolerance band of the command index.

HITS - the number of times the controlled index "hit" the arbitrary tolerance bank which surrounds the command index.

AVERAGE ABSOLUTE CONTROL MOVEMENT (A. A. C. M.) - the average distance in arbitrary numbers along the arc of the control stick handle that the control stick is moved within each experimental trail. $\sum_{i=1}^{N} \frac{|c_i|}{N} = \text{A.A.C.M.}$

CONTROL STICK COUNT - the number of times the control stick is moved through its null position.

Apparatus

The tracking task was performed by the subjects while seated in an F-86L flight simulator. The canopy of the simulator was modified to eliminate much of the stray light from entering the cockpit. The standard centerpositioned control stick found in the F-86L cockpit was used. Its movement was restricted to the pitch axis only; all other stick motions were frozen. The top of the control stick could be moved seven inches from the null position. The control loading computer simulated a six pound spring and the force-displacement gradient was linear.

The display elements were presented on the five-inch E-4 Fire Control System CRT mounted below the flight instrument panel directly in front of the subject and at the normal viewing distance of about 28 inches. A circle 1/3-inch in diameter in the center of the CRT served as a command index, and the subject-controlled index was a dot 1 mm in diameter which could be moved vertically across the entire CRT.

The forcing function was composed of three superimposed sine waves with frequencies of 0.5, 1.0 and 1.5 radians per second with a slope of six db per octave.

When making a response to the system display, the operator's output may be mediated in many ways depending on the position, rate and time characteristics of the control system. The following three types of control stick dynamics were selected for this study:

Dynamic I (D-I) rate control $\frac{K}{S}$

Dynamic II (D-II) lag control $\frac{K}{S + 0.2}$

Dynamic III (D-III) rate control with lag $\frac{K}{S(S + 5)}$

The F-86L Simulator was used in combination with a Mid-Century 500 general purpose analog computer and an Ampex SP 300 four channel FM tape recorder. The computer controlled and monitored the simulator conditions and fed data into the recorder. These data were made available as a simultaneous, real time, permanent record and consisted of the forcing function, the control stick movement, and the error of each subjects's performance. These data were digitized to 12-bit accuracy at a sample rate of 25 samples per second, and the resultant data analyzed on an IBM 1401 computer. Each subject tracked eight minutes of which minutes two through six were recorded. The first three and one-half minutes of the record were converted to error scores. Digitizing the data gave us great flexibility in performing the statistical analyses employed, and reduced the possibility of human error to a minimum. The sampling rate selected was based on the sampling theory of Susskind (5) and the derivation of percent of TOT on the method of Webber and Adams (6).

Subjects

Forty-two male, undergraduate college students began the series of three tracking runs of this experiment; of those, 37 completed all requirements and contributed to the data used in the study. These students ranged in age from 18 to 21 and were screened for visual acuity and phoria with the Bausch and Lomb Orthorator. Subjects were accepted if their scores were within the limits recommended for use as normal standards by the equipment manufacturer. A groove type steadiness test requiring the subject to pass an electric stylus through a gradually narrowing channel was used as a screen for detecting gross defects in motor coordination. None of the subjects had previous tracking experience or pilot training. In recruiting the subjects, the emphasis was on "an opportunity to earn money" because it was felt that focusing on a monetary incentive would narrow the qualitative aspects of the motivation.

To encourage their participation throughout the entire study, the pay of the subjects was fixed at five dollars for each of the first three-hour periods and ten dollars for the final period. The original 42 subjects were assigned randomly to seven groups of six subjects each. Group A was used as the control group to determine group gain throughout the study. The other six groups (Groups B, C, D, E, F, G) - one for each of the six possible sequences of the three stick dynamics - served as experimental groups.

Experimental Design and Procedure

Since each subject was required to perform tracking tasks involving three types of stick dynamics within a 90 minute period, the possible effect of the sequence in which the three dynamics were presented had to be taken into account. To counterbalance the effects attributable to differential learning and transfer, each of the experimental groups was assigned to one of the six possible sequences of the three dynamics. Each of the six subjects in the control group was assigned a different sequence to achieve the same objective.

Three subjects of the same group were scheduled for each of the three-hour data gathering periods which were spaced exactly one week apart. Throughout the three-hour period the subjects entered the simulator in rotation, with each subject having 20 minutes of rest after each 10 minutes of activity. His active period included tracking, getting in and out of the cockpit, and receiving instructions. In the approximately eight minutes of actual tracking, a three and one-half minute segment of data was recorded, beginning two minutes after the forcing function was activated.

The design and intent of this research were not known to the subjects. They were informed at the outset by tape recording (text in Appendix) only that we wanted "to check our equipment". After the control group completed all three runs, and the first subject in an experimental group was in the cockpit, a tape recording (text in Appendix) announced the incentive for the first time. The incentive consisted of a dollar a second for any increase in time-on-target (TOT) for each dynamic, over the time-on-target score he achieved in the final series of runs during the previous week. In addition, there was a prize of $50 for the subject making the greatest gain and $25 for the next greatest gain. Thereafter, the recording was played for each subject in the experimental group before he began his final tracking runs.

The schedule for experimental and control groups for the three testing periods is shown in Table I.

TABLE I.
SCHEDULE FOR CONTROL AND EXPERIMENTAL GROUPS FOR VARIOUS INCENTIVE CONDITIONS

Group	N	Sequence of Dynamics			Period 1 (First Week)	Period 2 (Second Week)	Period 3 (Third Week)
A (Control)	6*	One S for each possible Combination			No Incentive	No Incentive	No Incentive
B	6	I	II	III	No Incentive	No Incentive	Incentive
C	6*	I	III	II	No Incentive	No Incentive	Incentive
D	6*	II	III	I	No Incentive	No Incentive	Incentive
E	6*	II	I	III	No Incentive	No Incentive	Incentive
F	6+	III	II	I	No Incentive	No Incentive	Incentive
G	6*	III	I	II	No Incentive	No Incentive	Incentive

* One subject in each of these groups failed to complete the experiment.
+ Data for one subject in this group were lost due to damaged recording tape.

The purpose of Period 1 was to provide an opportunity to learn to track for each of the stick dynamics and to reduce the variability of performance. During this period, the subject tracked for two separate eight minute runs for each of the three stick dynamics, making a total of 48 minutes.

During Period 2, which followed the previous period by exactly one week, the subject repeated the tasks of Period 1. The first eight minute period in each dynamic supplemented the learning and performance stabilization requirements of the study, whereas the second trials served as a standard of performance for comparison with later trials.

In Period 3, the entire procedure was repeated as in the previous two trials. The first runs of the three dynamics were intended as a warm-up and means of maintaining experimental constancy with the two previous periods. This was followed by the introduction of the incentive to the experimental groups. The control group finished the study under non-incentive conditions.

RESULTS

To meet the objectives of this study, the data have been subjected to three types of analysis:

1. An analysis of the change in the control and experimental group in the seven measures of tracking performance which can be attributed to the motivation variable.
2. A two way analysis of variance was performed to determine the significance of main effects in changes in motivational level and type of control dynamics as well as the interaction effects between incentive and type of dynamic.
3. A correlational analysis of the seven performance measures was completed for each of the two incentive conditions under all three dynamics in order to determine the interrelationship of the performance measures, and the sensitivity of these relationships to changes in motivation, and changes in dynamics.

The Effect of Motivation on Performance Measures

To determine the nature and extent of performance change, an analysis was made of the significance of the difference in performance measures shown by a comparison of the reference trials of period 2 with those taken during the final trails of period 3. This analysis was performed for both the control and experimental groups. The control group was given the final trials without any change in instructions; whereas, the experimental group was informed of the monetary rewards before the final trials began. This analysis of "before and after" measures for the same subjects makes it possible to trace the performance for seven measures and for three types of control dynamics across two incentive conditions. Tables 1 and 2 show the results of this analysis for the control group and for the experimental groups, respectively.

These results using correlated measures indicate clearly the effect of the incentive on the various measures. It is shown that the control group did not make a significant change in any of the seven measures in any of the dynamics. On the other hand, the experimental group reduced significantly all error scores except RMS and increased significantly in TOT and HITS. It is interesting to note how the AACM in the experimental group apparently varied with motivation and dynamics. As task difficulty (as defined by the dynamic used) increased, the significance of the motivation effect decreased. This is suggested by the shift in significant levels from .01 in Dynamic 1 to .05 in Dynamic II and then to no significance for Dynamic III. However, caution must be used in drawing conclusions, the difference between the proportions .01 and .05 was tested, and in this case not found to be significant. It seems that the quantitative aspects of the interaction effect on AACM should be investigated further.

TABLE 1
Summary of t Test of Correlated Means of Seven Measures of Tracking Performance of Control Group (N = 5) for Reference Trials and Final (no incentive) Trials

Dynamic	Measure	Reference Trials Mean	Reference Trials SD	Final Trials (No Incentive) Mean	Final Trials (No Incentive) SD	Mean Diff.	t	p
	AE	9.65	8.25	7.83	3.27	- 1.82	0.48	-
	AAE	25.58	8.85	20.99	5.10	- 4.59	1.36	-
	RMS	58.06	41.69	51.64	18.98	- 6.42	0.34	-
I	TOT	16.10	0.96	19.82	3.80	+ 3.72	1.56	-
	HITS	109.80	10.76	128.00	17.47	+ 18.20	1.33	-
	AACM	61.29	5.12	52.32	2.05	- 8.97	1.81	-
	CSC	63.20	11.36	58.00	14.52	- 5.20	0.89	-
	AE	14.03	5.04	4.94	3.22	- 9.09	1.84	-
	AAE	28.12	4.95	19.10	2.82	- 9.02	1.57	-
	RMS	63.04	17.83	55.39	18.49	- 7.65	0.42	-
II	TOT	15.64	2.25	24.20	4.65	+ 8.56	1.81	-
	HITS	109.40	14.68	218.65	58.76	+109.25	1.60	-
	AACM	63.32	4.01	52.83	2.55	- 10.49	1.74	-
	CSC	64.20	10.44	66.00	14.44	+ 1.80	0.49	-
	AE	11.25	4.58	9.20	4.19	- 2.05	0.95	-
	AAE	40.10	9.77	26.46	7.46	- 13.64	1.91	-
	RMS	75.16	19.74	53.47	19.33	- 21.69	1.31	-
III	TOT	11.94	2.45	16.86	2.44	+ 5.72	1.96	-
	HITS	97.80	12.32	121.00	15.85	+ 23.20	1.83	-
	AACM	70.33	8.47	56.19	6.95	- 14.14	1.85	-
	CSC	92.60	19.96	88.40	19.70	- 4.20	1.16	-

TABLE 2
Summary of t Test of Correlated Means of Seven Measures of Tracking Performance for Experimental Group (N=31) for Reference Trials and Final (Incentive) Trials

Dynamic	Measure	Reference Trials Mean	Reference Trials SD	Final Trials (Incentive) Mean	Final Trials (Incentive) SD	Mean Diff.	t	p
	AE	11.93	4.66	5.55	3.94	- 6.38	4.41	.01
	AAE	17.47	3.88	13.52	3.93	- 3.95	3.73	.01
	RMS	25.92	13.87	34.44	23.13	+ 8.52	1.92	----
I	TOT	20.44	5.76	29.67	7.24	+ 9.23	4.46	.01
	HITS	112.45	29.34	215.52	55.67	+102.97	4.93	.01
	AAC	52.27	4.95	48.04	3.35	- 4.23	3.39	.01
	CSC	56.97	23.97	54.35	3.47	- 2.62	0.35	----
	AE	13.05	6.45	5.58	5.55	- 7.47	3.47	.01
	AAE	19.43	7.16	13.32	3.72	- 6.11	3.59	.01
	RMS	34.02	28.19	32.42	23.81	- 1.60	0.25	----
II	TOT	19.55	5.32	30.60	8.45	+ 11.04	4.12	.01
	HITS	102.00	22.14	218.65	58.72	+116.65	4.91	.01
	AACM	51.95	7.54	48.87	3.32	- 3.08	2.10	.05
	CSC	556.29	19.13	57.16	11.65	+ 0.87	0.20	----
	AE	13.95	5.37	6.48	4.63	- 7.47	4.51	.01
	AAE	24.54	5.92	19.53	4.98	- 5.01	3.88	.01
	RMS	37.16	17.31	38.81	21.97	+ 1.65	0.38	----
III	TOT	16.11	3.60	21.32	4.33	+ 5.38	4.39	.01
	HITS	101.90	19.18	165.90	37.61	+ 64.00	4.85	.01
	AACM	54.22	10.17	52.13	5.62	- 2.09	0.91	----
	CSC	86.03	27.90	68.77	14.20	- 17.26	2.55	.05

ANALYSIS OF VARIANCE DATA

The data were subjected to an analysis of variance in order to determine the effects of the motivation and control dynamics variables and their interaction effects. Results of this analysis are shown in Table 3.

The F tests performed on each of the seven measures of tracking per formance under two incentive conditions and three types of control dynamics were all significant except for the RMS measure. RMS did not reflect a change attributable to motivation.

Dynamic effects were significant ($p < .005$) in TOT and HITS, and to a lesser degree ($p < .05$ and $p < .01$) in AACM and CSC measures, respectively. The variance contributing to this result is traceable to Dynamics III. This is shown by Table 2 which indicates that changes in Dynamics I and II are almost identical in direction and magnitude and about twice the amount of gain shown for Dynamics III.

The significant interaction effects found only in TOT and HITS are also a function of Dynamics III. As shown in Table 2, the increase in means for Dynamics III under incentive conditions for TOT and HITS was only about 60% of the increase in means of Dynamics I and II.

TABLE 3
Summary of Analysis of Variance Data for
Seven Measures of Tracking Performance Under Two
Incentive Conditions for Three Types of Control Dynamics *

		df	MS	F	p
AE	Motivation	1	2,348.95	84.00	.005
	Dynamics	2	34.17	1.24	---
	Mot x Dyn	2	6.11	0.22	---
	Within	180	27.54		
AAE	Motivation	1	1,173.48	43.44	.005
	Dynamics	2	7.80	0.28	---
	Mot x Dyn	2	17.98	0.66	---
	Within	180	27.77		
RMS	Motivation	1	375.41	0.759	---
	Dynamics	2	965.20	1.95	---
	Mot x Dyn	2	409.59	0.827	---
	Within	180	494.52		
TOT	Motivation	1	5,711,626.93	1527.000	.005
	Dynamics	2	2,771,547.95	741.000	.005
	Mot x Dyn	2	2,674,378.72	715.000	.005
	Within	180	3,740.61		
HITS	Motivation	1	415,871.10	249.00	.005
	Dynamics	2	16,701.07	10.00	.005
	Mot x Dyn	2	11,578.46	6.94	.01
	Within	180	1,668.19		
AACM	Motivation	1	457.27	11.20	.01
	Dynamics	2	173.85	4.22	.05
	Mot x Dyn	2	17.77	0.43	
	Within	180	41.11		
CSC	Motivation	2	1,692.05	4.39	.05
	Dynamics	2	9,531.55	24.73	.01
	Mot x Dyn	2	1,490.77	3.86	--
	Within	180	385.39		

* Due to the lack of independence between groups, Box's correction was used to determine the degrees of freedom for p.

INTERCORRELATIONAL ANALYSIS

It is the intent of this analysis is using the Pearson r formulas to determine the extent to which the seven performance measures share a common variance and how this commonality is affected by changes in dynamics and motivation. Intercorrelation matrices for these seven measures applied to the experimental group are shown for the pre-incentive conditions in Table 4 and for the incentive groups in Table 5. Based on the matrices the following general observations can be made:

1. It will be seen that of the three error measures (AE, AAE and RMS), that RMS and AAE show considerable common variance whereas AE is more closely related to HITS and TOT than to the other two error measures.
2. The relationship between the two measures most commonly used as an overall measure of performance, TMS and TOT, ranges from moderately significant to insignificant, r. -45 to r. -.08. There is a tendency for the relationship to be more significant under incentive conditions. AAE shows a much stronger relationship than RMS with TOT with r's ranging from -.50. to -.77.
3. The highest and most consistent relationship is found between HITS and TOT. (r=.72 to r=.93)
4. Three measures (HITS, AACM and CSC) are related to the pattern of response used by the operator, and these showed several moderate and consistent relationships during the pre-incentive trials. However, after the incentive was introduced, no consistent or significant correlations remained. For example, HITS and CSC showed r's of .39, .50 and .59 for Dynamics I, II, and III, respectively in the pre-incentive measures. After the incentive, these same correlations were -.04, .16 and .11. This shifts away from significance may be due to changes in the approach to the task of many subjects upon the introduction of the incentive.

TABLE 4

Table of Intercorrelations Under Pre-Incentive Conditions for Seven Measures of Tracking Performance and Types of Control Dynamics

DYNAMIC I

	AACM	Hits	TOT	RMS	AAE	AE
CSC	.17	.39	.14	0	-.16	-.16
AACM		-34	-29	.19	.44	-.02
Hits			.89	-.08	- 64	-.58
TOT				-.20	-.72	-.62
RMS					.63	-.21
AAE						.40
AE						

DYNAMIC II

	AACM	Hits	TOT	RMS	AAE	AE
CSC	-.27	.50	.26	.06	-.13	-.06
AACM		-40	-.36	-.54	.51	-.15
Hits			.83	- 36	-.54	-.42
TOT				-.44	-.58	-.34
RMS					.93	.14
AAE						.42
AE						

DYNAMIC III

	AACM	Hits	TOT	RMS	AAE	AE
CSC	.74	.59	-.02	.01	-.03	.07
AACM		.17	-.43	.29	.44	.08
Hits			.72	-.26	-.57	-.46
TOT				-.45	-.80	-.46
RMS					.81	-.05
AAE						.33
AE						

df = 30

$\underline{p} < .05$ $\underline{r} = .296$

$\underline{p} < .01$ $\underline{r} = .409$

$\underline{p} < .005$ $\underline{r} = .449$

TABLE 5

Table of Intercorrelations Under Incentive Conditions for Seven Measures of Tracking Performance and Three Types of Control Dynamics

DYNAMIC I

	AACM	Hits	TOT	RMS	AAE	AE
CSC	-.44	-.04	-.06	.18	.07	-.12
AACM		-.14	-.14	.06	.19	.09
Hits			.93	-.43	-.72	-.49
TOT				-.24	-.62	-.65
RMS					.84	-.45
AAE						-.09
AE						

DYNAMIC II

	AACM	Hits	TOT	RMS	AAE	AE
CSC	.05	.16	.20	-.09	- 22	-.02
AACM		.02	.05	.10	.07	-.13
Hits			.84	-.26	-.71	- 52
TOT				.08	-.54	-.70
RMS					.75	-.64
AAE						- 09
AE						

DYNAMIC III

	AACM	Hits	TOT	RMS	AAE	AE
CSC	.33	.11	.05	.19	.15	.03
AACM		.23	.13	. 0	-.12	-.04
Hits			.87	-.47	-.80	-.58
TOT				-.32	-.77	-.63
RMS					.79	-.13
AAE						.30
AE						

df = 30

$\underline{p} < .05$ $\underline{r} = .296$

$\underline{p} < .01$ $\underline{r} = .409$

$\underline{p} < .005$ $\underline{r} = .449$

THE EFFECT OF DYNAMICS AND INCENTIVE CONDITIONS ON THE INTERRELATIONSHIPS OF SEVEN MEASURES OF TRACKING PERFORMANCE

Of the 126 correlations presented in Tables 4 and 5, some, of course, would be found significant by chance alone. Certain measures, however, are significantly interrelated consistently over the three Dynamics under both pre-incentive and incentive conditions. Three such measures (See Table 6) are HITS, TOT and AAE. Each of these measures correlates significantly ($p < .005$) with the other two, even with changes in Dynamics and conditions of motivation.

Of further interest is the consistent significant correlation ($p < .01$) of AE with HITS and TOT over all Dynamic and motivation conditions.

Finally, RMS shows a significant interrelationship ($p < .01$) with the AAE measure regardless of Dynamic or incentive conditions used in this experiment.

TABLE 6

Frequency of Significant Correlations ($p < .01$) between Seven Performance Measures

	AACM	HITS	TOT	RMS	AAE	AE
CSC	2	2	0	0	0	0
AACM		0	1	1	3	0
HITS			6	2	6	6
TOT				2	6	5
RMS					6	2
AAE						1

COMPARISON OF THE RELIABILITIES OF THE SEVEN TRACKING MEASURES

To decide which of several tracking measures to use when each is known to be sensitive requires that the "goodness" of the measures be compared. One measure of "goodness" in such situations is the reliability of the measure. Thus, selecting which of several measures to use depends upon which is more reliable. Table 7 shows the ranking of each measure as to its reliability in comparison with the other measures under two incentive conditions and three levels of control dynamics. This analysis was performed on the data of the experimental group only.

The measures were ranked according to their reliabilities. The reliabilities were computed by finding the standard error of each mean. Then, using the concept of the Coefficient of Variation, the percent each standard error was of its corresponding mean was calculated. These percentages were then ranked. This computation was necessary to prohibit the different magnitudes of means from yielding differing values of standard errors.

As can be seen from the table, the reliability of any given performance measure is relatively unchanged over the different dynamics and incentive conditions. Considering the overall picture, RMS and AE averaged lowest in reliability and AACM highest.

TABLE 7

Summary of Rankings According to Relative Reliabilities of Seven Measures of Tracking Performance Under Two Incentive Conditions for Three Levels of Control Dynamics

	Pre-Incentive			Post Incentive			Overall Mean Ranks
Measure	I	II	III	I	II	III	
AE	6	6	6	7	7	7	6.5
AAE	2	4	5	5	5	5	5
RMS	7	7	7	6	6	6	6.5
TOT	4	3	3	3	4	2	3
HITS	3	2	2	4	1	4	2
AACM	1	1	1	2	2	1	1
CSC	5	5	4	1	3	3	4

DISCUSSION

Based on longitudinal data for both the experimental and control groups, the effect of the incentives on the performance measures is quite apparent. In fact, six of the seven measures used in this study reflected significant changes when momentary incentives were introduced. These data show that the typical subject reacts in the following manner when promised a monetary reward for improved performance:

1. He reduces the average distance between the command and controlled index (lower AE and AAE).

2. Occasionally, he makes momentary wide excursions which result in a slight increase in RMS scores.

3. His time on target increases more significantly than any of the error measures.

4. He becomes more active and makes more frequent corrections when near the boundaries of the command index. This is shown by an increase in HITS.

5. He shows a tendency to reduce magnitude of control stick movements.

6. He does not change the number of times the control stick is moved through its null point.

In this representation of the typical subject's response to the incentive, the fact that the RMS measure made a slight gain in the unexpected direction is somewhat surprising. It must be noted that this measure is among those most frequently used in systems design and research. In some cases, the resistance of the RMS measure to motivational effects may be a definite asset, in others, it could lead to inaccurate conclusions.

Although Dynamic III was the most difficult of the three dynamics, there was a tendency for changes under the incentive conditions to occur to the same degree for all three dynamics. An analysis of variance showed significant interaction effects between incentive level and Dynamics only in the case of TOT and HITS. Likewise, an inspection of the correlation matrices and the reliability rankings fails to show large and consistent changes across Dynamics. The AACM ranked the highest in terms of reliability for dynamics and incentive conditions.

If a stable overall measure is required which is sensitive to the effects of motivation, the AAE is recommended over the other error measures. The TOT is perhaps equally suitable as a performance measure and may have an advantage of greater simplicity in the experimental situation. Although HITS showed many of the characteristics of a good performance measure, it cannot be considered satisfactory because the subject could develop a technique for producing a high HITS score which would nevertheless be a meaningless index of tracking skill.

The correlation between the number of HITS and TOT (.72 to .93) exceeds that of any other pairs of measures. A possible explanation for this relationship would be a tendency for good trackers in terms of TOT to remain near the boundary of the command index when not within it and thus make frequent passes through boundaries of the command index more likely.

The following implications for system design and research are evident in this study:

1. Because of its influence on almost all performance measures, motivation must be taken into account when determining the performance capabilities of system operators.

2. In choosing a single performance measure for the evaluation of a operator in a system, it must be considered whether such a measure should reflect or resist the temporary effects of motivation.

3. The findings of this research confirm the hypothesis that the various measures reflect psychological influences differentially. This makes it imperative that these measures be analyzed further for the effects of other psychological factors such as fatigue, stress, and operator overload.

SUMMARY AND CONCLUSIONS

This study was designed to determine the relative sensitivity of seven commonly used measures of tracking performance to the effects of changes in motivation. Data on a series of compensatory tracking tasks involving three types of control dynamics were collected on 36 subjects. Five of these were in the control group and 31 in the experimental group. After an identical period of practice, a base measure of performance was established for all subjects. The control group then repeated their tasks as before while the experimental group was promised a dollar per second for any increase in time on target scores on their subsequent trials. Analyses were made in terms of the Dynamics and incentive conditions involved, with the following results:

1. The control group measures showed no significant changes in their final trials while the experimental group scores changed significantly for all measures except RMS.

2. TOT and AAE emerged as the measures showing the most significant and consistent changes when incentives were introduced.

3. The motivational effects were not associated with any particular dynamic.

4. In terms of reliability across levels of incentive, AACM received the highest rank and AE and RMS were ranked as least reliable. An analysis of the intercorrelations between the measures shows AE to have a consistently low correlation with other error measures.

References

1. Bahrick, H., Fitts, P., and Rankin, R., Effect of Incentives on Reactions to Visual Stimuli, Journal of Experimental Psychology, Vol, 44, No., 6, Dec., 1952.

2. Box, G. E. P. (1954) Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems. II. Effects of Inequality of Variance and Correlation Between Errors in the Two-Way Classification Ann. Math Statist. 25 484-498.

3. Eysenck, H. J., Measurement of Motivation, Scientific American 208: May 1963

4. Lazarus, R., Deese, J., Osler, S., The Effects of Psychological Stress upon Performance, Psychological Bulletin, Vol. 49, No. 4, July 1952.

5. Noble, C. E., An Attempt to Manipulate Incentive-Motivation in a Continuous Tracking Task, Research Bulletin, AFPTRC-TR-54-45, 1954.

6. Susskind, A. K. (Ed) Notes on Analog-Digital Conversion Techniques. Cambridge, The Technology Press. 1957

7. Webber, C. and Adams J. Issues in the Use of an Analog-Digital Data System for the Measurement of Tracking Behavior AFOSR TN 59-528 April 1960

APPENDICES

APPENDIX A

INITIAL RUNS

We are interested in checking our equipment to see how well it performs for a wide range of subjects.

You will be taking turns sitting in the simulator cockpit. While in the cockpit you will be concerned only with the radar screen directly in front of you and the control stick.

The dot on the radar screen will move up and down, but can be stopped and controlled back and forth by movements of the control stick. Your task is to keep the dot in the circle, by making the proper control stick movements. Use only one hand, whatever hand you prefer but use the same hand throughout the experiment.

When you get the signal, try to center the dot within the circle, trying until told to stop. You will be in the cockpit for about ten minutes at a time.

The switch you will use to lower the canopy will be pointed out to you. You will be asked to raise the canopy with the same switch.

Try your best to perform in a steady consistent manner.

BETWEEN RUNS

Today we are repeating the tracking problem. We will be using the same procedures as before and the trial periods will be for the same duration.

Try your best to perform in a steady, consistent manner.

APPENDIX B

EXPLANATIONS GIVEN TO SUBJECTS OF THE VARIOUS TYPES OF STICK DYNAMICS

Dynamic I — You now have what is called rate control and the rate of movement of the white dot depends on the distance that you move the stick. Try moving the stick slowly back and forth. You will notice that the dot moves faster if the stick is moved a great distance than if it is moved a smaller distance.

Dynamic II — This is what is called a lag control. The movement of the stick controls the position of the dot, but there is a short time delay or lag between movement of the stick and final position of the dot. Move the stick to see how this will affect your control movements.

Dynamic III — We now have what is called a "rate with lag" control. You will find that the rate of movement of the white dot depends on the distance the stick is moved. You will notice the dot moves faster if the stick is moved a great distance than if there is a smaller movement of the control stick. However, you will notice that there is a short lag or delay between stick movement and the movement of the dot. Move the stick to see how the delay would affect the movement you would be required to make.

APPENDIX C

ANNOUNCEMENT TO SUBJECTS OF INCENTIVES TO BE USED

For the past few weeks you have made repeated runs of a tracking task in order to help us check out our equipment and to determine how well and how consistently an operator can be expected to perform.

Now we want a measure of your maximum performance. We want you to try your absolute best to keep the dot within the circle throughout the next series of tests. As an added incentive, you will be given a cash bonus of a dollar per second for any increase in your average time on target over the average you made last week in all of the three control stick dynamics. In addition, you will be competing against the other subjects of this experiment for grand prizes. These will be $50 for the greatest and $25 for the next greatest total gain in time on target score.

You will continue to use only one hand and use the same tracking techniques used in previous trials.

As before, each type of stick dynamic will be announced and there will be the usual 20 minute rest period between changes in control stick dynamics.

Are there any questions?