(4) $7.60 N-63-4-3

(19)

Technical Report: NAVTRADEVCEN 323-1

(5) 338600

(1)

(6) THE NATURAL PILOT MODEL FOR FLIGHT PROFICIENCY EVALUATION

(7)-(9) NA

(10) by
Ezra S. Krendel
Joel W. Bloom

Engineering Psychology Branch
Franklin Institute
Philadelphia, Pennsylvania

(11) 15 Apr 63,
(12) 70 p
(13) NA
(14) NA
(16) NA
(17) NA
(18) NTDC
(20) U
(21) NA

15 April 1963

DDC
RECEIVED
JUL 30 1963
TISIA A

Prepared for

U.S. NAVAL TRAINING DEVICE CENTER

Port Washington, New York

$7.60

(15) Contract ~~No.~~ N61339-323

THE NATURAL PILOT MODEL FOR FLIGHT PROFICIENCY EVALUATION

ABSTRACT

This report presents the development and rationale for a new approach to pilot proficiency measurement in operational flight trainers. It is based on a "natural pilot model" that identifies three criteria as being of prime importance to the understanding and measurement of pilot performance: consistency of system performance, human adaptability, and least effort in skilled performance. By means of these criteria -- which arose from an effort to apply the servo-mechanism theory of skilled performance to the study of pilot proficiency -- the investigators believe that the traditional impediments to valid measurement will be removed; and that the characteristics that most crucially differentiate the good from the poor pilot will be measured. Ways of quantifying these criteria and the implications to training and further research are discussed.

FOREWORD

This study was initiated because of the great need for pilot proficiency evaluation that could be carried out in an operational flight trainer and that could give results useful both to the individual pilot for gauging his progress and to the squadron commander for determining the readiness of his men for flight missions. Despite the vast and often excellent research that has been carried out on pilot proficiency measurement, it can be said without fear of contradiction that every existing or currently proposed system suffers from considerable theoretical or practical difficulties. Not the least of these is the fact that the validity and reliability of any system vary with the aircraft or simulator type, with the trainees' previous experience in simulators and with the personnel who are doing the evaluation. Even apparently satisfactory measurement methods in operational flight trainers are of doubtful or unknown value in predicting ultimate mission success.

The basic purpose of this study was to develop a rationale for a practical, objective and valid system of pilot proficiency measurement that could be applied in operational flight trainers, and that would be free of the just-mentioned shortcomings.

METHOD: A four-fold approach was taken to the solution of this problem.

1. Existing and proposed flight proficiency measurement systems were critically surveyed.

2. Interviews and observations were conducted at several Naval Air Stations.

3. The knowledge gathered in the above two steps was examined in the light of the servo-mechanism approach to skill measurement.

4. An effort was made to develop a theory that would fuse a servo-mechanism viewpoint of pilot proficiency with conventional ways of studying and measuring human performance.

RESULTS AND IMPLICATIONS: The research yielded the "Natural Pilot Model" for studying the uniquely human aspects of pilot behavior which are the human's particular contribution to the manned aircraft system. Three aspects or criteria of the model are emphasized:

1. Consistency of system performance

2. Human adaptability

3. Least effort in skilled performance

The empirical and theoretical work led the investigators to believe that measures based on these three variables would get to the very heart of flight proficiency and yield information having greater validity and reliability, and yet be more easily obtained than any heretofore proposed. While these criteria are primarily concerned with the flight control (tracking) aspects of flight skills, it is very possible that they indirectly yield summary information on procedural skills as well. An experimental program is needed to specify which of the various alternate ways of obtaining measures of these criteria would be operationally most useful. The theoretical development accomplished during this study has yielded guidelines to aid in the quantification of such measures.

George Chajet
GEORGE CHAJET
Project Psychologist
U. S. Naval Training Device Center

TABLE OF CONTENTS

Page

BRIEF OF STUDY 2

STUDY REPORT 5

I INTRODUCTION 6

II SELECTING PERFORMANCE CRITERIA FOR PILOT AIRCRAFT SYSTEMS 7

- Introduction 7
- Servomechanism Performance Criteria 7
- Mission Phase Requirements 14

III CONSISTENCY OF PERFORMANCE 20

- Introduction 20
- Consistency - The Foundation of Prediction 20
- Consistency - An Attribute of Skill 27

IV ADAPTABILITY 31

- Introduction 31
- Adaptability to Control Dynamics 31
- The Adaptive Optimizing Human 33
- Criteria and Measurements 35

V ECONOMY OF EFFORT 38

- Introduction 38
- Minimizing Effort 38
- A Servomechanism Approach to Skill Development . . . 43
- Time in Control 48

BIBLIOGRAPHIC FOOTNOTES 52

APPENDIX 55

I THE PILOT AIRCRAFT SYSTEM 55

- Introduction 55
- Flight Control 55
- Pilot Control 60
- Flying Qualities of Piloted Aircraft 61
- The Correlation of the Quantitative and Qualitative in Aircraft Handling Qualities 63

LIST OF FIGURES

Figure		Page
1.	Feedback Control System	7
2.	System Response Characteristics	10
3.	Measures of the Transient Performance of Second Order Systems	11
4.	Criteria for the Step Function Responses of Second Order Systems	12
5.	Lead Pursuit Course	14
6.	Lead Collision Course	17
7.	Chart Showing the Effect of Change in the Selection of Samples of 10	21
8.	Data for Illustrative Control Chart	25
9.	Shewhart Control Chart	26
10.	Schematic for "Time to Adapt" Measurement	36
11.	Pure Compensatory Display	45
12.	Successive Organizations of Perception	45
13.	Perceptual Display Progression Sequence	47
14.	Schematic for "Time in Control" Measurement	50
15.	Aircraft Axis System	56
16.	Definition of Angle of Attack and Angle of Side Slip	57
17.	Static Stability	59
18.	Servo Model of Pilot-Airframe Attitude Hold System	61
19.	Autopilot-Airframe Attitude Hold System	64
20.	Hypothesized Linear Model of Human Operator Performing a Continuous Closed Loop Task	65
21.	Iso-Opinion Plot of Pilot Opinion Data	66
22.	Human Dynamics-Handling Quality Correlation	67

BRIEF OF STUDY

The purpose of this report is to present a new rationale for the evaluation of the flight proficiency of Naval Aviators. The principal incentive for this study was the need for a scientific basis for flight proficiency evaluation that could be undertaken in simulators and predict success in actual flight.

While accepting the validity of the skilled instructor's subjective evaluation of piloting performance which is in current use in the fleet, we suggest a number of objective measures for validation against the opinion of expert judges. These measures are devices for facilitating, not supplanting, current subjective evaluation procedures. Their use will reduce the burden now placed on flight proficiency evaluators; thus freeing instructor and commander to concentrate on those problems which resist objective measures. In addition, these new measures suggested in the study report provide insights into the learning process which open up a variety of new possibilities for improved training methods.

The estimation of the handling qualities of aircraft is an analogous problem where the need existed for objective measures to correlate with the opinion of expert pilots. In aircraft handling qualities evaluations, "standardized" pilots test "unstandardized" aircraft in much the same fashion as the proficiency evaluator employs "standardized" aircraft to test "unstandardized" pilots. In the last few years, a relationship between the opinions of "standardized" pilots on aircraft handling qualities and objective servomechanism descriptions of the "unstandardized" aircraft's control dynamics has been discovered. It is our thesis that objective servomechanism descriptions of the "unstandardized" pilot are therefore appropriate in pilot proficiency evaluations.

This approach is easily implemented in the evaluation of mechanical autopilots; however, requirements on human pilots are both more subtle and more general. Hence, servomechanism descriptions must be tailored to this particular problem. In terms of the descriptive Navy dichotomy, we seek to separate the "natural" pilots from the "mechanical" pilots. For it is the "natural" pilot who possesses the unique human skills so vital to the Navy's mission.

To this end, we present the "Natural Pilot Model" to specify those aspects of pilot behavior which are the human pilot's unique contributions to the manned aircraft system; and which therefore can be used to discriminate among pilots of differing ability.

Our development focuses on three aspects:

1. Consistency of System Performance

   It is necessary that the pilot's behavior be in a state of statistical control in order that meaningful predictions and estimates may be made. We are, however, interested in consistency of system performance as a basic attribute of skill as well. Such consistency of system performance may be achieved only by human performance which changes so as to cope with changes in the characteristics of the machine, the human and the external environment. Stability of overall performance therefore implies the following two pilot attributes.

2. Human Adaptability

   The ability to generate adaptive behavior as a function of changing dynamic requirements is not only an effective discriminator between man and machine but also serves as a basic parameter of fundamental skill ... of "natural" pilot ability.

3. Least Effort in Skilled Performance

> Sustained and stable performance, especially under changing dynamic requirements or overloading, can be achieved to a greater extent the more the human operator's performance is characterized by effort conservation (principle of "least effort").

The foregoing characteristics imply general criteria which we have developed in the study report. Before these criteria can be converted to operationally useful evaluation techniques, an experimental program for the detailed specification and validation of these criteria is required. Such a program is presented in outline form in the study report.

STUDY REPORT

CHAPTER I

INTRODUCTION

The specific skills which Naval aviators are required to master in order to perform their missions form an impressive list. The All Weather Flight Manual devotes sections to flight planning, basic instrument flight, confidence maneuvers, radio range procedures, and aerology, to mention but a few. The flight training syllabus of a typical fighter squadron includes gunnery, aerial refueling, carrier landings, emergency procedures, broadcast intercept, fighter tactics and many others. Thus, Naval flying involves a large complex of judgmental, procedural and perceptual-motor activities. Nevertheless it is clear that the essential skills required of the pilot are the ability to stabilize and to control his aircraft, i.e., to maintain the airplane in a desired flight condition despite transient disturbances; and to command changes in the aircraft orientation as needed. It is also clear that these essential flight control skills must be applied while the pilot is occupied or even overloaded with his other tasks.

This investigation is devoted to an analysis of the central pilot task of flight control and its interaction with the other elements of the manned aircraft system. To this end, we have ordered the following chapters so that the reader may become familiar with the general problem involved in the control of a pilot aircraft system, with the problems in selecting system criteria, and with those uniquely human aspects of pilot behavior which constitute the human pilot's special contributions to a manned aircraft system.

CHAPTER II

SELECTING PERFORMANCE CRITERIA FOR PILOT AIRCRAFT SYSTEMS

Introduction

The selection of the proper criterion for evaluation is often the central problem in system's analysis. In this Chapter we discuss the construction of criteria in terms of the qualities which the criterion must measure and in terms of the general requirement which criteria should meet. In order to present the problem in the light of practical requirements, we have illustrated the choice of criteria in several tactical examples.

Servomechanism Performance Criteria

Adequate feedback control systems had been designed long before the advent of rational performance criteria, just as skilled men had flown before the development of flight proficiency criteria. As control system and aircraft performance requirements became more severe, more accurate performance criteria became necessary. In this section we shall discuss advances in criterion development as they relate to our problem of assessing flight proficiency.

It is useful to begin our discussion with a simple diagram of a simplified feedback control problem, Figure 1. Any system which succeeds in slaving an output r(t) perfectly to an input i(t) (despite possible variations in input and disturbances) is one for which:

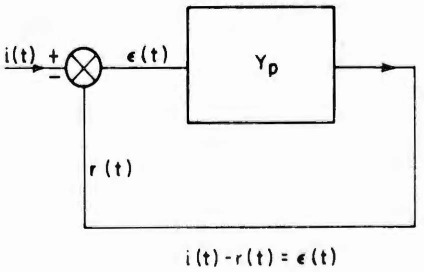


FIG. I. FEEDBACK CONTROL SYSTEM

$$r(t) = i(t) \qquad \text{(II-1)}$$

If the deviation or error in achieving this ideal control is defined by:

$$\varepsilon(t) = i\ (t) - r(t) \qquad \text{(II-2)}$$

then the ideal control system is one for which

$$\varepsilon(t) = 0 \qquad \text{(II-3)}$$

This control system may be specified mathematically by its closed loop transfer function, Y(s). A transfer function is a mathematically convenient device for specifying the relationship between an input and an output to a control system. The closed loop transfer function for Figure 1 is defined as the ratio of the Laplace transforms of the output r(t), i.e., R(s), to that of the input i(t), i.e., I(s), for initial values set equal to zero. Therefore:

$$Y(s) = \frac{R(s)}{I(s)} = \frac{Y_p}{1 + Y_p} \qquad \text{(II-4)}$$

No real system can satisfy the idealized performance specified by Equation 3 at all times. Instead the error usually has some finite value which is conveniently expressed as the sum of two terms; one corresponding to its operation in response to changes in command and the other to its operation in response to disturbances.

Servo system errors can be classified further as to whether they refer to events which have occurred after the system has settled down to steady values or whether they refer to transient behavior. We will first make some remarks about steady state errors, and then discuss transient behavior.

The form of Y(s) determines the steady state error of a servo system. In general, the transfer function of a linear system may be expressed in normalized form as

$$Y(s) = \frac{R(s)}{I(s)} = \frac{p_m s^m + \dots p_2 s^2 + p_1 s + p_o}{s^n + q_{n-1} s^{n-1} + \dots q_2 s^2 + q_1 s + 1} \qquad \text{(II-5)}$$

where $m \leq n$

Of the countless tranfer functions which can fill the requirements of the form of Equation 5, we will limit ourselves to three cases which are of considerable practical importance. These cases are those where the form of the numerator is:

1. 1
2. $p_1 s + 1$
3. $p_2 s^2 + p_1 s + 1$

Case 1. represents the condition where the steady-state system error after a step input is zero.

Case 2. describes a condition for which the steady-state system error is zero when the system is subjected to a step input of velocity.

Case 3. describes the condition where the error is zero to a input in acceleration.

The three systems are accordingly called zero position, zero velocity or zero acceleration error system, respectively. These terms are perhaps the most commonly used basic classifications for feedback control systems and they conveniently characterize the steady state performance of given devices. The type of system; i.e., the form of the numerator of Equation 5 selected for a specific task obviously is directly related to the type of command inputs anticipated; and the selection of a particular system type from the three described automatically supplies some criterion for design and subsequent evaluation.

While the three classifications listed offer reasonably adequate criteria and descriptions for steady state performance the total dynamic picture is neither specified nor clear. In order to obtain a total picture we must include a description of the control system's transient behavior as well.

In order to accomplish this more complete formulation of system criteria, Graham and Lathrop (1953) have argued for the convenience of some unitary figure of merit which might be used as a guide to general servomechanism design or evaluation, including transient performance. Three basic attributes that such a criterion should possess are:

1. Reliability
2. Ready Applicability
3. Selectivity

Clearly, these attributes provide a general basis for the evaluation of criteria and are not limited to this particular servomechanism context.

Subject to the necessary condition of stability, and limiting their discussion in the main to linear servomechanisms with zero steady state position error, Graham and Lathrop examined various possible criteria for transient responses to step inputs. The examples were mainly for second order systems, e.g., a mechanical system in which an accelerated mass is subject to a linear spring restoring force and viscous friction, or a series inductance, resistance, capacitance circuit. The normalized transfer function for such a second order system is:

$$Y(s) = \frac{1}{s^2 + 2\zeta s + 1} \qquad (\zeta \text{ is the damping ratio}) \qquad \text{(II-6)}$$

They continue with some discussion of third order systems, and a restricted mention of fourth and fifth order systems.

Three commonly used characteristics of system response which have been used as performance criteria are illustrated in Figure 2:

1. Time - for the error to reach its first zero (t_o)
2. Overshoot percentage - expressed in terms of the initial error
3. Solution Time - defined as the time for the error to decrease to and remain within 5 per cent of its initial value ($t_{5\%}$).

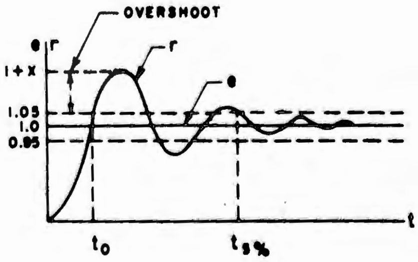


FIG. 2. SYSTEM RESPONSE CHARACTERISTICS

In Figure 3 these three qualities are compared for a second order system with the damping ratio, ζ as a parameter. That time to first zero and percentage overshoot are conflicting criteria is clear. Solution time appears to be a fairly good criterion. It is, however, excessively selective in that it gives an exaggerated picture of the difference between a system at the optimum and one quite close to this optimum. Among the several other criteria which have been suggested. we list:

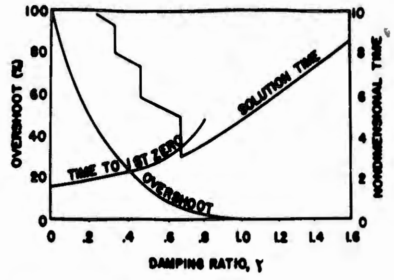


FIG. 3. MEASURES OF THE TRANSIENT PERFORMANCE OF SECOND ORDER SYSTEMS

GRAHAM AND LATHROP (1953)

1. Integral of Error

$$I_1 = \int_0^{\infty} \varepsilon(t)\, dt \qquad \text{(II-7)}$$

This criteria, also known as control area, is satisfactory for systems which do not overshoot. For systems which overshoot it gives an erroneous indication since an overshoot will decrease the value of the integral.

2. Time Weighted Control Area

$$I_2 = \int_0^{\infty} t\, \varepsilon(t)\, dt \qquad \text{(II-8)}$$

This provides for the weighting of the error with time, which provides an increasingly heavy penalty for a sustained error. This results in the same objection with respect to overshoots as the control area criteria.

3. Integral of Squared Error

$$I_3 = \int_0^{\infty} \varepsilon^2(t)\, dt \qquad \text{(II-9)}$$

The use of this form insures that the transient is not too slow since the values of the error at the beginning of the transient are predominant. The squared error criterion does, however, tend to favor underdamped systems because of the presence of a square under the integral sign which results in large errors being emphasized.

4. Integral of Absolute Error

$$I_4 = \int_0^{\infty} |\varepsilon(t)|\, dt \qquad \text{(II-10)}$$

The use of the absolute value instead of the error squared in this form would tend to eliminate the bias in favor of the underdamped systems.

5. Integral of Time Weighted Absolute Error (ITAE)

$$I_5 = \int_0^{\infty} t\, |\varepsilon(t)|\, dt \qquad \text{(II-11)}$$

The ITAE tends to penalize long duration transients by giving greater weight to error occuring late in the transient.

In Figure 4 these five criteria for transient performance are examined for a normalized second order system.

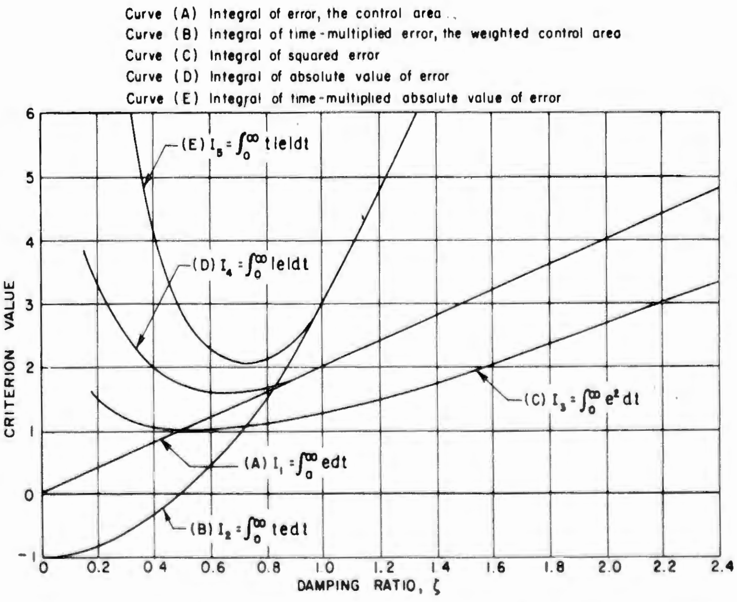


FIG. 4. CRITERIA FOR THE STEP FUNCTION RESPONSES OF SECOND ORDER SYSTEMS

Graham and Lathrop (1953)

Graham and Lathrop (1953) demonstrate that ITAE is the most convenient and desirable criterion to use from both a consideration of the previously listed desirable attributes as well as from the viewpoint of analog computer analysis. Figure 4 makes quite clear the good selectivity available with the ITAE.

A measure such as the ITAE is properly used as a measure of transient performance where the steady state value of the error is zero. The reason is:

If $|\varepsilon(t)| \neq 0$, then, as t becomes large without limit, the integral

$$\int_0^\infty t \, |\varepsilon(t)| \, dt$$

also becomes large without limit. Clearly this is not a desirable property for a criterion.

A proper criterion measure for many systems for which $\varepsilon(t)$ does not approach a steady value of zero is the rms value of the error. This measure is capable of handling conditions which are to be expected in practice when the system is subjected to a distribution of random inputs.

These random inputs exciting the system constitute a steady state condition in a statistical sense. As a consequence a unitary statistical measure such as the rms error

$$\varepsilon_{rms} = \sqrt{\frac{1}{T}\int_0^T \varepsilon^2 (t) \, dt} \qquad \text{(II-12)}$$

is indicated. A more detailed measure of the statistics of the error can be obtained from the spectral density of the error signal, $\phi_{ii}(\omega)$, which describes the error signal in terms of its predominant frequencies. The spectral density and the rms error are related as follows:

$$\varepsilon(t)_{rms} = \frac{1}{2\pi}\int_0^\infty \Phi_{ii}(w) \, d\omega, \qquad \text{(II-13)}$$

for ε_{rms} measured over long time durations. This frequency function based on the error signal is useful information for evaluating dynamic man-machine systems since the effects of various frequency components in the error are related to the resonant modes of the complex system dynamics. For example, it is clearly undesirable for a pilot to excite the structural modes of his controlled element.

Mission Phase Requirements

In the preceding section we have reviewed the development of general servomechanism criteria. In this section we shall outline the selection of criteria according to the demands of specific tactical situations.

We will begin with a typical intercept problem - the gunnery firing attack. During the final attack phase the pilot's control process consists of flying the airplane at zero sideslip on a lead pursuit course with the target, Figure 5. He must proceed on this course for a considerable period of time, including the two or three second period when he is actually firing his armament. During this last interval three conditions must be met:

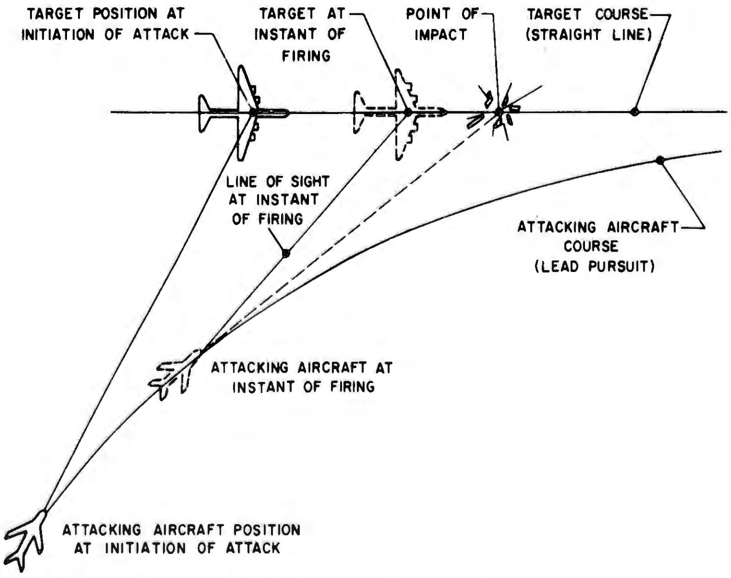


FIG. 5. LEAD PURSUIT COURSE

1. The sideslip is approximately equal to zero. (to reduce windage error)
2. The reticle, or other means of generating course, is tracking the target.
3. The range is correct.

The basic general criterion in this situation is, of course, kill probability. Kill probability computations involve the complex interactions of target vulnerability, weapon and fire control effectiveness, precision of tracking, and so forth. Assuming that the pilot is not employing tracers or other auxiliary means for correcting his course and must rely solely on target image in his gunsight, the best he can do is to fly his course adequately. The figure of merit then must be tied directly to his ability to fly the course within defined limits. An appropriate form of the error criterion would then be the rms error of his flight path during the actual firing period. This would take the form of:

$$\varepsilon_{rms} = \sqrt{\frac{1}{T}\int_0^T \varepsilon^2 (t)\, dt} \qquad \text{(II-12)}$$

Also appropriate would be the averaged absolute value of the error

$$\overline{|\varepsilon(t)|} = \frac{1}{T}\int_0^T |\varepsilon(t)|\, dt \qquad \text{(II-14)}$$

or the mean square error

$$\overline{\varepsilon^2 (t)} = \frac{1}{T}\int_0^T \varepsilon^2 (t)\, dt \qquad \text{(II-15)}$$

It will be noted that it is necessary to remove the algebraic sign from the function, as a miss in any direction is equally detrimental to the system. As the gunnery run is a total failure if the error exceeds a fixed value it is necessary to set this maximum allowable value for the error on the basis of a kill probability analysis.

The use of tracer ammunition complicates the problem somewhat. Although, a gunsight mechanism can perform effective fire control computation, kill probability is increased if the pilot can introduce corrections based on observations of where his shots are going. Indeed, as legend has it, tracer ammunition was not available in the Korean theatre, since it had been thought that the A1 Gunsight would make such ammunition obsolete. However, it was only after effective makeshift tracer ammunition became available, that the superiority of our A1-equipped aircraft over the MIG-15 became evident.

The inaccuracies of tracer information are well known, but the use of tracers modifies the problem in that the error can now be observed as a miss instead of an angular quantity seen on the sight. Additional information has been added for the pilot's use, and the criterion will now have to be some function of both miss and angular error. The miss is dependent upon the error existing one tracer time of flight, τ, in the past, i.e. $\varepsilon(t - \tau)$, so criteria based upon some function of $\varepsilon(t)$ and $\varepsilon(t - \tau)$ are appropriate. If the pilot operated on a purely linear combination the $\varepsilon(t - \tau)$ term would allow the helpful effect of lead, since it can be shown that the pilot's input can then be expressed in Laplace transform notation as follows (where K is a constant):

$$\begin{aligned} \text{CRITERION} = \varepsilon(s) - K\,\varepsilon(s)\,e^{-\tau s} &= \varepsilon(s)\,[1 - K(1 - \tau s + \ldots)] && \text{(II-16)} \\ &= \varepsilon(s)\,[(1 - K) + K\tau s] && \text{(II-17)} \\ &\simeq \varepsilon(s)\,K\,[1 + \tau s] \end{aligned}$$

If tracers are used and the miss at the target becomes so dominant that this quantity is effectively the error, i.e. K becomes very large, then it would be necessary to time weight this effective error since the miss distance near the end of the two or three second gunnery run period is most critical. Thus the form would be:

$$\text{ITAE} = \int_0^T t\,|\varepsilon(t)|\,dt \qquad \text{(II-18)}$$

This form differs from the conditions when only the sight is used - where with no external correction available an error at any point within the run is of equal weight. In both cases, however, an upper limit restriction on the error is still appropriate. Actually, the foregoing describes a conservative estimate of the improvement possible. A clever pilot could generate various non-linear modifications of his response to further improve performance over that predicted by the linear model.

A rocket bearing interceptor poses a somewhat different problem. The pilot ordinarily flies a lead collision course (Figure 6) and releases his armament at essentially a single instant along the course. Although the pilot must track the enemy over a period of time before release, in order to provide inputs to the computer system, the error at the instant of release is the significant error. So, for a given run, the figure of merit is simply the absolute value of the error at that time. In order to assess a pilot it would, of course, be necessary to score a number of runs. A series of such trials would yield a distribution of errors and the appropriate criterion would be an ensemble error averaged over the series.

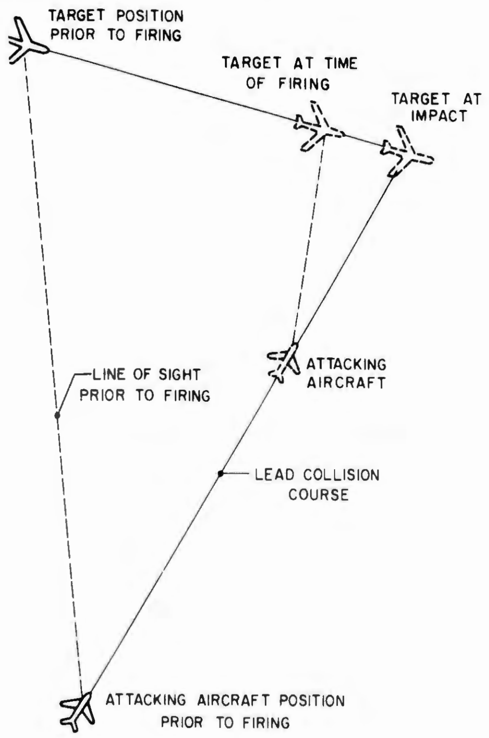


FIG. 6. LEAD COLLISION COURSE

AVERAGED ENSEMBLE ERROR

$$= \frac{1}{N} \sum_{i=1}^{n} |\varepsilon_i| \qquad \text{(II-19)}$$

Our criteria may be identical in form to these employed in the gunnery firing attack, but they would be generated over a series of single point trials rather than on a continuous basis.

If the interceptor is armed with a guided type weapon we have a situation quite similar to that of the rocket case. Here again the pilot usually flies a lead collision course. Actual launching, however, is at a considerably greater range than is possible with the rocket. The criterion is again based on point of launch with some latitude with respect to aim allowed as the weapon will receive guidance signals while in flight. The need to provide this guidance may, however, impose a limitation on the pilot's freedom to maneuver after weapon release if he must keep the enemy continuously surveyed by his radar.

High yield weapons for which the lethal volumes are very large imply a different type of criterion. Here there is a distance measure within which the error must fall; i.e.,

$$|\varepsilon| < \alpha$$

This constant α is determined by the lethal volume of the weapon, and the complexity and time requirements of the tactical breakoff maneuvers required by the attacking aircraft.

The carrier landing situation provides an example of a somewhat different form of error criterion. Landing on a carrier is an exceedingly difficult task. At a certain point off the fantail the aircraft's attitude and speed must be held within very closely restricted limits. In fact the point at which the pilot is committed by the LSO (Landing Signal Officer) is a "point of no return" - engine failure past this point for example, would almost certainly result in a crash. Not only must attitude and speed be closely controlled but the aircraft must be in a steady state condition, that is to say neither accelerating nor maneuvering when the LSO gives the "cut" signal. Thus the final stage is almost a ballistic drop in order to retain these conditions. Here as in the missile fire control situation there is a single point at which the pilot seeks to attain certain conditions. However, in the missile case if the launching were to take place while the aircraft was passing through the zone of correct conditions, a perfect score would be registered despite subsequent deviations. In the landing approach the problem is far more complex. The pilot is engaged in a final value synthesis task in which he seeks to minimize the system error at a future point in time. Thus the error must be very heavily time weighted for a meaningful criterion. For example, if we consider the process to begin at some relatively early point in the final approach path the penalty attached to an error increases sharply as the "cut" point is approached. The rationale here is that while an early error is correctable, correction becomes increasingly difficult and finally impossible. Thus an appropriate criterion might be:

$$\text{CRITERION} = \int_0^T \frac{|\varepsilon(t)|}{T-t}\, dt \qquad \text{(II-20)}$$

where T-t is the time to go. Or in somewhat simpler form but not so heavily time weighted:

$$\text{ITAE} = \int_0^T t\, |\varepsilon(t)|\, dt \qquad \text{(II-18)}$$

The second form is somewhat easier to mechanize.

We have illustrated how specific performance criteria are selected as functions of the pilot - aircraft tactical mission. These criteria are applicable equally to human pilots or to auto pilots. We have the problem of extending this development to arrive at criteria singularly appropriate to the evaluation of the human pilot. To do this we must have criteria which take into account his ability to profit from experience and manifest adaptive, optimalizing behavior - in a fashion no machine has yet approached. In short we seek criteria to separate the "mechanical" pilot from the "natural" pilots.

CHAPTER III

CONSISTENCY OF PERFORMANCE

Introduction

Consistency, i.e., statistical control, has two important aspects in our usage. First, that measurements are samples from a stable population is a prerequisite for meaningful measurement. Second, consistency in a man-machine system output illustrates a fundamental human ability. We examine the pattern of consistency in sufficient detail to show that consistency of performance in terms of the output of the overall man-machine system is achieved only by adaptable human performance which copes with the machine and environment variability. Statistical stability of overall performance can be derived from two hypothesized attributes of the pilot:

a) Ability to adapt behavior to meet the changing requirements imposed by the controlled element and the environment.

b) Minimum energy mode of operation so as to permit sustained performance.

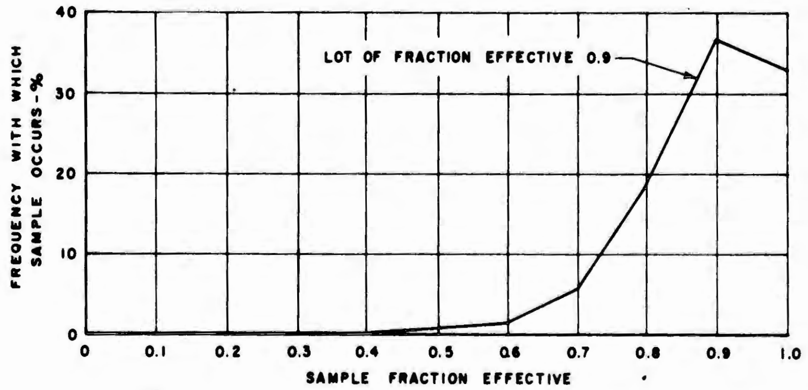
Consistency - The Foundation of Prediction

In the acceptance testing of ammunition by destructive means it is particularly clear that a selective sampling procedure is needed. The testing agency samples the product, subjects it to tests and generalizes about the untested ammunition on the basis of what was determined from the tests. Sampling is necessary in other acceptance and assessment procedures because of the obvious requirement imposed by feasibility and economy. Clearly, the assessment of Naval Aviators' flight proficiency must be based on a selective sampling procedure. Thus a primary criterion of flight proficiency prediction must be consistency of the system output.

The skilled instructor is able to assess consistency by an intuitive weighing of the time history of the observed flight processes. Indeed the current fleet practice of decreasing the number of trials for experienced pilots in instrument checks*, is an example of this. Intuitive judgements of this nature are often capable of remarkably high validity — witness the day to day successful operations of the Fleet Air Arm.

Some notion of the difficulties inherent in intuitive approaches to consistency measures through sampling techniques can be developed by returning to the ammunition testing situation. This deliberate simplification is used to clarify the basis upon which subsequent concepts will be structured.

Figure 7, which is derived from the binomial expansion for independent random sampling, from Simon (1941) shows the relationship in quality between a small sample and a large lot where the fraction effective (proportion of good pieces) of the large lot is 0.9. The pilot's total performance repertoire at a given stage of his skill



7

FIG. 7. CHART SHOWING THE EFFECT OF CHANCE IN THE SELECTION OF SAMPLES OF 10

Simon (1941)

---

* Consensus of interview data developed by The Franklin Institute personnel at NAS Patuxent River, Maryland. Although anecdotal in nature, such evidence is nevertheless pertinent at this point.

development corresponds to the large lot and his performance during test trials to the sample. Thus we are interested in what type of performance might be expected in, for example, ten independent trials out of a large population where the population probability of a success is 0.9. From the chart it is apparent that from a population of fraction effective 0.9 a sample of 10 will be perfect 35% of the time, 0.9 effective 39% of the time, and of poorer fraction effective than the lot only 26% of the time (100%-39%-35%).

It is possible to improve the efficiency of a small sample by the use of variable sampling; e.g. classifying the sample quality along a continuum, instead of attribute sampling which classifies dichotomously as either acceptable or unacceptable. This improvement is not always possible because of the difficulty of constructing appropriate scales for many of the complex parameters of the process for such assessment problems as a flight performance. In general, as Simon (1941) points out:

> "For most purposes sample sizes of the order of 300 to 600 are a minimum for achieving satisfactory working estimates of the lot fraction defective from the evidence of a sample and that very small samples are not only practically worthless for distinguishing between lots (except where defectives are overwhelming) but likely to be positively misleading."

As a consequence a statistician might be inclined to recommend a large number of samples of pilot behavior to measure flight proficiency with confidence. We are aware, however, of the day to day success of the experienced evaluator who employs samples of small sizes. Why this apparent divergence between sound theoretical statistical knowledge and empirical experience? The answer is quite simple, the experienced evaluator does not treat the samples as isolated events. He draws upon his memory of the history of the process — to enrich each sample.

In observing a student engaged in the process of learning to make a carrier landing the experienced instructor bases his evaluation not only on the results of the trial under consideration but on the student's behavior relative to the time history of the entire learning process. This intuitive and clearly valid and reliable practice provides an explanation of the correlations among subjective grades reported by many investigators - Ben-Avi (1947). Indeed the failings of the small sample item by item "objective check list" may in no small measure be attributed to precisely this. For while the "objective check list" in effect forces individual small sample evaluation it makes no attempt to employ data on past performance to supplement the information adduced from the sample.

Our goal is to develop an objective methodology by which the information available from both present and past flight performance can be used to effect maximum predictability of future performance. Converting this process from a highly developed individual skill to a set of systematic procedures will reduce the skill requirements of flight proficiency evaluation and free the instructors to concentrate on those aspects of their craft which defy objective statement.

A well tested technique employed in both industrial quality control and work measurement which accomplishes the foregoing is the Shewhart Control Chart - Abruzzi (1956).

The control chart provides a convenient technique for determining that a "constant system of chance causes exists" and to state that prediction within limits is then possible. In essence, we recognize that the behavioral processes of flying are always subject to a certain amount of variation within the pattern of a stable population and occasionally to "assignable cause variation" as well which indicates that a stable population cannot be assumed to exist. The control chart then serves to provide criteria to discriminate between these conditions.

The use of industrial work measurement techniques in contrast to aptitude testing and psychophysical measurement as the conceptual basis for flight proficiency measurement requires comment. Industrial work measurement involves a multifaceted behavioral process of relatively long periods of time involving man-machine interaction under conditions where the operator's objective is to maintain consistent output in the face of variation in tool, machine, material and personal fatigue. Aptitude testing and psychophysical measurement on the other hand involve severely restricted and controlled stimulus situations. Situations in which the operator is repeatedly exposed to homogeneous stimuli and expected to make almost homogeneous responses such as "yes", "no", "equal", "one-half", etc. The Naval Pilot's striving to maintain consistently high performance with his weapon system clearly suggest the applicability of work measurement techniques.

An example is of value to elucidate this notion. Let us assume that we are measuring a given parameter of pilot proficiency. The causes of variation in this parameter may change in magnitude over time or they may come and go due to changes in the pilot's physiological state, the aircraft, the recording equipment, and the like. In a short space of time it is reasonable to assume, however, that new "assignable causes" will not develop. If we then sample performance from a few successive trials, such variation as occur may well be attributed to chance variation. If successive groups of samples are taken at intervals sufficiently spaced in time, variation resulting from changes in the cause system, i.e., a change in the pilot's basic capability level, can enter the system. If we were to examine 25 samples each containing five elements, spaced at uniform intervals along the pilot's practice sessions, Figure 8, there would be an inter and an intra sample variation. It will be noted that the time order of the data has been preserved. We can estimate the standard deviation, which is part of the constant system of chance causes, from the samples themselves. If there are no changes in the cause system, then the arithmetic means of the samples will behave approximately like a normal distribution and almost all of the sample means will lie within plus or minus three standard deviations of the sample means from the grand mean of all the samples. If all the sample means lie within these limits, which we shall term "control limits", then we can predict that subsequent sample means will lie within these limits, providing, of course, that the underlying cause system is unchanged. This situation is readily portrayed on a Shewhart Control Chart, Figure 9, where the behavioral processes may be said to be in a state of statistical control. The actual limit criteria were computed as follows:

Sample	Element					$\bar{X}$	R
No.	a	b	c	d	e		
1	4	5	4	4	4	4.2	1
2	4	7	4	5	5	5.0	3
3	4	5	4	4	8	5.0	4
4	6	6	7	5	6	6.0	2
5	0	4	8	6	6	4.8	8
6	4	8	6	5	9	6.4	5
7	5	3	4	3	2	3.4	3
8	5	4	5	4	4	4.4	1
9	2	6	7	4	5	4.8	5
10	3	6	6	8	5	5.6	5
11	7	6	3	4	1	4.2	6
12	6	7	7	6	5	6.2	2
13	5	4	3	3	4	4.8	2
14	4	1	8	8	6	5.4	7
15	8	5	7	5	6	6.2	3
16	6	4	6	5	5	5.2	2
17	5	5	4	4	4	4.4	1
18	2	4	5	4	3	3.6	3
19	4	6	2	6	3	4.2	4
20	4	3	6	6	6	5.0	3
21	6	7	8	4	4	5.8	4
22	5	4	6	5	5	5.0	2
23	8	3	4	7	2	4.8	6
24	4	6	6	4	4	4.8	2
25	4	5	4	7	6	5.2	3

$$\bar{\bar{X}} = \frac{\bar{X}}{n} = 4.976 \qquad \bar{R} = \frac{R}{n} = 3.48$$

FIG. 8. DATA FOR ILLUSTRATIVE CONTROL CHART

$$UCL\ \bar{X} = \bar{\bar{X}} + A_2\bar{R}$$

$$LCL\ \bar{X} = \bar{\bar{X}} - A_2\bar{R}$$

where UCL and LCL are the upper and lower control limits respectively. $\bar{\bar{X}}$ represents the mean of the sample means while $\bar{R}$ represents the mean of the corresponding ranges. The A values are readily obtained from tables found in the quality control literature; see for example Grant (1952). Parenthetically, it should be noted that the plus or minus three standard deviations employed as the control criteria are merely a tentative suggestion which conforms to much current industrial quality control practice. The actual criteria selected should attempt to minimize the "type I" and "type II"

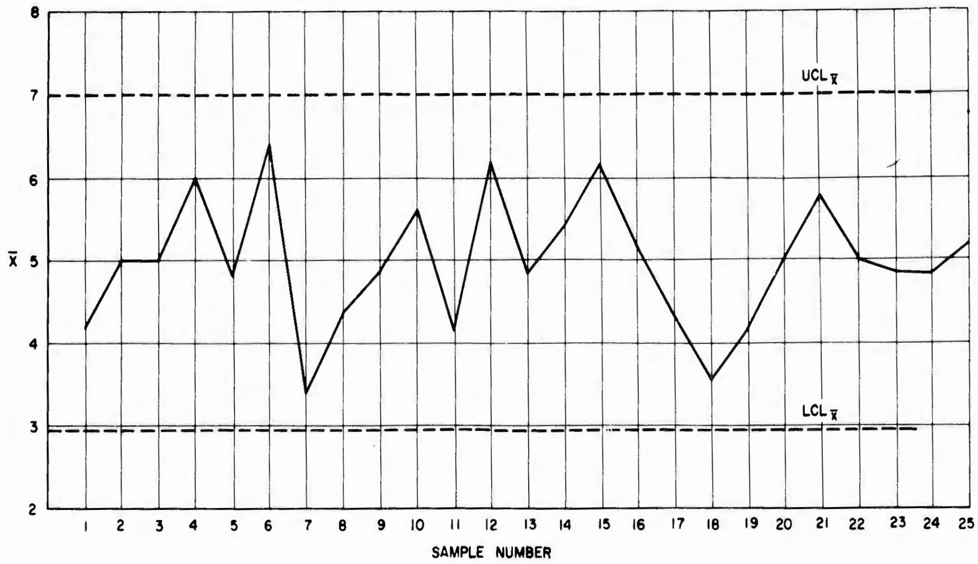


FIG. 9. SHEWHART CONTROL CHART

errors. A type I error involves claiming statistical stability exists when it really doesn't, and a type II error is the denial of statistical stability when it really exists.

As the costs associated with errors of these two kinds in a Naval training situation are quite different from those obtaining usual industrial practice, it will be necessary to establish stability criteria in the context of protocol and actual experimental evidence. In like manner, the choice of sample size represents an attempt to achieve a balance between two desirable and contradictory goals. The first goal is to minimize the chance of including significant variation within the sample. This goal demands small samples. The second goal, which is to have normally distributed sample means, requires large sample sizes, of course. Here again, the form of the actual populations to be encountered and the costs associated with the sampling process will determine the decision.

Consistency - An Attribute of Skill

On an intuitive basis it might appear that the performance of machines would always be more consistent than that of men. Interestingly enough, with complex equipments, the opposite is more often the case. Even such relatively simple devices as radar equipment often exhibit surprising variability compared with their human operators. In an unpublished study cited by Chapanis (1959), variations between individual radar sets of the same model were found to be much larger than differences among operators. Indeed, in his studies on detectability on radar scopes, Williams (1949) found so much variability in cathode ray tubes that he could not use electrical measurements of circuit parameters as the independent variables. In the case of entire aircraft systems the machine variability is, of course, substantial. As a matter of fact, pilots tend to characterize certain of supposed identical individual craft as "mushy" and others as "touchy". In a similar vein, Bray (1958), reports on experiments on interceptor gunnery performances as measured by actual firing on targets. Variability between pilots was normally of little importance compared to variability within and between aircraft; he goes on to state:

> "In other words, the hardware varies within itself
> and from one item of hardware to another so greatly
> as maintenance and adjustment are normally carried
> out that the pilot matters little, if at all."

The significant value of man in terms of his ability to increase <u>system</u> reliability — measure of consistency of system output — has been dealt with by McRuer, Ashkenas, and Krendel (1959). in connection with the current manned vs unmanned space craft controversy; and Anderson (1958), has presented a quantitative discussion of the increased system reliability afforded by the use of manned space vehicles in comparison with unmanned vehicles.

Thus it appears that consistency of system output is not only an appropriate measure in terms of insuring validity of measurement but is a performance measure particularly appropriate to the human operator in the man-machine complex. For it is man by the exercise of his peculiarly human attributes who contributes to overall system output consistency.

To examine the underlying human skill which results in consistency of system output we shall refer to industrial work measurement data. Whitehead (1938), in reporting the Hawthorne experiments noted that the production rates of experts were more stable than the production rates of novices. Loveday and Munro (1920), noted similar findings in reporting that highly skilled workers had more uniform production rates than other workers. Elton (1922), in studying the time required for each of a group of weavers to complete a series of warps concluded that:

> "...considering the large number of causes at work which affect output the consistency of these figures which have been taken at random and have not been specially selected is remarkable."

McRuer and Krendel (1957) examined of a number of tracking studies including those by Benepe, Narasimhan and Ellson (1954) and Elkind (1956) and concluded that variability in tracking tasks without controlled element dynamics was dependent on both task and degree of subject training. For our purposes, level of skill and degree of training may be considered to be correlated.

Shaw (1958) in a study of automotive driving skill tested three groups of drivers: skilled police patrol-car men, skilled civilian drivers, and below average civilian drivers (the assessments were made by an experienced police driving instructor). He found highly significant differences in consistency of performance with the police drivers exhibiting higher consistency than the skilled civilian drivers and the latter in turn more consistency than the unskilled civilian drivers. Shaw suggests that not all consistency is indicative of skill, "for obviously bad habits can be acquired" but that perhaps all inconsistency is indicative of lack of skill.

The foregoing studies then impute some generality to hypothesis that consistency of output is a necessary concomitant of skill. Recalling, however, that man in a complex system must interact with time variant machine and environmental factors , it is quite clear that this consistency of overall system output can only be achieved by an essentially variable behavioral process. Abelson (1953) reports on some conversational remarks by Shewhart to the effect that he "Has concluded from his own private observations that it is virtually impossible to find a series of stereotyped human performances which are in control."

Abruzzi's (1956) experimental work in the area of work measurement tends to substantiate this view. He found that there exists a consistency in skilled industrial workers in terms of their performances in achieving overall stability in their work output. This overall stability is, however, achieved by the individual worker by the use of "non-systematic purposive behavior" in coping with the environment. For example, on a motion level a fumble occurring during a productive motion of one hand will induce both hands to make an accomodating change in motion. On a higher level is the use of informal rest periods, smoking, etc., which compensate for the vagaries of the human mechanism to achieve an overall stable output on a daily basis. He goes on to suggest that these findings negate the classic notion of "one best way" theory of work, as optimum stable overall output can only be achieved by non stable individual work patterns. The denial of the "one best way" e.g., a predetermined pattern of specific motions appropriate to each task is of great interest. The "one best way" concept is analogous to the "simple amplifier" approach to continuous, closed loop manual tracking and its denial parallel to the recognition of man as an adaptive optimalizing servomechanism as described in McRuer, Ashkenas and Krendel (1959).

In industrial work measurement where the system output is simply units per day, individual behavior and system output can readily be measured independently. For example, conventional time study or the more elaborate motion picture techniques can be used to secure precise data on the individual worker's motion pattern while unit output measurement over relatively short time intervals can measure the statistical stability in internal work rate. In problems of aircraft control this is not the case; here operator behavior, aircraft system, and measuring equipment are often hopelessly confounded. We must approach the problem of behavioral as opposed to systems measurement in somewhat different fashion. In the subsequent chapters we shall present evidence for our hypothesis that consistent overall system output is in the main derived from two attributes:

a) Ability to adapt behavior to meet the changing requirements imposed by changes in the controlled element, the environment, and the operator himself.

b) Economy of effort so as to permit sustained performance.

If our hypothesis is borne out, then the direct measurement of these attributes will provide sensitive measures of performance of value to Fleet activities.

CHAPTER IV

ADAPTABILITY

Introduction

In the previous chapter we discussed the trained human operator's faculty for maintaining the system's output of a complex man-machine system relatively invariant. The human operator demonstrates this behavior despite changes in the characteristics of the system and despite changes or even deteriorations in his abilities. It would appear, therefore, that the human's ability to sense changes in his control behavior or in the dynamics of the system as well as his ability to initiate appropriate compensatory behavior are fundamental attributes of skilled human behavior. In this chapter, we discuss conditions when such adaptive behavior is essential for system functioning, and the criteria and methods for assessing this behavior. We also present a quantitative model for the adaptive human pilot.

Adaptability to Control Dynamics

Adaptation, which in our usage is the process by which the human operator compensates for changes in the requirements of his task, has long been recognized as an important attribute of human motor behavior without explicit formulation. For example, Bartlett (1948), has stated that all forms of skilled performances possess an "outstanding character of rapid adaptation". Hé cites experimental work on the "sense" or direction of movement in the manipulation of controls as follows:

> "It is possible to alter the sense of the manipulatory movement during the course of the experiment without the knowledge of the operator. Nearly everybody now makes a greater amount of error. But the increase is significantly greater in some cases than in others. Some operators make the display the 'key', and are delicately perceptive of any change in visually presented motion. These are readily adaptable, changing an upward into a downward bodily action very often without at all

> definitely knowing what they have done. Others make their 'key' the feel of the movement itself, and for these the required adaptation is relatively difficult."

The need for rapid adaptation by aircraft pilots can be understood in terms of capabilities and tactics of modern high performance aircraft which result in large changes in controlled element dynamics. The static and dynamic response characteristics of aircraft control systems are in general functions of time, mass loading, and flight conditions. The extension of the operating flight regime of modern military aircraft has resulted in ever wider variations in these characteristics. A number of examples from operational or combat conditions provide illustrations of these changes in dynamics. Thus:

1. The release of bombs will shift the center of gravity of the airplane and thus change controlled element dynamics.
2. A jet penetration requires the airplane to pass through atmosphere of rapidly changing density with changes, albeit no radical changes, in the aircrafts' dynamics.
3. Combat damage to control surfaces, structure, hydraulic system and so forth will change the aircraft's response in an unpredictable fashion.

In some cases it is possible to minimize these changing characteristics mechanically by the use of high gain feed back loops. Whitaker, Yamron, and Kezer (1958), however, point out that in an increasing number of cases, the required aerodynamic control moments and power are not available due to the limits imposed by other aspects of aircraft design. In other cases, it is possible to compensate for these changing characteristics by preprogramming adjustments in the control system parameters. If sufficient information about the airplane's flight characteristics is available these adjustments can be scheduled to take place based on measured flight condition; for example, when altitude or Mach number reach a certain value, the gain of a servo amplifier could be increased or decreased by a predetermined amount. Rath (1958) assesses the difficulty of such open loop adjustments and points out that if the scheduled adjustment is not precisely correct due to poor data, miscalculation, some unforeseen change in aircraft configuration, or even component aging, the adjustment might well lead to a poorer system response. He, therefore, presents the case for a closed loop type of adjustment which would continue adjusting until the desired response was reached.

The Adaptive Optimizing Human

The adaptive capabilities of human operators has been put on a quantitative basis by a recent study by McRuer and Krendel (1957) which reviewed the body of data on manual closed loop control. By examining the measured linearized transfer functions which describe human control behavior under various experimental conditions, the authors were able to detail and delimit human adaptive, optimizing behavior. The essential results are that there exist a limited number of transfer function forms which the human pilot can generate. The selection of the form is adapted to the particular control dynamics with which the pilot is confronted. Thus, on an intuitive level, it is clear that a pilot adapts his control behavior differentially to a sensitive high performance aircraft from his control behavior in flying a sluggish aircraft. Within a particular form which the pilot has adapted, he may optimize his behavior by effectively varying the parameters which describe the transfer function he has generated. Thus, although many pilots can fly high performance aircraft, some fly better than others. These superior pilots have optimized their behavior within the following constraint: fly a high performance aircraft.

We can express the human operator's adaptive capability in continuous closed loop control tasks quantitatively by stating the limits within which this adaptation may occur. This form was measured with the human operators subject to random visual inputs and it is consistent with all of the data available.

This linear adaptable model of the human operator for random appearing visual inputs is, in equation form:

$$Y_P = \frac{K_p e^{-\tau j\omega}(\alpha T_I j\omega + 1)}{(T_I j\omega + 1)(T_N j\omega + 1)} K_T \frac{a_T}{\sigma_T} \qquad \text{(IV-1)}$$

where

Reaction time delay, τ, is: $0.12 < \tau < 0.20$ second.

Neuromuscular lag, T_N, is partially adjustable for task.

Equalization, $\frac{(\alpha T_I j\omega + 1)}{(T_I j\omega + 1)} = \frac{(T_L j\omega + 1)}{(T_I j\omega + 1)}$, is adjustable with forcing function and controlled element.

Gain, K_p, is adjustable for over-all system stability and low frequency performance.

Indifference threshold, $K_T \frac{a_T}{\sigma_T}$, is a minor effect with adjustment and values not known.

Within the limitations of the above form, the operator selects a transfer function (lag-lead, lead-lag, pure lead, pure lag, or pure gain) which is adapted to the controlled-element system response in the presence of the inputs to the system. The transfer function form adapted is one consistent with stability and good low frequency control of the over-all system. The constants are adjusted; i.e., the optimization process, to some performance criterion akin to that of the rms minimization criterion of servo theory, and consistent with phase margins from 60 to 100 degrees. The transfer function adapted and optimized for a given task is very similar to the one that a servo engineer would select if he were given dynamics to control together with a "black box", having within it elements making up the describing function given by Eq. (IV-1) and knobs on the outside for adjusting the parameters.

As the authors point out, the precise criterion for the optimization cannot be stated explicitly. It may be that this elusive criterion is a combination of good over-all system performance defined by a minimized rms error and a least effort principle of pilot behavior.

Criteria and Measurements

Granted the individual human operator's performance criterion may defy an explicit statement at this time, if we are to evaluate human adaptive performance, we must select a criterion or criteria. The major desirable attributes for criteria discussed in chapter 2 are still applicable, but adaptation imposes an added requirement. This follows from the process of adaptation and optimizing in which the human may make false starts before he begins to converge on an adaptive optimized dynamic response. The path of convergence is important in this process. It is undesirable for the adapting and optimizing pilot to pass through unstable states which may result in structural damage to his vehicle. The time taken to converge on a proper solution is also important, for should the process take the pilot-aircraft system through an unstable region, then it is critical that the time in this dangerous state be kept as small as possible to minimize the risk of damaging or destroying the vehicle. Solution time must be compared with the criteria time intervals needed to accomplish the immediate tactical problem confronting a pilot; for example, in a dive bombing exercise the pilot must adapt sufficiently fast so that the mission can be carried out at all.

Before the adaptive optimizing capability of the human pilot can be used as the basis for a measure of the "natural" pilot, the concept must be given an empirical foundation. Without an experimental program, the foregoing development lacks contact with the real problem.

Experimentation to sharpen and refine these techniques can be conducted in a simplified part-task flight trainer or in an operational flight trainer. The face validity inherent in the latter, however, would limit the range of conditions under which adaptation could be studied.

For preliminary research, a simple continuous closed loop tracking system, such as Figure 10 presents, would be advisable.

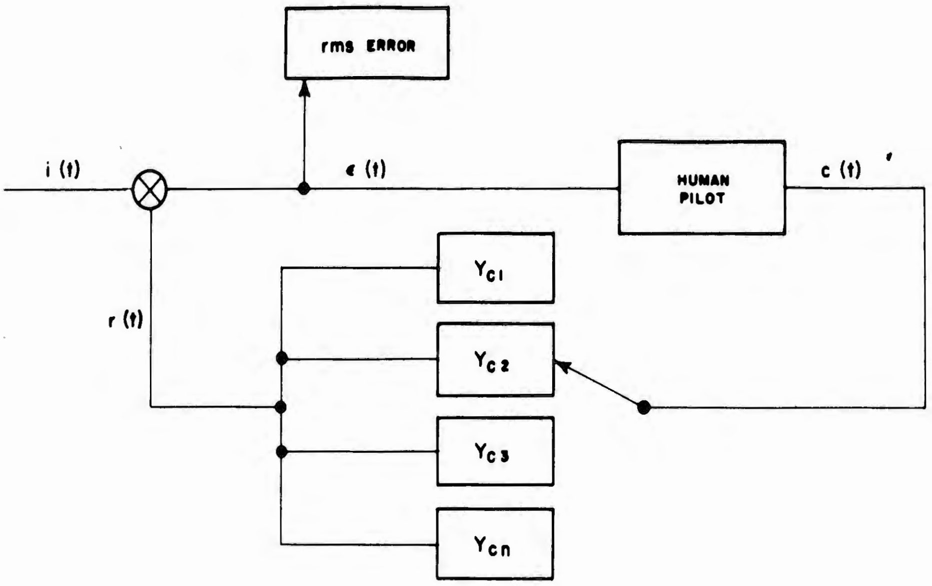


FIG. 10. SCHEMATIC FOR "TIME TO ADAPT" MEASUREMENT

In the illustrated experimental set-up the criterion variable would be rms error and the independent variable would be the controlled element, Y_{ci}, with which the subject was confronted. We have illustrated a situation in which the controlled elements were functions which changed in a discrete fashion.

Stripped to bare essentials, an experimental study of human adaptation to new control dynamics involves the following:

1. The selection of an appropriate random appearing input and the training of experimental groups on the several controlled elements available to a stable level of performance as defined by rms error.
2. The experimental switching without alerting the subject from one controlled element to another and measuring the time required by the subject to achieve the criterion level of performance.

The use of continuously variable controlled element dynamics which would be possible, subject to an established definition of the variables, could be accomplished with an analog computer. In such a case the criterion performance would have to be measured and defined for predetermined controlled element functions. The ranges of possible controlled element variations are considerable. In the simplest example we could reverse the sense of the controls, as in the Bartlett example mentioned at the beginning of this chapter; i.e., $Y_{c1} = -Y_{c2}$. In more complicated cases, changes in dynamics might be so subtle as to tax the pilot's power to detect and to respond to such changes.

CHAPTER V

ECONOMY OF EFFORT

Introduction

In the foregoing chapters we have developed some of the characteristics of the "natural" pilot. First, we specified his ability to maintain the output of the system which he was controlling in a state of statistical equilibrium. Second, we hypothesized a form of behavior - adaptive and optimizing - which would describe the process by which system output was maintained relatively constant despite changes in either the controlled element characteristics or in the characteristics of the pilot himself. By measuring a pilot's capacity to generate adaptive behavior we obtain a sensitive proximal criterion of flight proficiency, as well as a productive insight into the learning process.

Another insight into the learning process may be obtained by examining the criteria which guide the path of a pilot's progress toward greater skill. In this chapter we shall discuss one such guideline - economy of effort.

Minimizing Effort

The concept of "effort" plays a crucial role in many phases of industrial engineering and applied psychology. Effort is a concept in which psychological and physiological energy expenditure are contaminated. Depending on the type of task, one or the other type of energy predominates. Time study standards, for example, are based on the notion of equal pay for equal "effort" (assuming equal training and experience), while methods engineering seeks as its goal the reduction of work "effort". Despite its widespread use the concept of "effort" does not appear suspectible of operational definition and objective measurement.

As a first approximation, it might appear useful to attempt to measure "effort" in terms of the energy expenditure of the body. This approach to what may be termed the physiological cost of work may in general be carried out by measuring oxygen consumption. Oxygen consumption measurement is complicated by the fact that during short intense work periods the body in effect contracts an oxygen debt which is paid back during subsequent rest periods. Oxygen consumption is conveniently measured by an analysis of the subject's exhaled air. Until recently this method, although theoretically quite feasible, required that the subject breathe into a large rubber bag which not only encumbered his movements but severly limited the possible duration of the task. Improvements in the technique are noted by Greene, Morris and Wiebbers (1959) who report on the use of now lightweight gasometors and improved techniques for gas analysis.

Another approach to the measurement of energy expenditure involves the use of rather sophisticated techniques to measure the actual work output of the subject. For example, the Lauru "effort detector" to detect and record forces exerted by a worker occupying a platform have been employed at the Purdue Farm Cardiac Project and the E. I. Dupont de Nemours & Company Haskill Laboratory, among others. [Greene and Morris (1958)]

Unfortunately, metabolic measures, although of obvious interest, are clearly not the entire answer. [Ryan (1947)] For piloting, although requiring very little energy in the metabolic sense, may nevertheless demand an extremely high level of psychological "effort". Since both sedentary and muscular work require the functioning of many parts of the organism, it would be reasonable to expect that all "effort", or stress be it physiological or psychological, is accompanied by measurable physiological changes. To this end a tremendous variety of measures have been employed with varying degrees of reported success. Among them we may list changes in blood chemistry, urine both volume and character, saliva, blood pressure, vascular state of the skin, muscle tonus or tension and the like. Cohen and Silverman (1957) report on a program to measure the in-flight EEG of pilots in an effort to measure pilot "mental effort", and Eason (1959) has obtained some very promising correlations between electro myographic output integrated over ten second intervals and the subject's level of aspiration in a task.

The foregoing techniques for measuring the poorly defined concept of "effort" are not presently in such a stage of development where they can be applied to the operational problems which confront Naval Aviators because of the extent to which the measuring process obtrudes upon the task. In less demanding tasks than flying, however, many of the foregoing techniques have been used to measure human effort. In particular, we are interested in such effort measurements which illuminate the intuitively satisfying concept of a minimum energy mode of behavior as a concomitant of skill. Thus, "easy does it" has long been the characteristic of both the skilled craftsman and the skilled athlete, but we would like some quantitative evidence for this notion.

Studies of human efficiency in an energetic sense have not been in vogue recently because of the prevalence of mechanical power amplifying devices which humans now operate. In the early part of this century, however, when manual power was less likely to be amplified and scientific work assessment was developing, there was considerable interest in the efficiency of human labor. Amar (1920), for example, presents measurements he made on the work effectiveness of a skilled mechanic and an apprentice in the use of a file. On a daily basis, Amar demonstrated that the output of the apprentice was 15% less and his expenditure of energy 65% greater than the corresponding measures for the skilled mechanic. This finding was repeated under many other conditions. In the case of athletic endeavors, Karpovich and Millman (1944), found on comparing expert swimmers with novices that the novices expended two to five times as much energy as the experts for the same stroke and speed.

More recently, Crossman (1959) discusses the development of skill and suggests the action of a selective process which tends to strengthen quicker modes of performance in relation to slower ones. He notes that although he employed a "time" selection criterion in setting up his model, it is not easy to see how the psychological and neural mechanism of the body could produce this result. He speculates further that:

> "Instead of time, the selective variable could be work for if the operator is exerting a constant level of effort as set by the conditions of motivation, the work done in a 'method' would be proportional to the time taken (though of course a particular 'method' can be done faster by exerting more effort). Since saving work is presumably of more biological advantage than saving time, it would be a more likely choice for such a general mechanism."

Seashore (1951) summarizes the work in the area of efficiency as follows:

> "The empirical evidence seems to be that skilled workers in fine manual skills rarely need to press themselves to their physiological limits and that instead they attain their superior results by hitting upon qualitative patterns of action or work methods that make the work easier. Even a person with low physiological limits can surpass one with higher physiological limits if the latter is not using an equally effective work method."

It is reasonable to suggest that "minimum energy" or "least effort" is an important parameter of the skilled pilots behavioral process even though physical effort is not dominant in the pilot's task. Indeed, Birmingham and Taylor (1954) have employed the notion of "doing least" as a fundamental design criteria for the human operator in a machine system where physical effort was secondary. Their design philosophy is best expressed by the following quotation:

> "He is best when doing least. It becomes, therefore, a fundamental assumption...that the more complex the human task, the less precise and the more variable becomes the man. It is assumed that, within limits, the higher the number of integrations and/or differentiations required of the man the poorer will he perform. Conversely, it is hypothesized that the more

> the human operator is freed from the tasks of integration and differentiation, the more regular and precise will become the human output. Human control behavior, it is asserted, reaches the optimum when the man becomes the analogue of a simple amplifier (with pure time delay) as shown in the following equation:
>
> $$\theta_o(t + \tau) = K\theta_i(t)$$
>
> where t represents a value in time, and τ equals the human reaction time."

This philosophy of the simple amplifier is open to criticism in that it denies the human free run of the adaptive process, the importance of which we have previously stressed.

Our problem, in distinction to Birmingham and Taylor, is not one of system synthesis, but of pilot performance assessment. Here we find that an internal criterion such as "least effort" is particularly appropriate for it permits us to distinguish between achievement and the means of achievement. True, we must first measure achievement - number of hits, rms error in tracking and the like. Such measures may, however, as Bartlett (1948) suggests:

> "...throw no light on the means of achievement. Since we must not only know how well a thing is done but within what limits of success it can continue to be done, a measure of immediate achievement is not enough."

This emphasis on process as well as on results is echoed by other workers. Cooper (1957), in his review of a number of aircraft tracking tasks using the scope-presentation type of fire control display, concludes that:

> "...the end result, such as aim wander, is the same in an obviously poor system as in the best system. The thing that distinguishes the two systems is as yet unmeasureable, intangible 'pilot work' or concentration necessary to attain this level of performance."

Just as a pilot's psychological as well as physical effort, as measured by test pilot opinion, provides high resolution measure to distinguish between aircraft systems, similarly, pilot effort may be used to yield a high resolution test of human pilot performance. For under conditions of combat stress, sub-system failure and battle damage, these elusive differences in the underlying skill of the pilot may determine mission success or failure.

Similarly, Duddy (1956) distinguishes between means and ends and presents an example in which pilot opinion of a new yaw damper was favorable, but performance measures showed no difference due to the presence of the new yaw damper. In the Appendix we present a qualified argument relating pilot opinion to the transfer function which the pilot must generate in a given aircraft. Duddy, together with others, states the importance of the transfer function as an index of the process by which performance is achieved:

> "It is well-known that the transfer function of an operator, in this case the pilot, is not a fixed quantity but one that varies with the control system used and the task performed. It has also been said that the human being works best when he does least. This means that he produces the most accurate and consistent results when his transfer function is of the simplest kind. Therefore, the goodness of an aircraft control system for a set task might be measured by measuring the pilot's transfer function when satisfactorily performing that task. The best control system would be the one which demands the simplest transfer function from the pilot."

In the next section of this chapter we will relate transfer function form to skill level and pilot's effort.

A Servomechanism Approach to Skill Development

The argument has been made that pilot opinion is a sensitive measure of effort; that pilot opinion is correlated with the transfer function generated by the pilot; and finally, that increasing skill is characterized by relatively decreasing pilot effort. In this section we will present a model for skill development which predicts certain changes in the pilot's transfer function. These changes imply certain experiments which we will outline in the final section of this chapter.

Krendel and McRuer (1960) present a theory of skill development which in essence is a perceptual approach to skills in the sense recently outlined by Annett and Kay (1956, 1957). The perceptual process is defined broadly so as to include stimuli from a neutral external environment, the human operator himself, and a relevant external environment. The stimuli from the neutral external environment appear as unalterable signals on an external display which also includes signals generated by the operator's manipulation of his environment. Response associated stimuli from the musculature and nervous system are also included in this definition of perception, although these are not indicated as such on the external display. The context in which the theory has been developed is that of manual tracking with controlled element dynamics which are a constant. This is the process of minimizing visually perceived errors by exercising continuous control so as to match visually presented input and output signals with a tracking device. In this situation response learning in the sense described by Bartlett in Chapter 4 is minimized, and we can logically focus on the perceptual aspects of the problem. Goldman's (1959) remarks provide a clear summary of the position:

> "The task facing the subject in compensatory tracking is one example of tasks often investigated as perceptual-motor skills. In almost every instance, as in the task used in this experiment, the adult subject has long possessed the motor repertoire which is employed. In truth, the problem to be solved by the subject and the aspect of the task to be considered by the investigator is the perceptual aspect."

The foundation of the theory rests on the trained pilot's ability to take advantage of the coherence and predictability of his input signal. The central argument of the theory is that an external organization of the stimuli which confront the pilot has a parallel in the internal organization of the pilot's perceptual field. This internal organization is then *effectively* the same as the external organization.

There are three clearly separable external organizations of tracking displays: Compensatory, Pursuit, and Precognitive. We shall define them in turn.

1. Compensatory. In compensatory tracking the visually displayed efforts of the operator's actions are not distinguishable from the system input. The display is the system forcing function minus the modified control response and the operator can only detect the effects of his control motion under zero input conditions. Figure (11) illustrates the general nature of a compensatory display. The operator's task is to minimize the error signal by keeping the follower superimposed on the stationary dot. He is presented, however, only with an indicator showing the difference or error $\varepsilon(t)$ between the forcing function $i(t)$ and the system output $r(t)$. Conventional aircraft instruments are of the compensatory type.

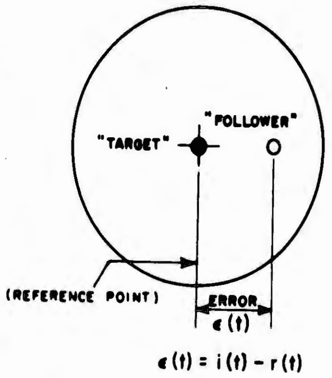


FIG. 11. PURE COMPENSATORY DISPLAY

2. Pursuit. In pursuit tracking the operator's corrective responses are displayed separately from his inputs so that he can distinguish between forcing function and system response. Figure (12) provides an illustration of such a display. Here again the operator must minimize the error between the circle and the dot, but here he can distinguish his own movements from those of the targets.

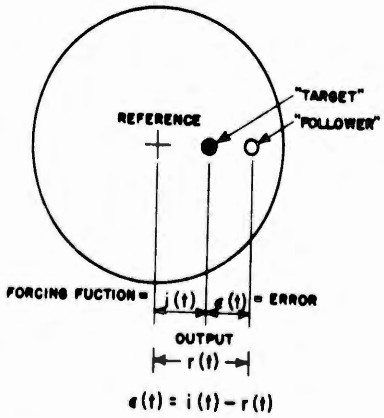


FIG. 12. PURE PURSUIT DISPLAY

3. Precognitive. Precognitive behavior is essentially open loop control and is thus actually not tracking. Here the operator does not maintain continuous control based on perceived error. Instead he employs his complete information about the inputs future to trigger off a reportory of practiced, properly sequenced response. Throwing

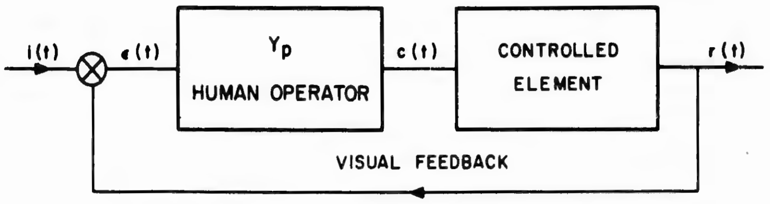
a dart or VFR flight in steady air are examples of precognitive behavior. One might characterize precognitive behavior as discrete control with long sampling intervals, such that open loop control is in effect during these periods.

Figures 13-A and 13-B which are functional block diagrams illustrate how a pursuit display enhances the pilot's ability to predict. This enhancement occurs because the pilot has three transfer functions: Y_{pr} Y_{pi} and $Y_{p\varepsilon}$ at his disposal in the case of Figure 13-B instead of just Y_p as in the case of Figure 13-A. Having more variables to manipulate, the pilot can shape his response so as to meet the requirements of his task more precisely.

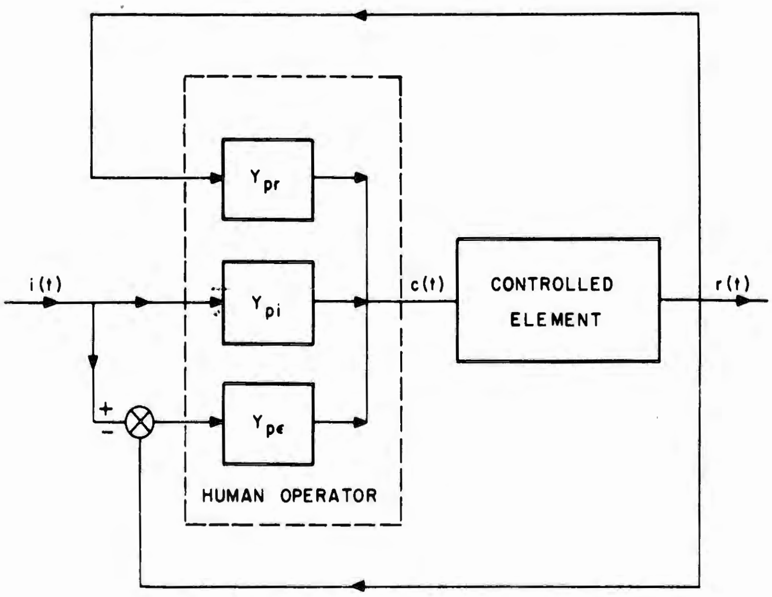
The development of the perceptual display may be represented by Figures 13-A, B and C, which can be interpreted as effective successive organizations of perception of a continuous closed loop control task. Starting with a physical compensatory display, the initial effective and physical form are shown in Figure 13-A. As the pilot or operator learns to extract regularities, he can make short-term predictions. This ability is operationally the same as tracking from the pursuit display Figure 13-B. Finally, when the pilot becomes highly skilled by dint of much training, he can take advantage of all of the internal coherence in the input signal and operate effectively as if he had an internal program generator as shown in Figure 13-C, which describes precognitive behavior.

Clearly, not all situations allow the operator to progress to a perception of the problem illustrated in Figure 13-C. The input signals must be deterministic, the pilot highly experienced, and the controlled element dynamics well-known and constant.

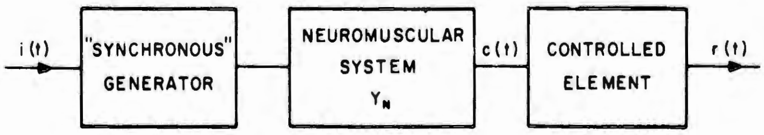
The effective open loop response pattern characteristic of the highly skilled pilot is not without its risks for a skilled pilot. McRuer and Ashkenas (1959) in a discussion of possible causes for pilot induced oscillations, P.I.O., present a case where a skilled pilot may



(A) COMPENSATORY SITUATION REPRESENTING THE INITIAL PHASE



(B) PURSUIT SITUATION REPRESENTING THE SECOND PHASE



(C) OPEN OUTER LOOP REPRESENTING THE SYNCHRONOUS PHASE

FIG. 13. SUCCESSIVE ORGANIZATIONS OF PERCEPTION

Krendel and McRuer (1960)

achieve precognitive functioning, but to the wrong stimulus! A pilot who has achieved the precognitive level of skill can track a sine wave input almost perfectly with little or no phase lag. Consider a case where the controlled element dynamics are such that when the system is excited part of the output is nearly sinusoidal, as would occur for a controlled element:

$$Y_c = \frac{K}{s^2 + \omega^2}$$

The skilled pilot runs the risk of confusing this sine wave output with a sine wave input to the system. If he proceeds to track this supposed input, he is in grave trouble because of the - 180° phase shift defined by the action of the differential in the compensatory tracking device. Thus until the pilot discovers the true source of signal, he will do exactly the wrong thing! If he manages to determine the origin of his error signals, he will reverse his output by 180° and live to fly again.

Despite the foregoing caveat, it is clear that achieving the precognitive level of skill means that the pilot is expending less effort because he need not maintain constant vigilance over all aspects of the problem; (he has learned what regularities in his display are important). His control outputs are deft ballistic responses more or less widely spaced in time.

Time in Control

The fraction of the total time available for the performance of a given control task has a long and honored place in the history and current practice of work load evaluation. In the simplest terms, the problem to which these techniques address themselves is: does the operator, skilled mechanic, or pilot have enough time to make the decisions and perform the actions which are required at the level of precision demanded by the task? Whether this approach is simply called a time line analysis, or is refined somewhat to become a second by second operational analysis, or is further elaborated by information theory consideration, or is mechanically performed on a digital computer, the essence remains the same. The underlying conditions of process stationarity, as discussed in Chapter 3, must be met before this or any other process measure may be applied. Furthermore, the segments into which the task must be analyzed before time measures can be taken must allow these time measures to be additive in order that the work load analysis have adequate meaning. (Abruzzi (1952))

In the context of our immediate problem, instrument flight doctrine, Anon. (1957) provides a direct contact with a sampling or time line analysis approach to work load, and by extension, to pilot skill. This flight doctrine requires that the pilot repeatedly scan his instrument panel and fixate on individual instruments for relatively short intervals in order to obtain the information required for his control actions. Thus, if task requirements could be overlaid in a quantitative fashion on eye fixation patterns, as suggested by Cole, Milton and McIntosh (1954), work load estimates could be synthesized. Such an effort has in fact begun and is being refined. (Lindquist and Gross (1958))

Our approach, based on the previous three chapters, argues that from <u>either</u> a work load or effort viewpoint, <u>or</u> from a servomechanism approach to the development of motor skills, the time that the pilot requires to control his problem is a sensitive measure of his level of skill. As the pilot develops successive organization of perception, he is capable of maintaining more and more "loose" a control because of his increasingly effective predictions. Time in control is a measure of the degree to which the subject has extracted the internal coherence of the input signal. For, during the time intervals when the pilot is not attending the display, he is clearly operating in precognitive open loop fashion.

In our pilot aircraft system situation, one might proceed to develop a skill level assessment technique in accordance with the schematic presented in Figure 14.

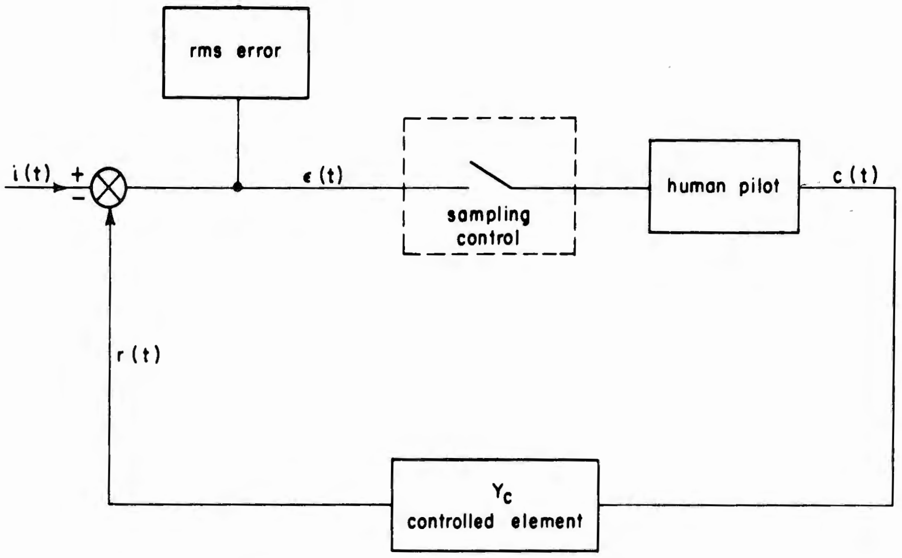


FIG. 14. SCHEMATIC FOR "TIME IN CONTROL" MEASUREMENT

The pilot would be isolated in an aircraft simulator-like device and assigned a problem; for example, a tail chase at Mach 1 at an altitude of 30,000 feet. The problem input to the pilot would be through a scope type presentation simulating a radar gun sight. The controlled element dynamics would include the appropriate aircraft dynamics and the geometrical relationships between the target and the tracking aircraft. A forcing function would be generated to simulate atmospheric turbulence or target evasive action, or both. It would be experienced by the pilot as a target disturbance from the scope fiduciary mark. The normative (satisfactory from a tactical standpoint) rms error would be established using a continuous presentation. With this performance as a base line, and using an "on" time which will be in the neighborhood of .25 seconds, the scope would be blanked out using increasingly long periods of "off" time so as to reduce the percentage of viewing time or "time in control". A pilot's score would than be the minimum "time in control" at which he is able to maintain a satisfactory rms tracking error. Similar experiments can, of course, be devised for other mission phases and partial to full instrument panel rather than scope presentation can be employed. In the latter case, "time in control" will be implemented by systematically blanking out the instrument lights.

The use of an aircraft simulator should not be regarded a hard and fast requirement since much simpler synthetic training and evaluation devices may be constructed to either train or evaluate flying personnel. The essential requirements for such devices are input functions of controllable internal coherence (e.g., a variation from completely deterministic sine wave inputs through a series of decreasingly sinusoidal appearing inputs as would result from passing a broad band noise through increasingly broad band pass filters); and a capacity for variable "on" and "off" interval blanking of schedules. There is a fruitful potential for the Navy in investigating such synthetic devices since they imply the possibility of developing less complicated training and evaluation devices than the present operational flight trainers. There are many advantages in such an eventuality:

a. Increased utilization time because of decreased system complexity and a consequent decrease in maintenance requirements.

b. Greatly decreased cost.

c. Greater availability of equipment.

BIBLIOGRAPHIC FOOTNOTES

Abelson, R.P. (1953). *Spectral Analysis and the Study of Individual Differences in the Performance of Routine Repetitive Tasks*, Educational Testing Service Research Bulletin 53-2.

Abruzzi, A. (1952). *Work Measurement*, New York, Columbia University Press

Abruzzi, A. (1956). *Work, Workers and Work Measurement*, New York, Columbia University Press.

Amar, J. (1920). *The Human Motor*, London, Routledge and Sons.

Anderson, R.O. (1958). *Design for Space - Vehicle Control System Reliability*, WADC TN 58-156.

Annett, J., and Kay, H. (1956). Skilled Performance, *Occup. Psychol. 30*, 112-7.

Annett, J., and Kay, H. (1957). Knowledge of Results and Skilled Performance, *Occup. Psychol., 31*, 69-79.

Anon. (1954). *Military Specification, Flying Qualities of Piloted Airplanes*, Mil - F - 8785 (ASG).

Anon. (1957). *All-Weather Flight Manual*, Office of The Chief of Naval Operations NAVAER 00-80T-60.

Bartlett, F.C. (1948). The Measurement of Human Skill, *Occup. Psychol.* *22*, 83-93.

Ben-Avi, H. (1947). Studies of Subjective Measures of Flying Proficiency in *Psychological Research on Pilot Training*, (Miller, N.E., ed) Washington, D.C. U.S. Government Printing Office

Benepe, O.J. (1954). Narasimhan, R., Ellson, D.G., *An Experiment Evaluation of Harmonic Analysis*, WADC TR 53-384.

Birmingham, H.P. and Taylor, F.V. (1954). *A Human Engineering Approach to the Design of Man Operated Continuous Control Systems*, NRL Report 4333.

Bray, C.W. (1958). AFPTRC Studies of Pilot Proficiency Measurement in *Aviation Psychology in Western-Europe and a Report on Studies of Pilot Proficiency Measurement*, Amsterdam, Publication No. 1 of the Western European Associations for Aviation Psychology.

Chalk, C.R. (1958). *Additional Flight Evaluations of Various Longitudinal Handling Qualities in a Variable Stability Jet Fighter*, WADC TR 55-229 Part I and 55-299 Part II.

Chapanis, A. (1959). *Research Techniques in Human Engineering*, Baltimore, The Johns Hopkins Press.

Cohen, S.I. and Silverman, A.J. (1957). *The Measurement of Pilot Mental Effort*, AGARD Report 148.

BIBLIOGRAPHIC FOOTNOTES (Continued)

Cole, E.L., Milton, J.L., McIntosh, B.B. (1954). Routine Maneuvers Under Day and Night Conditions Using an Experimental Panel Arrangement, WADC Technical Report 53-220.

Cooper, G.E. (1957). Understanding and Interpreting Pilot Opinion, Aero. Eng. Rev., 56, 47-51.

Crossman, E.R.F.W. (1959). A Theory of the Acquisition of Speed Skill, Ergonomics, 2, 153.

Doyle, G.B. (1958). Pilot Techniques for Stability and Control Testing, Maryland, Naval Air Test Center.

Draper, C.S. (1955). The Forty-third Wilbur Wright Memorial Lecture - Flight Control, J. of The Royal Aero. Society, 59, 451-77.

Duddy, R.R. (1956). The Quantitative Evaluation of an Aircraft Control System, AGARD Report No. 29.

Eason, R.G. (1959). Private communication.

Elkind, J.L. (1956). Characteristics of Simple Manual Control Systems, MIT Lincoln Laboratory Technical Report III.

Elton, P.M. (1922). An Analysis of the Individual Differences in the Output of Silk Weavers, London, Industrial Fatigue Research Board.

Gilruth, R.R. (1943). Requirements for Satisfactory Flying Qualities of Airplanes, NACA Report 755.

Goldman, A. (1959). Sensory Interaction and Response Capacity, NAVTRADEVCEN 789-11-1.

Graham, D. and Lathrop, R.L., (1953). The Synthesis of "Optimum" Transient Response Criteria and Standard Forms, AIEE Trans. 72, Part II.

Grant, E.L. (1956). Statistical Quality Control, New York, McGraw-Hill Book Company.

Greene, J.H. and Morris, W.H.M. (1958). The Force Platform: An Industrial Engineering Tool, The J. of Ind. Eng., 9, 128-132.

Greene, J.H., Morris, W.H.M., and Wiebers, J.E. (1959). A Method for Measuring Physiological Cost of Work, The J. of Ind. Eng., 10, 180-4.

Harper, Robert P., Jr. (1955). Flight Evaluations of Various Longitudinal Handling Qualities in a Variable Stability Jet Fighter. WADC TR 55-299.

Huff, W.W. (1958). Supplementary Notes on Longitudinal and Lateral Directional Stability and Control, Maryland, Naval Air Test Center.

Karpovich, P.V. and Millman, N. (1944). Energy Expenditures in Swimming, Amer. J. Physiol., 142, 140-4.

BIBLIOGRAPHIC FOOTNOTES (Continued)

Krendel, E.S., and McRuer, D.T. (1960). A Servomechanisms Approach to Skill Development, *J. of the Franklin Inst.*, 269, 24-42.

Lindquist, O.H., and Gross, R.L. (1958). *Human Engineering Man-Machine Study of a Weapon System*, Minneapolis-Honeywell Aero Report R-ED 6094.

Loveday, J., and Munro, S.H. (1920). *Preliminary Notes on the Boot and Shoe Industry*, London, Industrial Fatigue Research Board.

McRuer, D.T., and Ashkenas, I.L. (1959). *Pilot Induced Oscillations, Basic Causes*, Systems Technology Inc., TM 27.

McRuer, D.T., and Krendel, E.S. (1957). *Dynamic Response of Human Operators*, WADC TR 56-524.

McRuer, D.T., Ashkenas, I.L., and Krendel, E.S. (1959). A Positive Approach to Man's Role in Space, *Aero/Space Eng.*, 18, No. 8, 30-36.

Newell, F., and Campbell, G. (1954). *Flight Evaluations of Variable Short Period and Phugoid Characteristics in a B-26*, WADC TR 54-594.

Philips, W.H. (1948). *Appreciation and Prediction of Flying Qualities*, NACA TN 1670.

Priebe, V.J. (1957). *Directional Stability and Control*, Maryland, Naval Air Test Center.

Rath, R.R. (1958). *Summary and Status of Adaptive Control System Program*, WADC TN 58-330.

Ryan, T.A. (1947). *Work and Effort - The Psychology of Work*, New York, The Ronald Press Company.

Seashore, R.H. (1951). Work and Motor Performance in *Handbook of Experimental Psychology*, (S.S. Stevens, Ed.) New York, John Wiley and Sons.

Shaw, W.J. (1958). Objective Measurement of Driving Skill, *Traffic Safety*, *2*, No. 4, 13-16.

Simon, L.E. (1941). *An Engineers' Manual of Statistical Methods*, New York, John Wiley and Sons, Inc.

Webb, W.B. (1958). Measurement of Learning in Extensive Training Programs in *The Measurement of Learning*, Dept. of Psychology, Washington University.

Westbrook, C.B., and McRuer, D.T. (1959). Handling Qualities and Pilot Dynamics, *Aero/Space Eng.*, 18, No. 5, 26-32.

Whitaker, H.P., Yamron, J., and Kezer, A. (1958). *Design of Model-Reference Adaptive Control Systems for Aircraft*, MIT Instrumentation Laboratory

Whitehead, T.N. (1938). *The Industrial Workers*, Cambridge, Harvard University Press.

Williams, S.B. (1949). Visibility on Radar Scopes in *Human Factors in Underseas Warfare*, Washington, National Research Council.

APPENDIX

THE PILOT AIRCRAFT SYSTEM

Introduction

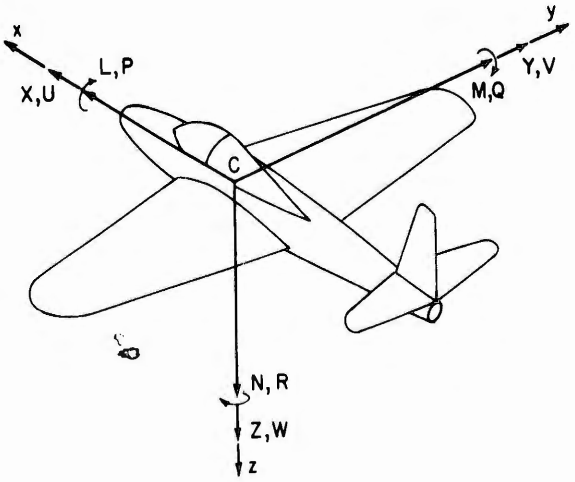
In this appendix we discuss the dynamics of aircraft flight and the pilot's role in the system. We then consider the assessment of the handling qualities of piloted aircraft and its relation to flight proficiency measurement. Finally, we present a recent theoretical advance which has explained the aircraft handling qualities judgments made by test pilots in terms of the dynamics of the aircraft and the resulting requirements imposed on those dynamics which the pilot must generate. In subsequent chapters we shall explore the implications of this finding, and we will develop a technique to assess pilot performance in terms of the extent to which he generates the appropriate control dynamics to optimize the performance of the aircraft-man system.

Flight Control

In order to understand the requirements imposed on a pilot by the aircraft he flies, it is necessary to achieve some familiarity with the underlying principles of aircraft flight.

For the purposes of this discussion an aircraft can be approximated by a rigid body. This rigid body is described in terms of two axis systems. One system--the moving axis system--is fixed to the aircraft with its origin at the center of mass and rotates and translates with the aircraft. The second system is fixed on the earth. Since the rotation of the earth is assumed to be negligible for the purpose of this discussion, the earth reference system is effectively an inertial reference system. The use of the moving axis system provides a convenience in description by enabling us to describe the aircraft's angular velocities in terms of constant moments and products of inertia.

The velocity of the aircraft's center of mass, the origin of the moving axis system, $\bar{v}_c$, is in general inclined with respect to the X axis, or forward facing direction of the aircraft. In Figure 15 we present the aircraft axis system X, Y, Z, and the corresponding velocity vectors U, V, W, for the velocity of the aircraft's center of mass.



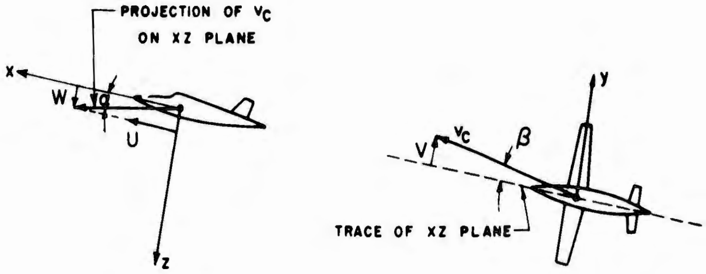
L = rolling moment P = rolling velocity
M = pitching moment Q = pitching velocity
N = yawing moment R = yawing velocity
[X, Y, Z] = components of resultant aerodynamic force
[U, V, W] = component of velocity of C

FIG. 15 AIRCRAFT AXIS SYSTEM

In Figure 16 we illustrate the angular differences between the velocity vector along the flight path and the orientation of the aircraft. These two angles; angle of attack, α, and angle of side slip, β are defined as follows:

$$\alpha = \arctan W/U$$

$$\beta = \arcsin V/\bar{v}_c$$



ANGLE OF ATTACK

ANGLE OF SIDESLIP

FIG. 16. DEFINITION OF ANGLE OF ATTACK AND ANGLE OF SIDESLIP

The flight path of the aircraft is defined by the motion of the center of mass, C, of the aircraft as expressed in the earth reference system. The orientation of the aircraft is described in terms of the earth reference system following a transformation from the moving coordinate system.

The flight path of the airplane is determined by the following physical factors:

1. The inertial characteristics of the aircraft.
2. The earth's gravitational field.
3. The propulsive force of the power plant.
4. The aerodynamic forces and moments.

The aerodynamic forces and moments on the aircraft are functions of its velocity with respect to the airstream, the density of the air, the angle of attack, and the aircraft's configuration.

The limits on the family of flight paths which the aircraft can assume are imposed by the configuration and structural strength of the aircraft and by the propulsive system available. Within the envelope of flight paths of which the aircraft is capable, the pilot exercises control by affecting the aerodynamic forces and moments through the control surfaces of elevator, rudder, and aileron, and by affecting thrust by throttle control of the propulsion system.

The flight control problem which the pilot solves is conventionally divided into two aspects:

1. stability
2. control

Stability is the tendency of the aircraft to maintain a state of dynamic equilibrium. Thus, a stable aircraft which is flying straight with wings level will tend to maintain this condition.

Control is the act of directing aircraft from one arbitrary flight path to another arbitrary flight path. In the process of exerting control transient non-equilibrium stages will be encountered.

The concept of equilibrium is central to an understanding of stability. The equilibrium state of an aircraft in flight is a pure translational motion. This is because the angular orientation of the aircraft about its flight path establishes the aerodynamic forces which are required to balance the aircraft's weight in an equilibrium condition. Stability is a property of the equilibrium state. There are three types of static stability:

1. An equilibrium state is positively stable if when subjected to a slight disturbance the initial state is ultimately reestablished. This situation occurs when the potential energy of a system which is acted on by conservative forces is at a minimum. (Figure 17A)
2. A body which is acted on by conservative forces is statically unstable when the potential energy is at a maximum. In this case a slight disturbance from the equilibrium position results in a continuous decrease in potential energy and a consequent continuous motion. (Figure 17B)
3. A body which is acted on by conservative forces is neutrally stable if its potential energy does not change as a function of its position. A disturbance from neutral equilibrium results in establishing a new position of neutral equilibrium without motion. (Figure 17C)

Dynamic stability is a somewhat more elusive concept for intuitive description. As with static stability it is related to the nature of the motion of the system following an initial disturbance. A detailed examination of dynamic equilibrium requires knowledge of the differential equations describing the system. Although these equations are seldom linear in nature; they may be approximated by linear equations for <u>small</u> disturbances.

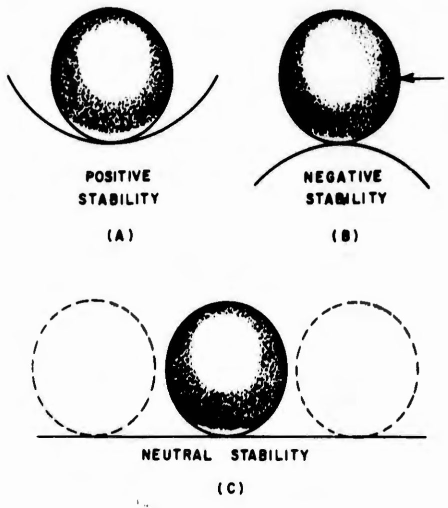


FIG. 17. STATIC STABILITY

1. A dynamic system is stable if an imposed small disturbance ultimately subsides.
2. A dynamic system is unstable if an imposed small disturbance leads to an ever increasing or divergent output. In a practical case, of course, the diverging output will be eventually stopped by a system nonlinearity or physical limitation.

The assessment of the stability properties of a given dynamical system can proceed by any of a number of related analytic techniques for examining the equation determinants (Routh's criterion), and by using servomechanism analysis, (Nyquist plots, Bode plots, Root locus plots).

The two aspects of the flight control problem which we have been discussing have been summarized by Draper (1955) as follows:

1. <u>Stability</u> - a characteristic that implies the continuous existence of an equilibrium orientation and an equilibrium path of motion to which the flying entity returns after disturbances have caused deviations.
2. <u>Control</u> - changes to the equilibrium state by command inputs.

From an overall standpoint the flying system - human, propulsion, plant control system, airframe, etc., must cope with three situations:

1. Stabilization - to keep the actual orientation and velocity acceptably close to the equilibrium orientation and velocity.
2. Navigational Flight Control - to cause the aircraft to follow equilibrium flight programs from takeoff to landing.
3. All-Maneuver Flight Control - to follow arbitrary complex paths to perform operations such as aerial combat, air-to-ground combat, aerial refueling, etc.

The reader should bear in mind that the stability and control in which we are interested is that characterizing a piloted aircraft. Thus a pilot may be required to generate such control dynamics as are necessary to stabilize the flight of an otherwise unstable aircraft. The Naval Aviator must be able to assess the flying qualities (controlled element dynamics) of his aircraft and then produce the appropriate human dynamics which will result in stable, high performance, controlled flight. This is one of the most essential aspects of flying proficiency

Pilot Control

As a very simple example of continuous control by the pilot, consider the following process:

> A pilot attempts to establish a certain rate of climb. He pulls back on the stick and holds it steady. In response to this command, the airplane smoothly takes up a steady climb. However, the climb rate is not precisely what the pilot wants, so he applies a slight forward force to the stick. The airplane then pitches forward, and the rate of climb is reduced. But it is reduced too much, so the pilot applies back pressure again, and again the plane climbs too fast.

In this example the pilot-aircraft combination functions as a continuous closed loop control system. The pilot observes some output quantity of the aircraft, such as its pitch attitude, compares it with the desired quantity and then operates the controls to reduce the error or difference between the actual and the desired attitude. The airplane then changes attitude and the pilot repeats the process of observation, comparison, and control operation until the error is reduced to zero. This servomechanism model of the pilot-airframe attitude-hold system is illustrated in Figure 18.

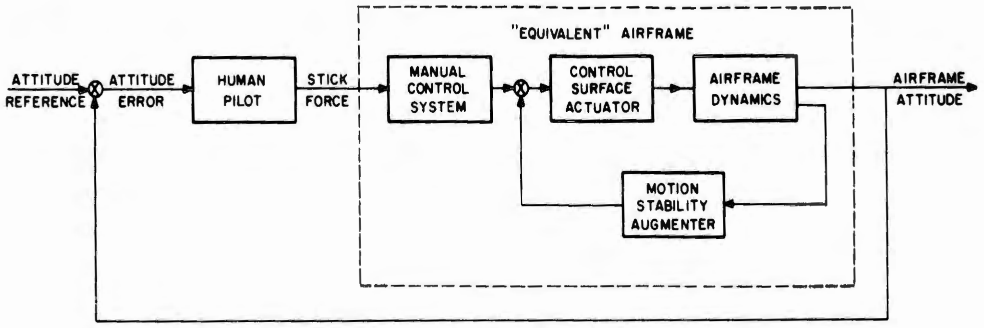


FIG. 18. SERVO MODEL OF PILOT-AIRFRAME ATTITUDE HOLD SYSTEM

Flying Qualities of Piloted Aircraft

In seeking a method for analyzing the performance of the pilot airframe system it is logical to examine the procedures by which new aircraft are tested and approved by the military services. A striking symmetry exists. In aircraft acceptance tests, pilots who have achieved certain normative standards, i.e. test pilots, fly non-standard aircraft to assess the characteristics of the aircraft; while in flight proficiency measurement, pilots who have not achieved normative standards fly standard aircraft, i.e. operational aircraft, and the characteristics of the pilot are determined.

Military specifications for the handling qualities of piloted aircraft began with Signal Corps Specification 486 for the procurement of a heavier-than-air flying machine. This specification stated that "during the trial flight of one hour it must be steered in all directions without difficulty and at all times be under perfect control and equilibrium". In the early forties an Air Force designers' handbook required that "the stability and control characteristics should be satisfactory". Modern concepts of quantitative criteria began with the R.B. Gilruth's (1943) NACA report: "Requirements for Satisfactory Flying Qualities." From this paper and others, such as W.H. Philips (1948) "Appreciation and Prediction of Flying Qualities", grew the Navy Bureau of Aeronautics Specification SR-119 and the Air Force Specification C-1815. In 1954 both of these specifications were superceded by MIL-F-8785 (ASG). This specification seeks to detail a set of requirements for an aircraft-pilot system which will result in stable, controllable flight. The requirements in general must specify both maximum and minimum levels of the various stability parameters; for as Doyle (1958) points out, " ... in stability and control, unlike performance, having too much is frequently as undesirable as having too little". The specification sets forth requirements for longitudinal and lateral stability and control. Thus, requirements for short period longitudinal oscillations, long period longitudinal oscillation (phugoid), lateral directional oscillations, spiral stability, control requirements and the like are covered in some detail. In general, a combination of qualitative and quantitative requirements are set forth, for example: Anon. (1954)

> "Long-Period Oscillations - Although there is no specific requirement for damping of the conventional long-period, or phugoid oscillation which occurs at approximately constant angle of attack, there shall be no objectionable flight characteristics attributable to apparent poor phugoid damping. In addition if the period of a longitudinal oscillation is less than 15 seconds, the oscillation shall be at least neutrally stable."

The following general statement on the qualitative requirements is of some interest:

> "Interpretation of qualitative requirements - In several instances throughout the specification re-

> quirements, qualitative terms, such as 'objectionable flight characteristics', 'unacceptable flight conditions', 'unusual pilot technique', etc., have been employed as a means of permitting latitude where absolute quantitative criteria might be unduly restrictive. Final determination of compliance with requirements so worded will be made by the procuring activity."

The not uncommon contrast between quantitative specifications and pilot opinion is also brought out in the paper by Doyle (1958), which he quoted at an AGARD meeting:

> "Where a conflict of 'numbers' and pilot opinion is involved we normally report the results as meeting the requirements of the specification, but being 'unsatisfactory' or 'unacceptable' depending on the seriousness of the deficiency under consideration. Although this may seem contradictory, this is not as unusual as it may sound. In your past experience you can easily recall cases where an airplane's characteristics appeared satisfactory from a quantitative comparison with accepted standards, yet pilot opinion of the airplane as a weapon was of a low order. Conversely, the F8F was definitely marginal insofar as stability characteristics were concerned, yet the fleet pilots loved the airplane; and the fact that it consistently won the annual gunnery competitions proved that it was an effective weapon."

In the U. S. Navy, stability and control testing is a responsibility of the Naval Air Test Center, Patuxent River, Maryland. What might best be termed "calibrated" test pilots are developed by subjecting experienced fleet pilots to a rigorous training program in the Test Pilot Training Division. In this program, test pilot trainees are taught highly detailed test techniques which are fully documented by Priebe (1957), Doyle (1958) and Huff (1958).

The Correlation of the Quantitative and Qualitative in Aircraft Handling Qualities

Since the early 1950's a program to relate the qualitative and quantitative requirements for aircraft stability and control has been underway under USAF sponsorship. It is well known that servomechanism theory and methods provide a scheme for predicting the characteristics of an autopilot-airframe system. Figure 19 shows a block diagram of an attitude-hold autopilot system. By an analysis of a closed loop system such as this it is possible to determine the controllability (response to inputs) and the stability (oscillation characteristics) of the aircraft. The similarity between this system and the human pilot-aircraft (fig. 18) previously discussed is apparent. The possibility of analyzing the human pilot-aircraft system as a servomechanism immediately comes to mind.

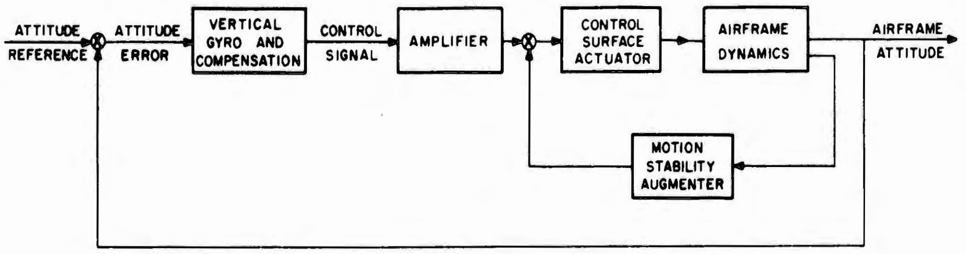


FIG. 19. AUTOPILOT – AIRFRAME ATTITUDE HOLD SYSTEM

The application of this servomechanism model of the pilot aircraft system permits an explanation of the judgments of aircraft handling qualities made by test pilots to be made in terms of the dynamic requirements which these aircraft impose. Analogously, in the reciprocal problem, we plan to assess pilot performance by the extent to which he generates control dynamics which are effective in optimizing the performance of the man-aircraft system.

A summary discussion of this advance in our understanding of handling qualities, together with an historical perspective on handling qualities specifications, has been presented by Westbrook and McRuer (1959).

The implementation of the servomechanism approach requires a detailed knowledge of the dynamic response characteristics of the human pilot in a language which is compatible with linear servomechanism theory. McRuer and Krendel (1957) drawing and elaborating on the work of Tustin, Mayne, Russell, and Elkind, have examined all of the applicable available data on human response characteristics. From these data a consistent usable mathematical representation of the human operator has been obtained. Figure 20 shows a block diagram of a hypothesized linear model of the human operator performing a continuous closed loop task. Further details are presented in Chapter 4.

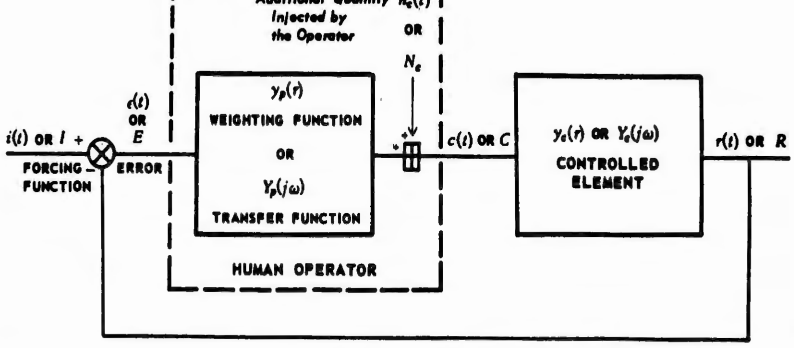


FIG. 20. HYPOTHESIZED LINEAR MODEL OF HUMAN OPERATOR PERFORMING A CONTINUOUS CLOSED LOOP TASK

McRuer and Krendel (1957)

Tracking in the continuous closed loop sense illustrated in Figure 20 exists in an extremely important series of flying tasks. For examples of flight conditions requiring manual tracking consider the following:

1. aspects of the collision type courses flown in rocketry and approach situations.
2. pursuit courses flown in gunnery
3. "constant range" tracking in formation flying
4. "infinite range" tracking which approximates in many control problems such as rough weather flying, horizontal turn entries and so forth.
5. visual landing and take-off where the pilot tracks the runway.
6. instrument flight where the pilot tracks the instrument.

On the other hand, certain of the essentially routinized and preprogrammed tasks which a pilot performs are open loop control functions; e.g., contact flight in smooth air in a stable aircraft.

Extensive flight tests have been conducted and pilot opinion data obtained at the Ames Research Laboratory of the NASA and at Cornell Aeronautical Laboratories for various aircraft dynamics. Newell and Campbell (1954), Harper (1955), Chalk (1958).

Figure 21 presents a so-called iso-opinion plot of **pilot** opinion data obtained with constant stick force and displacement per unit normal acceleration. In an effort to put the expert pilot opinion on a more nearly quantitative scale, Cooper (1957) devised a ten point rating scale by which variable stability aircraft may be rated in terms of "calibrated expert pilots". These rankings of aircraft handling qualities in terms of aircraft dynamics, together with our knowledge of human dynamics provide the basis for rationalization of these pilot opinions of handling qualities.

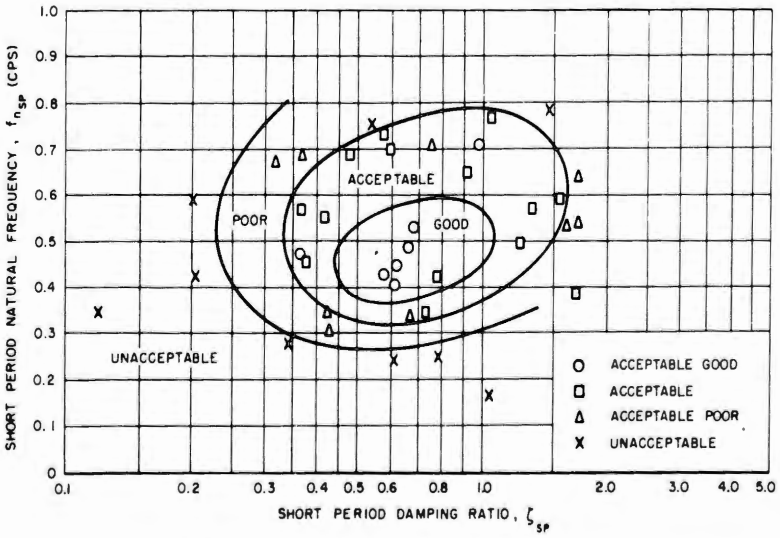


FIG. 21. ISO-OPINION PLOT OF PILOT OPINION DATA

Harper (1955)

In Figure 22 the same iso-opinion data is presented. There are three regions where tracking was important; A, B, C. Using the linear human pilot model (which is discussed in greater detail in Chapter 4), with values inserted for the various parameters it was found that the model appeared to be capable of predicting the boundaries established by pilot opinion data. The human-pilot-airframe combination would have to have an unstable boundary along the dashed line in Figure 22. If the pilot-airframe system became unstable, ratings of unacceptable would be expected. In essence, the rather close agreement between the computed unstable boundary line and the unacceptable iso-opinion line shows that the pilot opinion of an aircraft configuration is closely related to the closed loop performance of the pilot-aircraft system and the dynamic requirements imposed on the pilot.

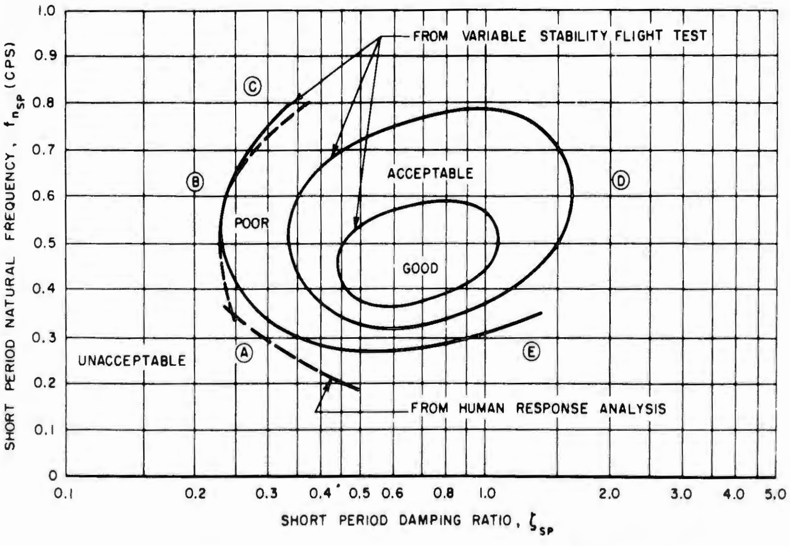


FIG. 22. HUMAN DYNAMICS – HANDLING QUALITY CORRELATION

Westbrook and McRuer (1959)

Westbook and McRuer (1959) state:

> "One can further show analytically that the 'good' region corresponds to conditions where the 'analog pilot' is allowed considerable variation in lead time constant and closed-loop response without having a closed-loop airframe-pilot system instability. In other words, the 'good' region is one where the analog pilot can easily be adjusted to obtain fast, stable time responses. The 'acceptable' and 'poor'

> regions correspond to those conditions where greater restrictions must be placed upon the pilot analog than is the case in the 'good' region to obtain even moderately fast and stable closed-loop response.
>
> While these findings are by no means ironclad, they do point up an exceedingly important concept - quite simply, that pilot opinion of an airframe configuration is closely correlated with closed-loop performance, and hence to some extent with the transfer characteristics and parameters adapted by the pilot to control the configuration. Essentially complete verification of this statement, and consequently evidence of close correlation between the two approaches to handling qualities, has been obtained for the longitudinal short-period case."

These results are of considerable importance to our present program as they are a demonstrated meeting ground between expert observer opinion and truly objective measurements. Our problem is to determine an objective measure of pilot performance which will correlate in an analogous fashion with the subjective ratings of pilot proficiency made by skilled instructors, in current use in fleet activities and in civil aviation.

Indeed, in what is perhaps the critical test of Naval pilot proficiency - the carrier landing - the split second subjective judgment of the skilled LSO dominates the scene. The successful execution of tens of thousands of such landings in day and night, in fair and foul weather provide ample testimony to the reliability and validity of overall composite subjective measures of pilot performance.

The success of subjective judgement has gradually lead to a reaction in flight proficiency research against atomistic assessment. For example, Webb (1958) in discussing U.S. Navy thinking in this area, emphasizes the need for overall compsite subjective evaluation:

> "Learning to be an aviator or a better executive are examples of .... the type of situation that seems to require a subjective type measurement - situations that are quite complex. Although in both these instances quite specific skills may be necessary but the situations in which the skills are applied and the combinations and sequence in which the skills are employed are likely to be quite varied. It is, in the case of the pilot, quite necessary that he learn to make a smooth coordinated turn. However, even more important is that he know when to begin such a turn, where each turn will end up, and how such turns are related to the varying requirements of flying. These latter capabilities cannot be judged from the isolated ability to perform a coordinated turn under standard conditions. Instructors feel that they can tell from observing a student engaged in the general process of learning to fly a given task (e.g., landing) how 'good' a flyer the man is likely to be in the vastly varied types of situations which will be posed by landings. He would prefer this general wholistic evaluation to using part task after part task involved in landing and observing the man's precise deviations from these part tasks."

Bray (1958) comes to a very similar conclusion after summarizing an eight-year U.S.A. Force program in flight proficiency evaluation:

> "Thus the effort to develop job performance measures as criteria for aptitude test and other predictors, or for use in predicting future job performance may well be self-defeating, and we cannot help but wonder whether subjective judgements of pilot capacity, such as those made by peers and instructors, do not somehow analyze out and take into account the lack of control and homogeneity in stimuli and responses and the effect of the system on the response being measured."

This striking parallel between researcher opinion and operational practice is reassuring, for it is the fleet pilot not the aviation psychologist to whom we must address our efforts.

UNCLASSIFIED

UNCLASSIFIED