N82-13668

D3

A Study on Task Difficulty and Acceleration Stress

D. W. Repperger D. B. Rogers

Air Force Aerospace Medical Research Laboratory
Wright Patterson Air Force Base, Ohio 45433

SUMMARY

The results of two experiments are discussed which relate to task difficulty and the effects of environmental stress on tracking performance. The first experiment involved 5 different sum of sine tracking tasks which humans tracked both in a static condition and under a 5 Gz acceleration stress condition. The tasks were designed in such a manner as to investigate workload measures and to compare our hypothetical design to subjective evaluations. The tasks were required to satisfy 5 criteria specified in mathematical terms.

The second experiment involved similar environmental stress conditions but in this case the tasks were constructed from deterministic functions with specially designed velocity and acceleration profiles. In both parts of this experiment, subjective evaluations were obtained and compared to the assumption that difficulty is related to magnitudes of velocity and acceleration profiles of the target tracking task. Phase Plane performance analysis was conducted across 7 subjects to study potential measures of workload or tracking difficulty.

INTRODUCTION

In the study of manual control theory, the systematic characterization of task difficulty has been a problem of considerable interest for many years. An extensive amount of work has been done in this area and a variety of studies indicating different measures related to workload are available in the Human Factors and Psychological literature. In the engineering literature, the classical paper by Cooper (reference 1) illustrates the motivation for such a characterization of task difficulty - a subjective rating scale for human tracking. The extent at which this subjective rating scale can be used to predict pilot response is best illustrated in reference 2 where a thorough study has been done to investigate and pinpoint the exact cause-effect relationships between pilot subjective ratings and handling qualities of aircraft. This study uses a decision tree type of analysis procedure to investigate the responses.

At the Air Force Aerospace Medical Research Laboratory, it is of interest in our research program to develop standard tasks or levels of tracking difficulty and to be able to estimate levels of difficulty associated with human tracking. Once a consistent set of tasks are developed which provide a basis or standard for tracking behavior, it is then possible to more closely evaluate the effects of stress on human tracking performance. The criteria for the design of the tracking tasks must be such that each task is required to be a sensitive indicator of performance change (between each task number) and, in addition, the requirement is made that the task is to be sufficiently sensitive as to show a performance decrement between the stress-non stress condition.

This study consisted of two separate experiments. Both parts of this investigation involved human tracking for target forcing functions with different acceleration and velocity profiles. It was desired to study a critical task concept (reference 3) based on a hierarchy of difficulty associated with the different target forcing functions. This approach differs from the classical critical task concept considered by Jex, et al. (reference 4) in which the controlled element would have dynamics that change. In our studies, the controlled element (figure (1)) remained the same; the tracking tasks varied in levels of difficulty based on our hypothesis of different velocity and acceleration profiles associated with each target forcing function. The motivation for this work was due to an interesting paper by Verplank (reference 5) in which he equated difficulty and stress in studying human response behavior within a vigilance paradigm.

SYMBOLS

$f(t)$ = The Target Forcing Function Signal
$e(t)$ = The Closed Loop Error Signal
$x(t)$ = The Output of The Plant (Controlled Element)
R = Radius in the $\dot{f}$ versus $\ddot{f}$ plane = $(\dot{f})^2 + (\ddot{f})^2$
μ = Median of the distribution of the error window histogram
e_{RMS} = Root Mean Square error score
$\bar{X}$ = mean e_{RMS} value
σ = standard deviation of e_{RMS} value
$\bar{X}_d$ = The deviation of a difference from the mean of the differences.
M_d = Mean of the n differences of paired observations.
$\underline{p}$ = Probability
$\bar{t}$ = t test statistic
t = time

METHOD

Subjects - Seven male United States Air Force volunteers participated in this experiment. They had prior training in both the G type of stress exposures and manual tracking tasks.

Design of The Target Tracking Task - Part I

The objective of this study was to develop the tracking tasks of different levels of difficulty and to study their ability to produce performance decrements between tasks (for a given experimental condition) and between experimental conditions (for the same task). For the first part of this study, it was decided to design five different tasks with the following constraints:
(1) Each tracking task will be zero mean, constant variance, sum of sines.
(2) Each forcing function when presented as a replication will have a random

initial phase angle for each frequency component. The phase angle must be a prime multiple of the fundamental frequency and not a linear constant multiple of any other frequency component.
(3) Each forcing function will have a random phase angle between each frequency.
(4) Due to human physiological exposure limits in the design of the acceleration experiment, the length of each task was set at 15 seconds.
(5) The amplitudes of all the sinusoids are scaled so such that they all have equal power and produce the same displacement on the CRT (display). The open loop and autopilot runs of this study which verify this fact are presented in the sequel.
(6) The component frequencies of the sinuosids are "relatively prime" multiples of a fundamental frequency.
(7) A shift in frequency content is required so that $ff_i > ff_j$ is true if $i>j$, $i,j=1,2,..5$ where the frequency content of ff_i is higher than that of ff_j. The procedure for obtaining this desired result is discussed in reference 6.

Using a measure of difficulty denoted as R (the distance in the target phase plane (figure (2))) where R satisfies:

$$R^2 = (\dot{f})^2 + (\ddot{f})^2 \quad (1)$$

Then table I illustrates the values of R obtained for the 5 different tasks chosen in Part I of this study.

Table I Results of The Open Loop and Autopilot Simulations

Forcing Function (or Task) Number	Open Loop Error RMS * 2351.	Autopilot Error RMS * 2351	R^2(mean) For The Forcing Function	R^2(s.d.) For The Forcing Function
#1	718.6	477.8	0.352	0.290
#2	718.6	477.8	0.662	0.611
#3	718.6	477.8	1.212	1.140
#4	718.6	477.8	2.151	1.897
#5	718.6	477.8	3.509	3.2309

In this table the results of the open loop and autopilot runs are also displayed. These results (columns 2 and 3) illustrate the consistency of the normality conditions imposed in this study on the task numbers.

Design of The Tracking Task - Part II

In this design, the object was to design a different type of target forcing funcion. Figure (3) illustrates the shape of the functions used in this part of the study. In this case the objective was to have forcing functions of varying difficulty. The assumption is that the radius R is a metric of dispersion about the origin defined by equation (1) and tasks with larger R values are more difficult to track. The design of the function in figure (3) is a result of the need to have target tracking tasks that varied the value of R as a function of time. To create the shape of the diagram in figure (3), the following exponential functions were chosen based on a set of time intervals:

Time Interval(seconds)	Function Chosen	
$(t_1, t_2) = (0,5)$	$f_1(t) = \int_0^t a\, e^{-(s-2.5)^2/2\sigma^2} ds$	(2a)
$(t_2, t_3) = (5,10)$	$\dot{f}_2(x) = \dot{f}_1(5) + \int_0^x (-ae^{-(s-7.5)^2/2\sigma^2}) ds$	(2b)
$(t_3, t_4) = (10,15)$	$f_3(y) = f_2(10) + \int_0^y (-ae^{-(z-12.5)^2/2\sigma^2}) dz$	(2c)
$(t_4, t_5) = (15,20)$	$\dot{f}_4(z) = \dot{f}_3(15) + \int_0^z (ae^{-(t-17.5)^2/2\sigma^2} dt)$	(2d)

Where:

$$x = t-5 \quad (3a)$$
$$y = t-10 \quad (3b)$$
$$z = t-15 \quad (3c)$$

With some manipulation, the relationships (2a-d) and (3a-c) can be shown to produce the trajectories displayed in figure (3). The value a can be adjusted to sweep out a range of values. Table II illustrates the values chosen for part II of this study:

Table II - Forcing Function Design For Part II

FF #	a	b
1	0.1	.04
2	0.2	.08
3	0.3	.12
4	0.4	.16
5	0.5	.20

Randomization of The Presentation of The Tasks

Reference 6 describes the procedure chosen to ensure that the subjects would not know the order of presentation of the five different tasks at any time during the experiment. This was true for Parts I and II for both the static and stress portions of the experiment.

Apparatus

A 19-foot arm centrifuge (figure 4) was used to establish a 5 Gz stress condition for the subjects. In Air Force applications this acceleration force is in the z direction (down the spine of the subject) and is termed Gz. The centrifuge rotated at an angular speed of 27.5 RPM with the cab vectored at 78 degrees about a line in the z axis of the subject. The subjects wore standard Air Force helmets, gloves, and an Anti-G Suit with a G-valve. The Anti-G Suit-G-valve delivers a specific air pressure to the bladders of the Anti-G Suit.

Training Orientation and Data Exposures

During this training orientation, the subjects were required to asymptote to the five tracking tasks (performance training) and also to acclimate to the G stress (physiological adaptation). In the final design of this experiment, each day's run consisted of five component parts or phases (figure 5 illustrates one day's run for data collection). Phase I comprised of the presentation of the five tracking tasks in the static condition (no stress). Phase II consisted of the presentation of the five tracking tasks at an acceleration stress level of 5 Gz with a 20-second preliminary warm-up run at 4 Gz. After the five exposures at 5 Gz, the centrifuge was brought to a stationary position and the subject again performed five tracking tasks presented in random order in the static condition (Phase III). Phase IV of the daily run consisted of five tracking tasks presented in random order again under the five Gz stress as in Phase II. In Phase V of this experiment, the five tasks were presented in the static mode. Again, as with all the previous tasks, all forcing functions were presented in a random sequence. Four data days were collected after the subject progressed satisfactorily in the indoctrination period. During the data collection phase of the experiment, the subject never experienced more than 300 seconds per day of 5 Gz exposure nor more than two daily exposures per week. After the 4 data days were collected, a questionnaire was administered on the fifth day with the subject sitting in the centrifuge but with no machine motion. The questionnaires recorded subjective ratings of the task difficulty hierarchy.

Questionnaire

One definition of workload (reference 7), indicates that it is a function of increased performance requirements plus additional attention requirements. To get a true subjective evaluation, it was necessary to ask the subjects how they rated the tracking tasks. On the last day of the experiment the subjects were presented 25 tasks in random order. After the first task, each subject was asked to compare the task he was presently tracking with the previous one. The subject was asked whether the present task was more difficult, less difficult, the same, or not possible to rate. Thus the subject, was not knowledgeable as to which forcing function number was presented and would only give relative ratings between tasks.

RESULTS

CDF Performance Results From The Data

As discussed previously, after the 5-day indoctrination period the seven subjects tracked until they trained to an asymptotic level of performance for the different tracking tasks. One criterion used to define asymptotic performance is that on three consecutive days, the RMS performance scores do not decrease more than 5% on daily exposures of 25-50 static presentations of the five random targets per day. After this level was reached, the subjects were assumed to be trained. In part I of this study, the first question to be asked concerns the adaptation of the subjects to learning the tracking task and acclimation to G levels.

To address the question of learning and adaptation to stress, a table based on error scores was constructed across all seven subjects and four replications of each stress condition. Table III illustrates these results:

Table III - Stress Data $\bar{x}/\sigma$ Ratios For 7 Subjects, 2 Replications/Day

	Day 1	Day 2	Day 3	Day 4
ff#1	1.9	5.2	2.9	4.6
ff#2	2.6	4.9	3.6	4.0
ff#3	5.3	4.5	13.1	10.6
ff#4	8.2	10.0	17.1	13.5
ff#5	10.1	5.9	24.3	20.4

If any trends did exist in the data runs, due either to further performance training (reduction of tracking error) or possibly to further acclimation to G stress (physiological adaptation), they would be shown by a gradual increase in the ratio $\bar{x}/\sigma$ across a row for a given forcing function number. Since there appear to be no apparent trends for this stress acclimation, it is assumed that the subjects had adjusted to a steady state physiological conditioning and tracking performance level.

The next question to be addressed here is whether the forcing function number was correlated with measures of performance degradation. From the CDF figures (similar to figure 6), using data from all seven subjects (five replications), it was desired to conduct tests to investigate if a significant performance decrement exists dependent on forcing function number for both the static or the stress conditions. The following statistical test would determine this effect:

$$H_0: \quad \mu_{i+1} > \mu_i \quad i=1,\ldots 4 \qquad (4a)$$

versus

$$H_1: \quad \mu_{i+1} \leq \mu_i \quad i=1,\ldots 4 \qquad (4b)$$

where μ corresponds to the 0.5 line on the CDF in figure 6. This figure is illustrated here to show how the median point u is obtained. This corresponds to a "median window" size for the tracking error signal. The test was conducted for both the static data and the stress data. The results using a t statistic are displayed in Table IV:

Table IV

Hypothesis test on values	$\bar{t}$ for static data	p	$\bar{t}$ for stress data	p
$ff_2 > ff_1$	5.174	<.01	11.48	<.01
$ff_3 > ff_2$	18.39	<.01	19.91	<.01
$ff_4 > ff_3$	7.15	<.01	9.64	<.01
$ff_5 > ff_4$	4.55	<.01	5.04	<.01

The $\bar{t}$ values used in Table IV were the t statistic (2-tailed test) for correlated data (references 8 and 9) which satisfies:

$$\bar{t} = Md \Big/ \sqrt{\frac{\sum_{i=1}^{n} x^2{}_d}{n(n-1)}} \qquad (5)$$

where Md = the mean of the n differences of paired observations and x_d = the deviation of a difference from the mean of the differences. This test is for paired samples; they are not independent but are correlated due to the five replications involved with all seven subjects in the static case, and four replications with the stress data. The results from Table IV indicate performance decrements correlated with the forcing function number. The performance decrement is significant at an .01 level for an increase of forcing function number in both the static and stressed condition. The test given here corresponds to changes in medians (i.e., the 0.5 point on the CDF curve). Using the CDF method, this analysis could have been performed for any window size or any other point on the CDF curve. This is emphasized here because in other types of applications it may be desirable to look at a specified level of the CDF function (e.g. CDF$\neq$ 0.5) or at a specified window size. Finally, the tests illustrated here hold over both the static and stress conditions.

Another question to be addressed is whether the effects of the physiological stress induce a performance change for each task number. Using the data from the seven subjects and four replications of the stress condition, Table V illustrates the effects of stress on tracking performance.

Table V Comparisons of Stress vs Static Conditions

Hypothesis test on μ values	$\bar{t}$ for this test	p
ff_1 stress > ff_1 static	3.34	<.05
ff_2 stress > ff_2 static	1.54	<.10
ff_3 stress > ff_3 static	2.81	<.05
ff_4 stress > ff_4 static	5.83	<.01
ff_5 stress > ff_5 static	3.14	<.05

The t statistic used in this test is the same as in equation 5. One can now see the impact on performance degradation as noted by the effect of stress on tracking in Table V.

In part II of this study, it was desired to study this sensitivity effect for the second class of tracking tasks. Table VI illustrates the actual error score results for Part II as well as the equivalent values found in Part I:

Table VI

	Part I				Part II			
	Static		Stress		Static		Stress	
ff#	$\bar{x}$ mean	$\bar{x}$ s.d.	$\bar{x}$ mean	$\bar{x}$ s.d	$\bar{x}$ mean	$\bar{x}$ s.d.	$\bar{x}$ mean	$\bar{x}$ s.d.
1	94.6	20.4	127.7	29.6	22.67	3.87	37.73	7.14
2	162.4	48.7	179.0	33.2	38.42	6.71	59.14	12.02
3	283.9	49.4	306.3	29.7	54.32	11.40	79.91	14.88
4	383.6	22.9	414.7	37.4	106.32	26.63	121.77	34.89
5	6[illegible]4.5	23.9	649.5	26.4	223.09	38.82	239.96	48.58

The results of the statistical tests indicate a performance decrement under various conditions. The subjective data from the questionnaire are presented next.

Results From The Questionnaire

In the questionnaire, the subjects were asked to compare the relative difficulty of the task they were presently tracking with the previous task. Responses of "more difficult", "less difficult", "the same", or "couldn't tell" were then correlated with the task numbers presented. These results are displayed in table VII for Parts I and II of this study:

Table VII - Correlation of Task Numbers with Subjective Responses*

	Part I		Part II	
Subject #	# Correct/Total	% Correct	# Correct/Total	% Correct
1	24/25	96%	21/25	84%
2	23/25	92%	25/25	100%
3	24/25	96%	19/25	76%
4	25/25	100%	25/25	100%
5	23/25	92%	23/25	92%
6	23/25	92%	24/25	96%
7	23/25	92%	25/25	100%

The subjects also commmented that as they were presented tasks with higher forcing function numbers, the tracking tasks required more attention. This corresponds to the description of workload cited earlier in which higher performance requirements coupled with more stringent attention requirements increase workload.

SUMMARY AND CONCLUSIONS

This study used sensitive tracking tasks to evaluate performance degradation under acceleration stress. The tasks designed here had to satisfy certain criteria. First they had to be zero mean, constant variance, sum of sines. Second, open loop scores for all five tasks had to be identical. Third, the autopilot runs also had to yield a consistent score for all five tasks. When the human was tracking these tasks, however, a performance decrement had to be observed dependent on forcing function number for static tracking. In addition, the performance decrement had to occur as a function of the experimental conditions stress versus non-stress for each forcing function number.

At the conclusion of the experiment the subjects were given a questionnaire to rate the different tasks. Subjective ratings of each task in order of difficulty were necessary in order to verify the workload definition used here, which requires both a performance decrement and an attention requirement for arranging tracking tasks in order of increasing difficulty.

* Due to different subject pools in Parts I and II of this experiment, subject #N in Part I may not be the same person as Subject #N in Part II (N=1,..,7).

REFERENCES

1. Cooper, G. E.,1957, "Understanding and Interpreting Pilot Opinion". Aeronautical Engineering Review, March, pp. 47-56.

2. Cooper, G.E. and Harper, R.P., Jr., "The Use Of Pilot Rating In The Evaluation of Aircraft Handling Qualities" AGARD Report 567, April, 1969.

3. Repperger, D.W., Ward, S.L., Hartzell, E.J., Glass, B.C., and Summers, W.C., 1979,"An Algorithm To Ascertain Critical Regions of Human Tracking Ability", IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-9,No. 4,April, pp. 33-196.

4. Jex, H.R., McDonnell, J.E., and Phatak, A.V., "A Critical Tracking Task For Man-Machine Related To The Operator Effective Delay Time I: Theory and Experiments With A First Order Divergent Controlled Element", NASA CR-616, October, 1966.

5. Verplank, W.L., 1977, "The Facilitating Effects of Uncertainty In Long-Term Manual Control", Proceedings of The Thirteenth Annual Conference on Manual Control, June 15-17, MIT, pp. 101-105.

6. Repperger, D.W., Rogers,D.B., "A Task Difficulty - G Stress Experiment", Submitted for Publication.

7. Levison, W.H., 1979, Comments on Workload, at the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, Ohio.

8. Snedecor, G.W. and Cockran, W.G., 1967, "Statistical Methods", The Iowa State University Press, 6th. Edition.

9. Guilford, J.P., 1956, "Fundamental Statistics in Psychology and Education", 3rd. Edition, McGraw-Hill.

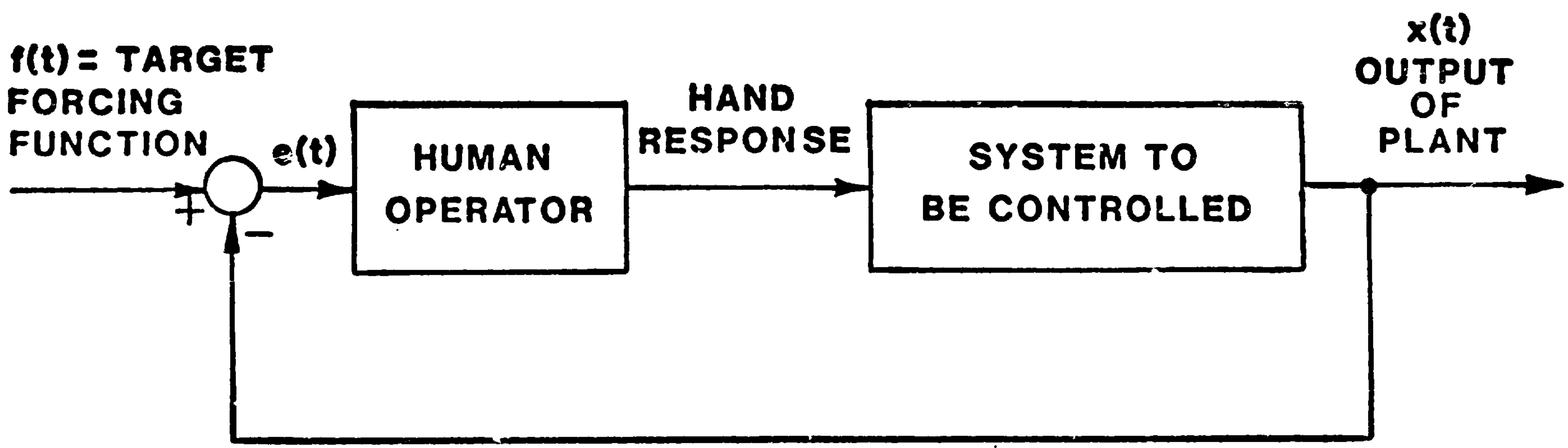

Figure (1) - THE MAN-MACHINE SYSTEM

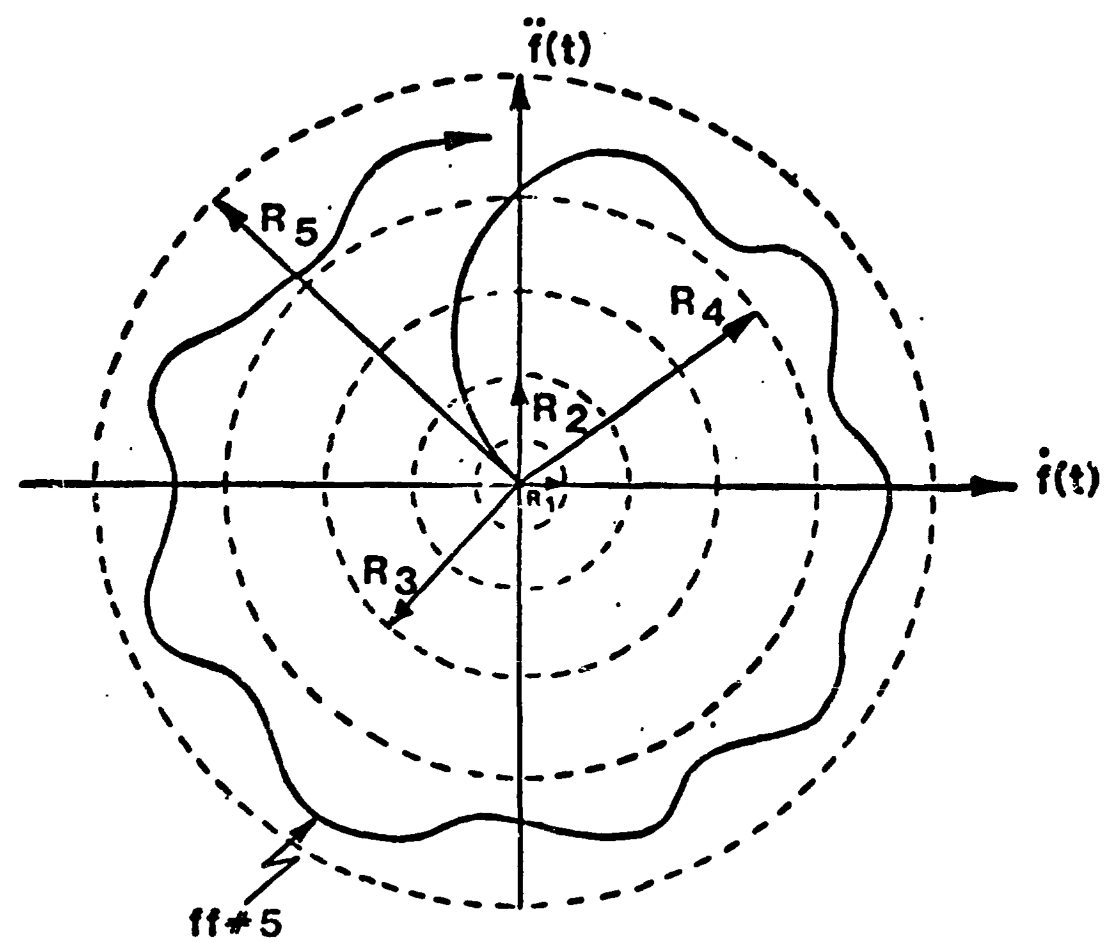

Figure (2) - A METRIC OF TRACKING DIFFICULTY

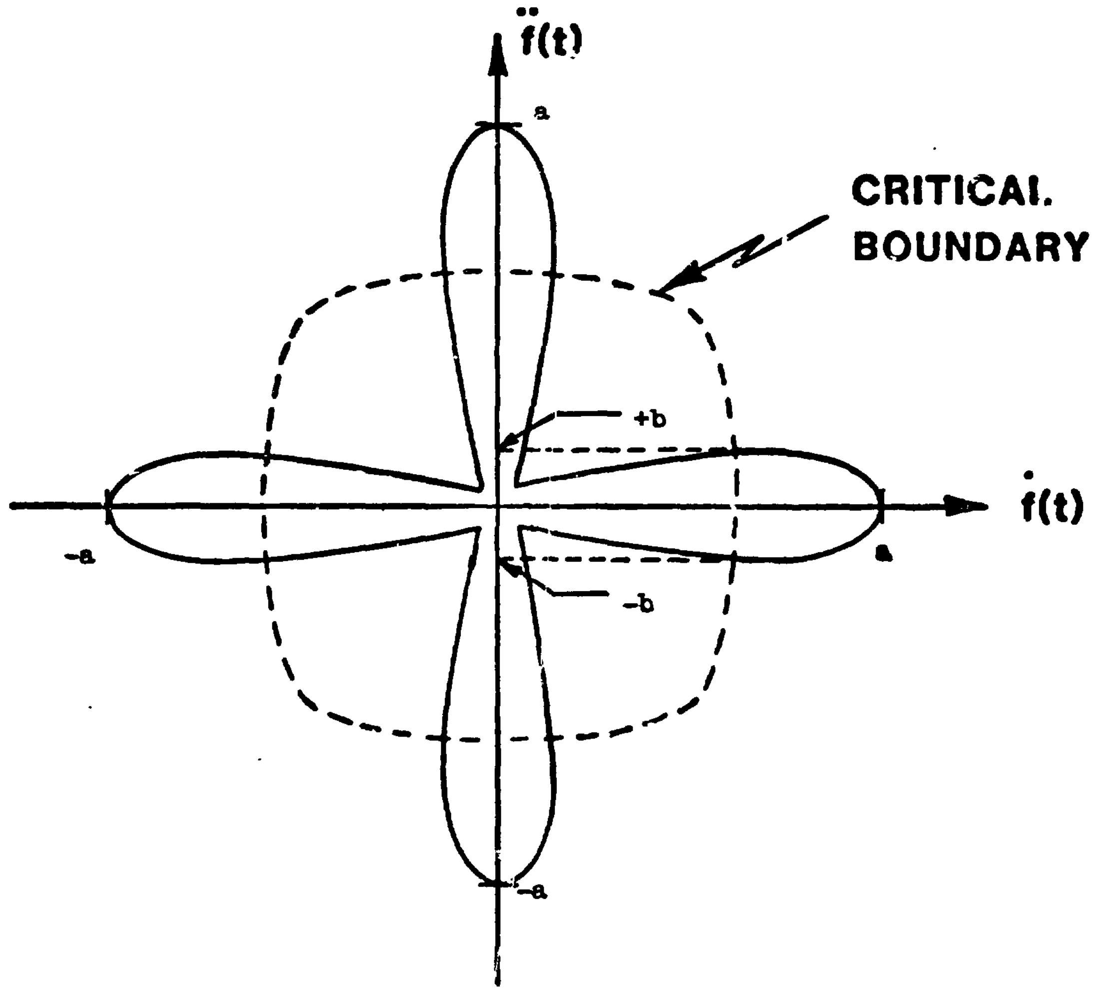

Figure (3) - SWEEP FUNCTIONS TO INVESTIGATE CRITICAL BOUNDARIES

ORIGINAL PAGE IS
OF POOR QUALITY

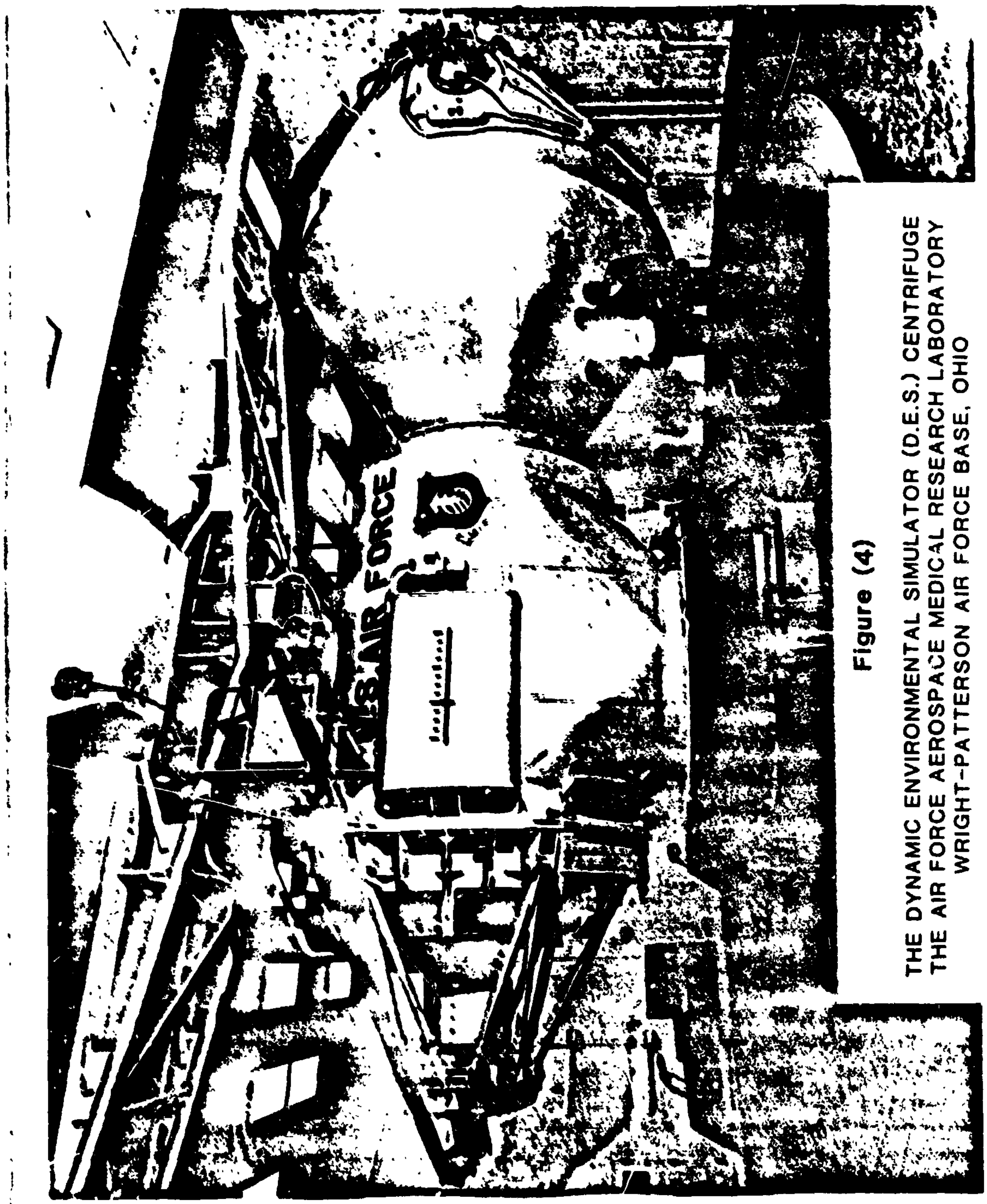

Figure (4)

THE DYNAMIC ENVIRONMENTAL SIMULATOR (D.E.S.) CENTRIFUGE
THE AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

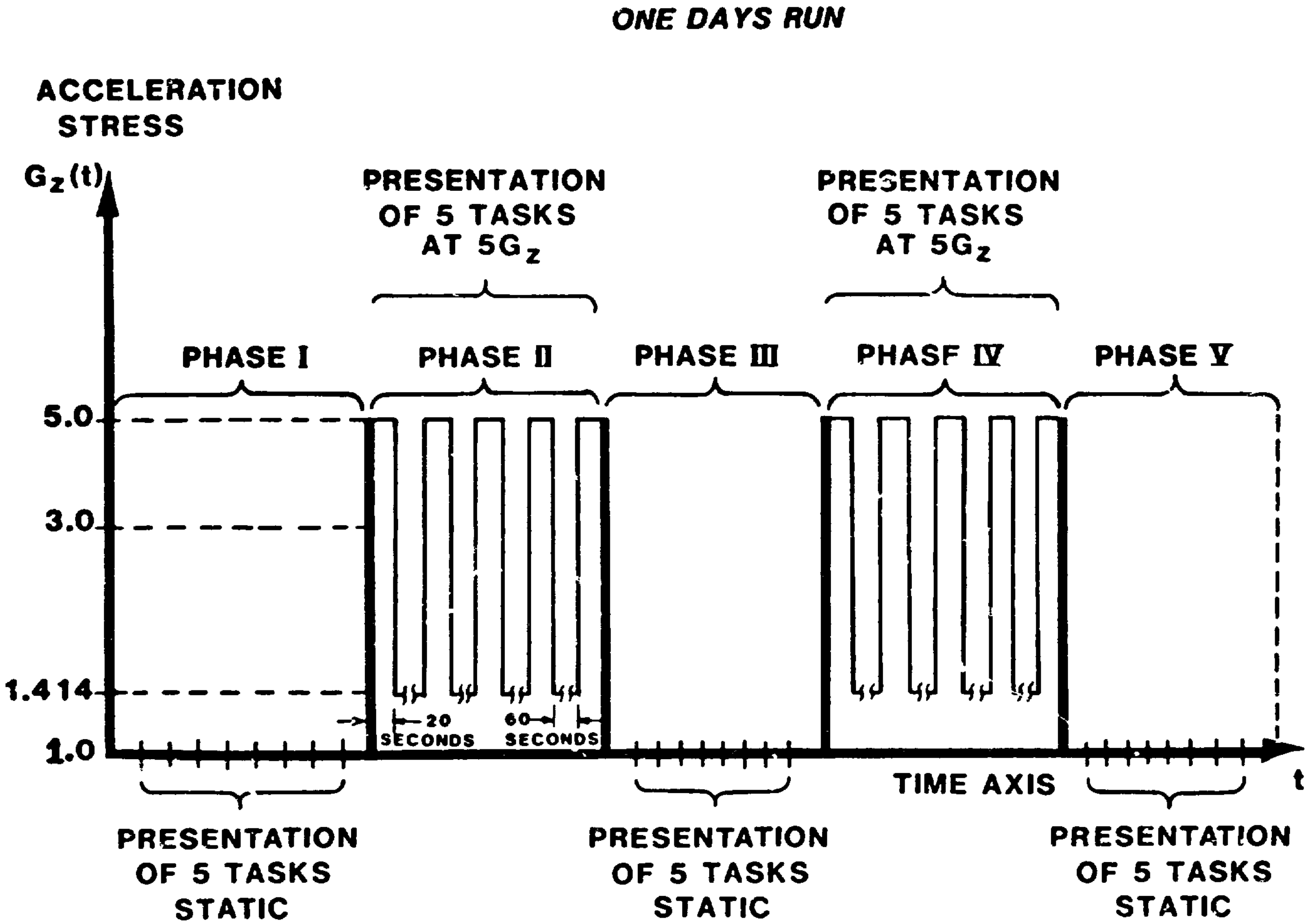

Figure (5) - A TEMPORAL DESCRIPTION OF THE G STRESS EXPOSURE

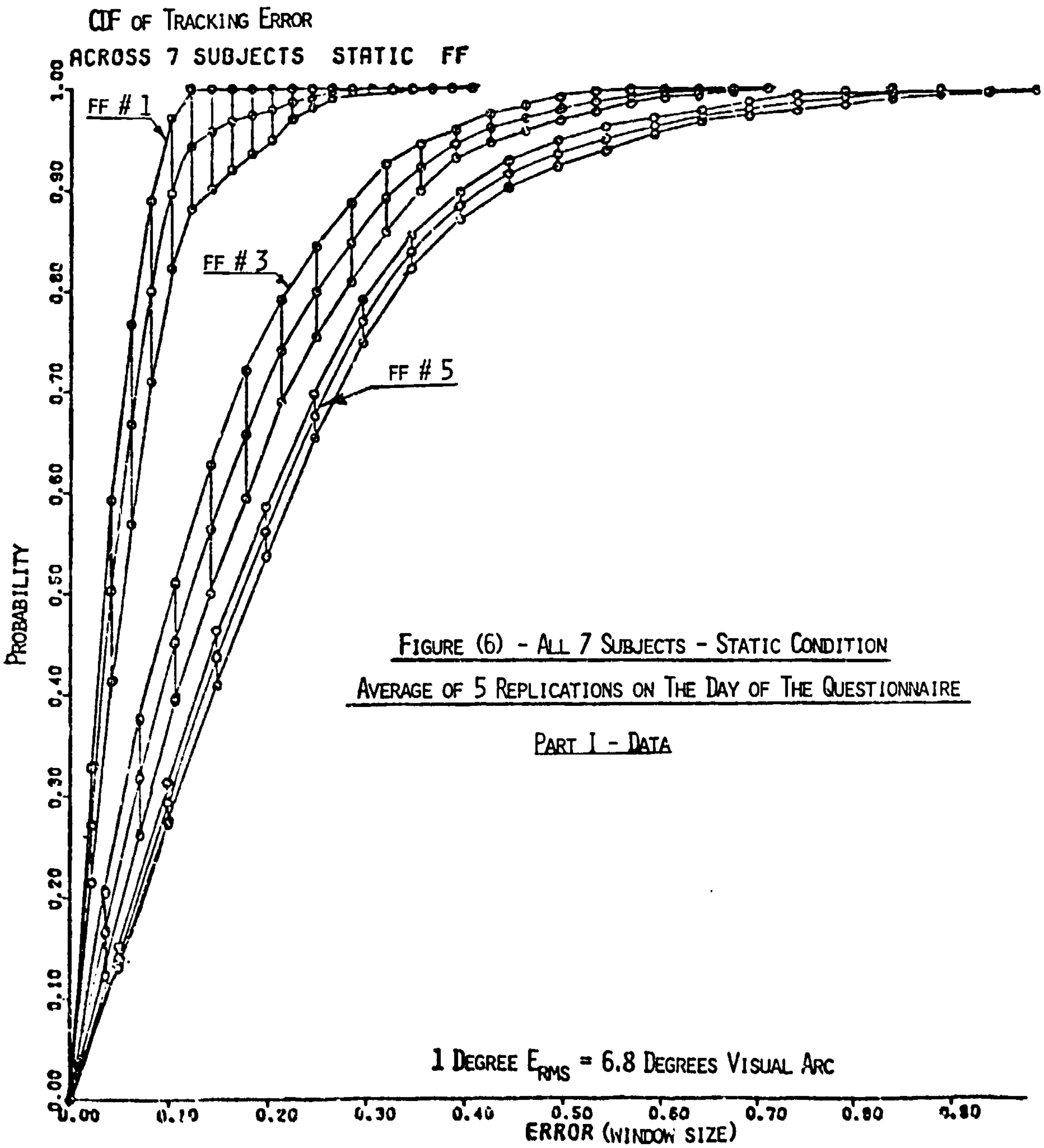

FIGURE (6) - ALL 7 SUBJECTS - STATIC CONDITION

AVERAGE OF 5 REPLICATIONS ON THE DAY OF THE QUESTIONNAIRE

PART I - DATA