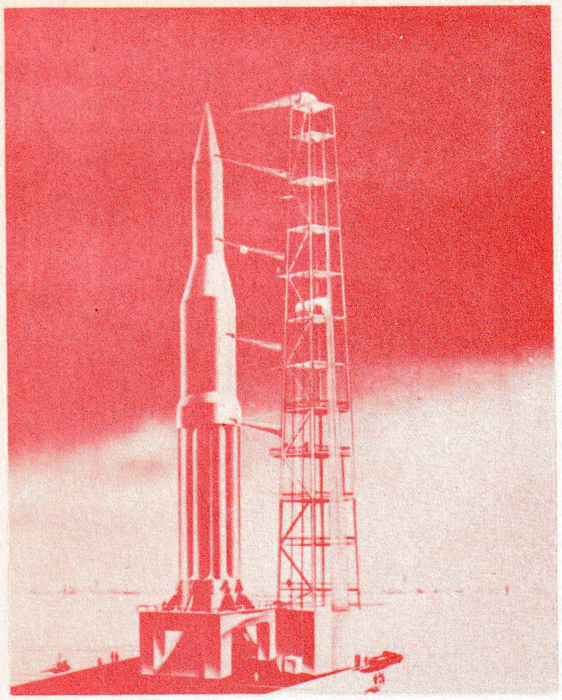
Title
Human Factors Testing in Weapon and Space Systems

Authors
Cooper, Joel
Rigby, Lynn
Spickard, William

Publication Date
1961-03-15



American Rocket Society

MISSILE AND
SPACE VEHICLE
TESTING CONFERENCE

MARCH 13-16, 1961
BILTMORE HOTEL
LOS ANGELES, CALIFORNIA

HUMAN FACTORS TESTING IN WEAPON AND SPACE SYSTEMS

by
Joel Cooper, Lynn Rigby and William Spickard
Northrop Corporation, Norair Division
Hawthorne, California

1646-61

American Rocket Society
500 Fifth Avenue
New York 36, N. Y.

HUMAN FACTORS TESTING IN WEAPON AND SPACE SYSTEMS

by

Joel Cooper
Lynn Rigby
William Spickard

Northrop Corporation, Norair Division

Hawthorne, California

Introduction

The development of system design is a process in which the designer formally or informally predicts that his design will satisfy the performance requirements of the system and then subjects the design to test in order to determine whether the design has met performance requirements. The process is iterative in that, if the design does not meet the performance requirements, it becomes necessary to determine the degree of discrepancy and feed this information back through the system for a design "fix." Although much of this testing is accomplished through formal test programs (qualification, acceptance, etc.), a great deal of analytical work that is done in the design process is essentially testing. Things such as design review boards, etc., are fundamentally "head" testing, in that they call on, to great extent, individual or group experience as to the results of former test programs.

For the designer, the data from test programs constitute the bases for the design "fixes" that must be accomplished. However, he previously reported data in the form of handbook information that has been established through theories and their proofs; i.e., Ohm's Law, Young's Modulus, etc., as well as tables of tube curves, metal strengths and so forth. From these information sources, the designer sets the system design and then determines the performance test requirements necessary to demonstrate the adequacy of the design. Although he uses failure data as a source of information, this is supplementary to performance test data in the adequacy determination.

In contrast to this approach, the human factors specialist, although he acts as a designer in determining or consulting on the system design for human performance, generally relies on failure data or observational data for his design determinations and recommendations. Thus, though the human factors specialist fundamentally has the same types of sources of background data, albeit less precise, he essentially fails to take advantage of the possibilities of methodical performance testing as a source of design information.

Since the human factors specialist has traditionally used malfunction or observational data as source of information, it may be well to examine the problems and uses of these sources as compared with those of performance test data.

Malfunction Data Problems

A study conducted at Stanford Research Institute[3] has indicated that for seven missile systems examined, from 20 to 53 percent of all failures can conservatively be traced to human-initiation. Additionally, hold data, which were available on two of these systems, showed that 23 percent of all unscheduled missile holds were human-initiated. Further, in each of the systems examined, contractor test personnel revealed that at least one disastrous launch or flight failure was human-initiated.

This record of human initiated failures should supply the human factors specialist with a good deal of information as to the performance of the human in the system, yet suprisingly, this does not seem to be borne out by the facts. In the study reported above an examination was made of malfunction reports in two of the systems studied. The reports were separated into two malfunction classifications; one in which the malfunctions could definitely be traced to human-initiation in the immediate causal chain, and another in which human-initiation was in any way doubtful. This would allow for a conservative estimate of the malfunctions which were classified as human-initiated. In each of the systems examined a code provided for a classification entitled "Human Error." Below is a table from the same report in which is shown a comparison of code reporting with analysis of the reports.

Missile	Number of Malfunctions	Malfunctions Labeles as Human Error on Reports	Malfunctions Analyzed as Human-Initiated
A	1391	3	322
B	977	0	193

* From Shapero, et. al. [3/]

Two points seem of importance in the apparent discrepancy of the data above. The first is that malfunction data reporting systems are, in the main, hardware oriented. Reports generally come into data analysis by way of a "failed hardware" tag of sorts. There seems to little if any formal reporting of human-initiated failures unless there is a resulting piece of failed hardware. Most reporting systems examined were limited to a list of failed parts (to varying detail of description), use time, replacement parts, area of occurrence, failure sypmtons, cause of failure, and corrective action taken. Data such as these, however, prove of little use to the human factors specialist for the analysis of the human contribution.

The second point is that personnel shy away from reporting on themselves or their fellow-workers. Rappaport and Cooper[2/] obtained critical incidence data from narrative type malfunction reports and compared these data with data from interviews with test personnel for the same system. They found that written reports reveal mainly those failures to which a "failed equipment" tag can be attached. The verbal reports, on the other hand, indicated that there were a great many failed operations that went unreported. Additionally design deficiencies were practically undetected in written reports yet occupied a sizeable portion of verbal reports.

If the written reports were analyzed alone, it would seem that the most fruitful areas for correction would be in preventing faulty construction. The interview data, however, would demand that faulty operations be looked at first. It seems evident that the information which can be obtained from a reporting system is, to a great extent, a function of the reporting form itself which is, in turn, a function of for whom the reporting form is originated as well as who will report on the form. Therefore, unless the human factors specialist can tailor the malfunction forms to yield specific data for his needs, he can expect little information from malfunction reporting ~~as~~ it is presently practiced.

2copper

Test Performance Data for the Human Factors Specialist

To some extent, human factors specialists have for years been using human test performance data. These have originated in the psychological laboratory and field experiments. These data are limited but if the limitations are realized and used properly, they quite often form a solid basis for a design hypothesis. Further, many of these hypotheses have been validated to yield much of the human factors engineering principles that are of use today. Too often, however, the human factors specialist has simply furnished the design engineer with the principles that have evolved from these experiments. From these principles the design engineer is expected to design a system in which the human can perform adequately. At other times, the human factors specialist runs an experiment in which he determines a choice between two or more given alternative designs. In the latter case, there is quite often no determination whether either of the alternatives are adequate to satisfy performance requirements of the system.

Present test programs are designed to yield performance test information on a variety of design predictions in order to feed back data for redesign of the system. Since information from these programs is usually in response to specific demands of the aerodynamicist, electronic engineer, and various other hardware-oriented engineers, the human factors specialist rarely finds data that are useful to him, except some information that is inadvertantly revealed during the course of test. Usually his presence at the test site, if he indeed is allowed there, is limited to what he can observe while keeping out of the way.

Many human factors specialists complain that they have no opportunity to obtain specific data from performance tests. However, there is a real question as to whether they can specify the test form and the data requirements for their needs. Unless there is a method for clearly specifying their needs, as well as the form and method which can supply this, it seems doubtful that there will be any sizeable amount of testing in the area of human performance. The needs for performance data, as indicated by the findings on human-initiated malfunctions, seems incompatible with the likelihood of getting this information until a method of approach to the problem can be delineated.

An Approach to Human Factors Performanct Testing

It is apparent that any suggested approach cannot be revolutionary in that it completely revamps present system practices. However, the mission of any system can be, and usually is, divided into a series of submissions or functions which will be performed for test or operational use of the system. These functions are essentially operations and it is in an operational context that the human enters and performs in the system. The operational context is dynamic, the human is a dynamic element in the system and it is logical that if there is to be performance testing for the human in the system it sould be in this dynamic context.

There is logic behind the use of operations as a basis for testing; the dynamics previously mentioned, the fact that operations can be chosen at any convenient level, the difficulties involved in complete system test, the fact that one part of a system can be demonstrated before the rest is complete, the sequential dependencies that occur between functions, and the ability to demonstrate system interactions for all system elements concerned in the operation. However, it is necessary to state some hypothesis to be tested and the academic training of the human factors specialist is ideally suited to set the hypothesis and form the experimental design to validate.

For the formulation of this hypothesis he needs the historical data that the previous laboratory and field experiments have supplied.

The results of such performance testing would add to the store of useful historical data, lead to further hypothesis for test, and establish the human factors specialist in a society (engineering) which implicitly states a hypothesis in each design decision that it makes. Further, this society demonstrates the efficacy of the prediction (design decision)by setting the test to prove the prediction.

Since there is an expression of operations presently in use in systems, it seems logical that if these operations were objectified and analyzed they could be subjected to the following measurements:

a. The mean-man time in which the individual operation could be performed. This would require either references to known times for given tasks or estimates of such times. But such estimates can be reasonably made, and much more exacting times can readily be obtained via experimental techniques.

b. The cost of operations in terms of the required instrumentation, expendables, etc. Cost is an ambiguous criterion, for cost ramifications are frequently elusive. However, estimates of the cost of a given operation could be made, or costs could be computed with greater or lesser degrees of accuracy, depending on the issue at hand. The fact remains that mantime cost, for instance, can be readily compared to automation costs and that similar cost measures would be extremely useful in accomplishing design trade-offs.

c. The variability with which operations can be performed. In a sense, variability could be measured in terms of the error bias indicated in the performance of the task type. Probably a simple nominal scale would provide both a means of estimating and evaluating variability as an operations criterion.

d. The accuracy with which the operation can be performed; e.g., the level of confidence regarding the degree to which the task or function can be performed reliably to specified tolerances. This, too, could be measured in an exacting fashion, but again, such exacting measurement would prove impractical, for such measurement would require considerable experimentation on each and every operation. This would not be feasible when a designer was waiting to make a design decision. Here, too, a simple nominal scale would probably serve the purpose.

A look at measurements of operations such as proposed should eventually lead to a store of quantitative data for standard handbooks for human factors engineering.

Choice of Operations to be Examined.

Ideally, the course of system development would be such as to examine each operation that will be performed. Practically, the complexity, time limitations, economic factors, and pure lack of knowledge of what operations will be performed preclude an undertaking as voluminous as this. It becomes necessary, then, to make a decision as to the choice of operations to be examined. To look at operations in this context it is necessary to examine them within a system. Some operations are unique in the system in that, though they may be affected by previous operations, they themselves affect no succeeding operation. Others are sequential and affect succeeding operations as well as being affected by previous operations. As such, it may be well to ask of any operation.

4 coord

How many other operations are affected by the failure of this operation? Secondly, though the operation may affect many other operations, questions as to the effects of the loss of the operation must be asked. These questions can be stated thus:

Will the loss of the operation abort the mission?

Will the loss of the operation degrade the mission, and to what degree? Additionally, it must be recognized that operations are subject to types of failures, such as:

Omission of operation or neglect to accomplish.

Accidental activation, interference or damage of related features.

Inaccurate performance of operation.

Each operation should be examined in terms of its susceptibility to these types of failures. To summarize, there are two areas for consideration of operational malfunction; likelihood of malfunction and results of malfunction. If it is found that an operational failure is critical, it is wise to ask:

Can the operation be performed differently?

Can the operation be performed at a different time?

On the other hand, if the operation is found to be non-critical, it is in order to ask:

Is the operation necessary?

The level at which the operation is chosen will certainly have an effect on the answers to the proposed questions. For example, if the operation is chosen at a level as broad as "launch missile," the loss of the operation would certainly abort the mission. If instead, the operation were chosen at a level as low as "tighten screw on access door" and there were sixteen screws in the door, the failure to tighten one screw would probably have little effect on the success of the mission. Where an operational failure will cause a system failure, it may be necessary to overlook certain efficiencies in order to increase system reliability.

System Testing

The consideration thus far has proposed an abstraction of elements of the system for consideration. Inasmuch as the problem that the human engineer faces is one of total system performance, it becomes necessary to examine human performance under system operational conditions; i.e., field test. The development of equipment follows a prescribed course of design, development etc. with testing at each level. "Knowns" in this context are not usually subject to test, but unknowns and interactions are specifically tested through the course of development. At a later stage the total system is field tested under conditions which attempt to simulate "real" conditions as closely as possible.

The field test allows, for the first time, a test of the total man machine system with all its interactions, coordination and concomitant problems. Meister[1] sums the reasons and results of human engineering field testing thus:

1. Evaluation of the adequacy of the human engineering recommendations made during design and development.
2. Resolution of human factors problems revealed by testing, which will lead to improvement in the design of equipment and procedures.
3. The new test environment presents special problems of its own because man-machine

5 COOPER

system elements must now function in a coordinated fashion. Special emphasis must be placed on procedures for interrelating operation, maintenance and communication functions. This is not to say that no thinking has been done about these interrelations prior to the test phase; but these interrelationships must now be checked out and improved.

4. To measure the operability, maintainability and communications adequacy of the system.

In the field test, time, error and variability must be recorded; the means for accomplishing are dependent on the particular situation. Control and manipulation of variables, such as would be done under laboratory conditions, is virtually impossible for the human engineer. Observation, recorded data, and interviews with personnel must be obtained without changing test conditions. Thus field testing becomes, to a great extent, a means of verifying the prediction of the human engineer.

Hopefully, the original predictions and consequent design recommendations will be sufficiently valid so as to impose little necessity for design or procedural change. Those changes which result from the interaction effects should provide some basis for predicting interaction effects for future systems.

Space systems present the same problems in testing albeit in a somewhat different environment. The most logical choice of a test facility seems to lie in an adaptation of a "Space Crew Holding Facility" (See Astronautics, February 1960). Since the concept itself is one of simulation of the operating conditions in space, it is particularly adaptable to providing the test information short of actual space conditions.

The human engineer then must assume the obligation of using the knowledge that he has available to make predictions of design adequacy in terms of time, error and variability, test the individual predictions in a dynamic content, feed back the error in terms of design changes, and verify the prediction in a setting that includes the problems of interaction between system elements.

BIBLIOGRAPHY

1/ Meister, D., Human Engineering in the Field, Unpublished Paper, Convair Astronautics, San Diego, July 1960.

2/ Rappaport, Maurice J., and Cooper, Joel I., A Preliminary Study of the Human Factors Problems Associated with the Operation of the (name is confidential) Missile System, SRI Project No. ID-2274, November 1957.

3/ Shapero, Albert; Cooper, Joel I.; Rappaport, Maurice J.; Schaeffer, K. H.; and Bates, Charles, Jr.; Human Engineering Testing and Malfunction Data Collection in Weapon System Test Programs, WADD TR 60-36, February 1960.

bcooper