DOCUMENT RESUME

ED 114 640 CE 005 506

AUTHOR	Foley, John P., Jr.
TITLE	Criterion Referenced Measures of Technical Proficiency in Maintenance Activities. Final Report.
INSTITUTION	Air Force Human Resources Lab., Wright-Patterson AFB, Ohio. Advanced Systems Div.
SPONS AGENCY	Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO	AFHRL-TR-75-61
PUB DATE	Oct 75
NOTE	12p.; Tables 1, 2, and 3 will not reproduce in microfiche due to smallness of type; Paper presented at the Annual Meeting of the Human Factors Society (Bedford, Texas, October 14-16, 1975)
EDRS PRICE	MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS	*Criterion Referenced Tests; Electronic Technicians; *Equipment Maintenance; Machine Repairmen; Maintenance; Measurement Instruments; *Military Personnel; Military Training; *Performance Tests; Speeches; Task Performance; *Test Construction; Test Results; Test Validity
IDENTIFIERS	*Air Force

ABSTRACT

The paper concerns instruments used in the determination of task performance efficiency of mechanical and electronic maintenance men. Composite results of unvalidated knowledge tests, theory tests, school marks, and supervisors' ratings are presented which indicate that these measurement instruments have low empirical validity. The limitations of traditional systems effectiveness measures are also discussed. Although job task performance tests (JTPT) have higher empirical validity, they have had limited use due to their relatively high cost factors. A series of efforts to develop better criterion referenced JTPT, and to attempt the development of paper and pencil symbolic substitute tests of high empirical validity, were made. Described is the model battery of 48 criterion referenced JTPT covering all key maintenance activities (checkout, align/adjust, remove/replace, troubleshooting, test equipment, and soldering), intended for use on the job, in training, and for validation of symbolic substitute tests. Developed batteries of graphic and video symbolic substitute tests were given limited validations. Suggestions are made for use of the criterion referenced JTPT in improving maintenance efficiency. Fourteen references are listed, and an appended section describes limitations of system effectiveness measures. (Author/MS)

AFHRL-TR-75-61

NOV 3 1975

AIR FORCE

HUMAN RESOURCES

CRITERION REFERENCED MEASURES OF TECHNICAL PROFICIENCY IN MAINTENANCE ACTIVITIES

By

John P. Foley, Jr.

ADVANCED SYSTEMS DIVISION
Wright-Patterson Air Force Base, Ohio 45433

October 1975

Approved for public release; distribution unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Advanced Systems Division, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base, Ohio 45433, under project 1710, with Hq Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

GORDON A. ECKSTRAND, Director
Advanced Systems Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFHRL-TR-75-61	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CRITERION REFERENCED MEASURES OF TECHNICAL PROFICIENCY IN MAINTENANCE ACTIVITIES		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) John P. Foley, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Advanced Systems Division Air Force Human Resources Laboratory Wright-Patterson Air Force Base, Ohio 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 17101007
11. CONTROLLING OFFICE NAME AND ADDRESS Hq Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE October 1975
		13. NUMBER OF PAGES 10
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Paper presented at Annual Meeting of the Human Factors Society, Bedford, Texas 76201, 14–16 October 1975.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

symbolic substitute job tests	system effectiveness measures
Job Task Performance Tests	personnel selection
criterion referenced tests	psychology individual and group
video testing	measurement and evaluation technical training
maintenance effectiveness	measurement and evaluation maintenance training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper concerns instruments used in the determination of how efficiently maintenance men perform the various tasks of their jobs. Currently, a great deal of reliance is placed on unvalidated knowledge tests, theory tests, school marks, and supervisor's ratings for such determination. This paper presents a composite of the results of such instruments. These results indicate that these measurement instruments have low empirical validity. The limitations of traditional systems effectiveness measures are also discussed. The limitations of all these measurement procedures are not widely known by maintenance supervisors, training managers and maintenance instructors. Although job task performance tests (JTPT) have higher empirical validity, such tests have had very limited use because of their relatively high cost in time, personnel and equipment. As a result of the foregoing findings, the Air Force Human

(Over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 19 (Continued):

measurement and evaluation electronics training
measurement and evaluation vocational education
human subsystems
human factors measures

Item 20 (Continued):

Resources Laboratory (AFHRL) supported a series of efforts to develop better criterion referenced JTPT, and to attempt the development of paper and pencil symbolic substitute tests of high empirical validity.

The paper describes the model battery of 48 criterion referenced JTPT, which has been developed to cover all key maintenance activities such as checkout, align/adjust, remove/replace, troubleshooting, test equipment and soldering. During this development many factors were considered including the identification and classification of tasks to be measured, the hierarchal relationship of maintenance tasks, the most effective order of their measurement and the ease of test administration. This battery was developed as a model of JTPT to be used on the job and in training.

It was also intended as a battery of criterion tests for the validation of paper and pencil symbolic substitute tests. Batteries of graphic and video symbolic substitute tests were developed and given limited validations. The validation of graphic symbolic substitute tests indicated that the symbolics for all activities with the exception of soldering, have promise. The paper discusses the requirements for additional refinement and validation for the various graphic tests. An unsuccessful effort to develop video symbolic substitute tests is also mentioned. Suggestions are made for the application of criterion referenced JTPT for the improvement of maintenance efficiency.

PREFACE

This report represents a portion of the exploratory development program of the Advanced Systems Division, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base, Ohio.

The preparation of this report was documented under Task 171010, Evaluating the performance of Air Force Operators and Technicians of Project 1710, Training for Advanced Air Force Systems. The effort represented by this volume was identified as work unit 17101007. The author was the task scientist. Dr. Ross L. Morgan was the project scientist.

CRITERION REFERENCED MEASURES OF TECHNICAL PROFICIENCY IN MAINTENANCE ACTIVITIES

John P. Foley, Jr. Ed.D.

Air Force Human Resources Laboratory
Wright-Patterson Air Force Base, Ohio

ABSTRACT

The paper concerns instruments used in the determination of how effeciently maintenance men perform the various tasks of their jobs. Currently, a great deal of reliance is placed on unvalidated job-knowledge tests, theory tests, school marks, supervisors' ratings for such determination. This paper presents a composite of the results of many research and development efforts concerning the empirical validity of such instruments. These results indicate that these measurement instruments have low empirical validity. The limitations of traditional systems effectiveness measures are also discussed. Although job task performance tests (JTPT) have much higher empirical validity, such tests have had very limited use because of their relatively high costs in time, personnel and equipment. As a result of the foregoing findings, the Air Force Human Resources Laboratory (AFHRL) supported a series of efforts to develop better criterion referenced JTPT, and to attempt the development of paper and pencil symbolic substitute tests of high empirical validity. The paper describes the model battery of 48 criterion referenced JTPT, which has been developed to cover all key maintenance activities such as checkout, align/adjust, remove/replace, troubleshooting, test equipment and soldering. During this development many factors were considered including the identification and classification of tasks to be measured, the hierarchal relationship of maintenance tasks, the most effective order of their measurement and the ease of test administration. This battery was developed as a model of JTPT to be used on the job and in training. It was also intended as a battery of criterion tests for the validation of paper and pencil symbolic substitute tests. Batteries of graphic and video symbolic substitute tests were developed and given limited validations. The validation of graphic symbolic substitute tests indicated that the symbolics for all activities, with the exception of soldering, have promise. The paper discusses the requirements for additional refinement and validation for the various graphic tests. An unsuccessful effort to develop video symbolic substitute tests is also mentioned. Suggestions are made for the application of criterion referenced JTPT for the improvement of maintenance efficiency.

INTRODUCTION

Human factors specialists can point with great pride to their achievements in the area of the interface between machines and humans. However, these achievements involve primarily the operators of the hardware system and not the maintenance of the system.

Part of this emphasis on the human operator seems to be due to the fact that the operator has much greater visibility. He performs when the machine performs It is quite obvious that the performance of the system usually is directly related to the operator's performance.

Many actions are taken to maximize effective and efficient performance of the operator. Work stations are human engineered to maximize the efficiency and comfort of the human operator. Major training facilities are provided so that operators can enjoy a large amount of supervised practice in performing typical tasks of his job. Graduation from training is based primarily on demonstrated ability to perform job tasks. And, periodic checks are made of the operator's ability to perform the critical tasks of his job. These, of course, are not all of the many efforts made to maximize the performance of human operators.

The important contributions of the maintenance man to the successful performance of the man-machine system are less visible than the operator's contribution. Most of the maintenance man's performance takes place while the man-machine system is not performing its mission. The effective end-product of the maintenance man's efforts, which are performed during this downtime, is an effectively operating machine for its next mission. There is no doubt that such maintenance must be effective. But, given sufficient resources, maintenance need not be *efficient* to be effective. And, at present, very little is done to ensure such efficiency.

SYSTEMS EFFECTIVENESS MEASURES

A critical examination of systems effectiveness parallels the above discussion. For the man-machine system to be efficient and reliable during a mission, the machine subsystem must be efficient and reliable during the mission. The maintenance man makes his contribution to these measures before the performance of each mission by ensuring that the machine subsystem is in prime operating condition when the mission is started. This contribution must be effective or the mission probably will not begin.

Once the mission starts, the efficiency of the system depends on the performance of both the machine and its operator(s). If the mission succeeds in an efficient manner both have operated efficiently. However, if the mission is performed inefficiently, the cause may be either the human subsystem or the machine subsystem. And, if the inefficiency is traced to the machine subsystem, it may be further traced to ineffective maintenance. The faulty maintenance can be corrected. However, the major point is that even when faulty maintenance is discovered and corrected, it need only be effective and not necessarily efficient. The usual measures of system effectiveness include indices of the effectiveness and efficiency of the human operator, but they represent only the effectiveness, and not the efficiency, of maintenance personnel. (A more detailed discussion of the relationships between systems effectiveness measures and the efficiency of maintenance is presented in Appendix A.)

This acceptance of effectiveness without efficiency seems to "set the stage" for the acceptance of inefficiency in maintenance. Unfortunately, methods used to select, train and promote maintenance personnel also contribute to inefficient maintenance.

MAINTENANCE PERSONNEL EFFECTIVENESS MEASURES

The personnel system, which includes formal training, has developed its own measures for making initial selection, for ascertaining effectiveness of training and for the promotion of maintenance personnel. Most of these measures involve tests which are of the paper-and-pencil, job-knowledge variety. The effectiveness of formal training for the *mechanical maintenance specialities* is measured mainly by scores obtained from such paper-and-pencil job-knowledge tests, even though the students in these training programs have received at least some "hands-on" practice on many mechanical maintenance tasks.

The measures of effectiveness of formal training programs for the *electronic maintenance specialisties* include scores from paper-and-pencil job-knowledge tests, as well as theory tests. Students in these electronic maintenance courses receive little if any "hands on" practice in their maintenance tasks.

The selection tests for both mechanical and electronic maintenance specialities have been standardized against composite scores from paper-and-pencil tests. This means that the people selected for the maintenance specialities have been selected not on their aptitude for performing the tasks of their maintenance jobs but on their aptitude for making high scores on paper-and-pencil, theory and job-knowledge tests.

The speciality knowledge test (SKT) and the promotion fitness examination (PFE) used for advancement up the maintenance career ladders also are paper-and-pencil job-knowledge tests. At the present time, throughout his whole career, a maintenance specialist is not required to demonstrate on formal job task performance tests (JTPT) that he can perform efficiently and effectively the tasks of his job (Foley, 1974).

This state of affairs would not be so bad if there was a high empirical relationship between *scores* received on paper-and-pencil theory and job-knowledge tests and *scores* received on JTPT, but this is not the case. Research evidence gives a rather low rating to all of these paper-and-pencil tests. Table 1 shows correlations that have been obtained by comparing job task performance tests (JTPT) to theory tests, and to job-knowledge tests. The latter two are paper-and-pencil tests. Table 1 also includes correlations of JTPT with school marks. As indicated earlier, school marks have been heavily weighted with the paper-and-pencil test scores. An examination of this table indicates that the correlations of JTPT scores with theory test scores are generally somewhat lower than with job-knowledge tests. None of these measures are sufficiently valid for use as substitutes for JTPT (Foley, 1974).

Table 1. Correlations Between Job-Task Performance Tests and Theory Tests, Job Knowledge Tests, and School Marks

Researchers	Type of Job Task Performance Test (JTPT)	Theory Tests	Job Knowledge Tests	School Marks
Anderson (1962a)	Test Equipment JTPT			.18–.33
Evans and Smith (1953)	Troubleshooting JTPT	.24 & .36	.12 & .10	.35
Mackie et al., (1953)	Troubleshooting JTPT	.38		.39
Saupe (1955)	Troubleshooting JTPT		.55	.56
Brown et al., (1959)	Troubleshooting JTPT	.40		
	Test Equipment JTPT		.29	
	Alignment JTPT		.28	
	Repair Skills JTPT		.19	
Williams and Whitmore (1959)	Troubleshooting JTPT			
	(Inexperienced Subjects)	.23		
	(Experienced Subjects)	.15		
	Adjustment JTPT			
	(Inexperienced Subjects)	.02		
	(Experienced Subjects)	.21		
	Acquisition Radar JTPT			
	(Inexperienced Subjects)	.03	.36	
	(Experienced Subjects)	.14	.22	
	Target Tracking Radar JTPT			
	(Inexperienced Subjects)	.24	.33	
	(Experienced Subjects)	.20	.38	
	Missile Tracking Radar JTPT			
	(Inexperienced Subjects)	.09	.15	
	(Experienced Subjects)	.19	.32	
	Computer JTPT			
	(Inexperienced Subjects)	.08	.24	
	(Experienced Subjects)	.06	.14	
	Total JTPT			
	(Inexperienced Subjects)	.14		
	(Experienced Subjects)	.20		
Crowder et al., (1954)	Troubleshooting JTPT	.11	.18–.32	

A conclusion that can be drawn from the discussion, thus far, is that there is nothing in the job structure of individuals in the maintenance subsystem that requires them to perform maintenance tasks efficiently. Systems effectiveness measures requires effective, but not efficient, performance of maintenance tasks. The paper-and-pencil measures used for selection, training and promotion require no actual performance of maintenance tasks. Wallace (1965) expressed this problem very succinctly as it applies to electronic maintenance personnel. A somewhat similar statement could be made for mechanical maintenance personnel. I quote:

> All of this is prelude to my main thesis which is in no sense revolutionary, original, or controversial. I state it because it is honored in the breach. It is that the nature of our proficiency measures determines how we select, classify, training, maintain and assess our human resources. If the measures are largely irrelevant to the jobs we want done, we will select the wrong men, classify them incorrectly, and train them wrong. This is true because these proficiency measures are, or should be, the criteria against which we validate our selection and classification procedures and evaluate our training content and methodology or our supervisory techniques. Thus, if I use a test of advanced electronics theory as the proficiency measure for electronics maintenance and as the criterion against which to evaluate a test for selecting men to go into maintenance training, I will end up choosing a selection test which rejects men who are not well above average in both reading and arithmetic ability. In the process I might reject a great many who are outstanding in their ability to get their hands on a piece of machinery and make it work. I might also accept a number who (like myself) are so lacking in the simplest manipulative ability that their hands could have been cut off at the wrists at birth without seriously affecting their outputs. So, when I decided what proficiency measures to use, I also decided what kind of men I was going to put into training for the job.
>
> But It doesn't end there. For when I now approach the problem of how to train men to perform the tasks involved in the job, I must make decisions about what should be taught and what methods should be used in teaching it. The only way I have of reaching such decisions (except by divination which is, admittedly, not a rare procedure) is to measure and compare the performances achieved with various curricula and methodologies. So, in the case of the electronics maintenance course, I put in lots of reading about electronics theory and I produce graduates who can read and write electronics theory, while their equipment deteriorates in hopeless inoperativeness (Wallace, 1965, p.4).

APPLICABLE RESEARCH AND DEVELOPMENT

Starting in 1969, the Advanced Systems Division of the Air Force Human Resources Laboratory supported a modest program to provide the Air Force with the necessary tools for measuring the ability of maintenance personnel to perform the key tasks of their jobs. This program had two objectives: (1) to develop a model battery of JTPT together with appropriate scoring schemes for the measurement of the task performance ability of electronic maintenance personnel (an effort was to be made for the development of JTPT which could be easily administered), and (2) using this battery as criteria, to develop and try out a series of paper-and-pencil symbolic substitute tests that would hopefully have high empirical validity.

Criterion Referenced Job Task Performance Tests

A model battery of 48 criterion referenced JTPT and a test administrator's handbook were developed for measuring ability to perform electronic maintenance tasks. Copies of the actual instructions for test subjects together with the test administrator's handbook are available from the Defense Documentation Center (DDC) as AFHRL Technical Report 74-57(II), Part II (Shriver et al., 1975). The test administrator's handbook was developed with step-by-step detailed instructions so that an individual with a minimum of electronic maintenance experience can administer the tests.

The battery includes separate tests for the following classes of job activities: (1) equipment checkout (2) alignment/calibration, (3) removal/replacement, (4) soldering, (5) use of general and special test equipment, and (6) troubleshooting. The Doppler Radar AN/APN-147 and its computer AN/ASN-35 were selected as a typical electronic system. This system was used as the test-bed for this model battery. The soldering and general test equipment JTPT are applicable to all electronic technicians. The other tests of the battery apply to technicians concerned with this specific doppler radar system. A detailed description of the development and tryout of these JTPT is given in AFHRL-TR-74-57(II), Part I (Shriver and Foley, 1974a).

Each class of activity for which JTPT was developed contains its individual mix of behaviors, but it is not mutually exclusive. There are dependencies among the classes. As a result a four level hierarchy of dependencies can be stated: (1) checkout, removal/replacement, and soldering; (2) use of general and special test equipment; (3) alignment/calibration; and (4) troubleshooting. For example, troubleshooting may include all of the activities mentioned before it.

After considering product, process, and time as to their appropriateness for scoring the results for each activity, it was decide that a test subject has not reached criterion until he had produced a complete, satisfactory product. This was a go, no-go criterion. Tests were prepared for seven classes of tasks; namely, (1) checkout, (2) physical skill tasks (soldering), (3) remove and replace, (4) test equipment, (5) adjustment, (6) alignment, and (7) troubleshooting.

Table 2 summarizes the number of tests, problems and scorable products by class developed for the AN/APN-147 and AN/ASN-35. The simple addition of numbers shown in Table 2 indicates that there are 48 tests, 81 problems, and 133 scorable products. But, these numbers tell us nothing in terms of the content of the tests. To say that one test subject accomplished 100 scorable products while another accomplished 90, tells us nothing about the job readiness of these individuals or that one is better than the other. The varieties of scorable products are so diverse that any combination of them without regard to what they represent is meaningless. The only meaningful presentation of such information must be in terms of a profile designed to attach meaning to such numbers. A sample of such a profile is shown in Figure 1.

Table 2. Tests, Problems, and Scorable Products

Class	Code	Tests	Problems	Scorable Products
1. Checkout	CO	2	2	2
2. Physical Skill Tasks (soldering)	PT	2	5	17
3. Remove and Replace	RR	10	10	20
4. Test Equipment	SE	7	37	67
5. Adjustment	AD	6	6	6
6. Alignment	AL	10	10	10
7. Troubleshooting	TS	11	11	11
Total	7	48	81	133

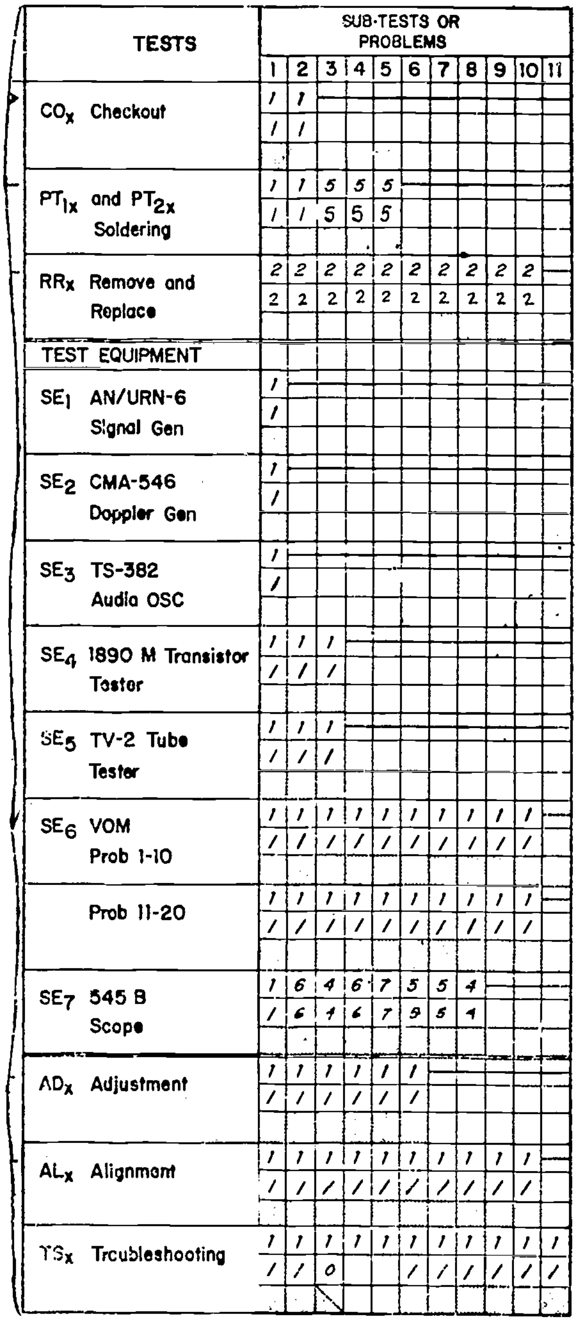

Figure 1. A profile for displaying the results obtained by an individual subject from a battery of Job Task Performance Tests concerning an Electronic System – the AN/APN-147 and the AN/ASN-35. This represents the profile of an individual who has successfully completed most of the battery.

This profile is not presented as the final solution to the profile problem for JTPT for electronic maintenance. It does contain most of the important information regarding a test subject's success on the full range of tests. It gives a meaningful picture of the subject's job abilities as measured by the test battery, indicating the subject's strengths and weaknesses.

An examination of the profile (Figure 1) indicates that most of the tests in this battery contain only one problem. For example, there are two checkout tests, having one problem each and there are eleven troubleshooting tests having one problem each. There are two soldering tests; one has two problems and the other has three. The voltohmmeter (VOM) test has 20 problems.

The subject receives no "credit" for a problem unless he obtains all of the expected products. No attempt is made, to combine these scores in terms of meaningless numbers.

The hierarchy of dependencies mentioned previously has implication for the order in which tests are administered as well as for diagnostics. For example, since troubleshooting includes the use of test equipment and other activities in the hierarchy, logic would dicate that administration of the tests for the sub-activities would precede the troubleshooting tests and that a test subject would not be permitted to take the troubleshooting tests until he had passed these other subtests. These dependencies are displayed on the left hand side of the profile. This portion of the profile is shown in Figure 2.

Due to the unavailability of a sufficient number of experienced test subjects at the time of the tryout of the JTPT battery, the tryout was not as extensive as planned. The limited tryout did indicate that the tests as developed are administratively feasible. Their continued use, no doubt, would result in further modifications and improvements.

Development of Symbolic Substitutes

Because of the time, equipment and personnel requirements of some of the JTPT, the availability of empirically valid symbolic substitute tests would be highly desirable. Even though such tests had not been developed in previous efforts, it was our opinion that much more work could be done to improve symbolic maintenance tests as substitutes for JTPT. It was hypothesized that higher correlations possibly could be obtained by a different approach to the development of symbolic substitute tests. A study of the Tab Tests (Crowder et al., 1954, see Table 1) indicated that the JTPT used as the criterion measures contained many distractions and interruptions to the subject's troubleshooting strategy (cognitive process), such as, using test equipment to obtain test point information. In addition to such interruptions to the cognitive process, the subject can obtain faulty test point information by the improper use of his test equipment. In the symbolic substitute Tab Tests, all of these potential pitfalls of the actual task were avoided. The subject was given a printed test point readout. It was hypothesized that the injection of job equivalent pitfalls into symbolic substitutes possibly would increase their empirical validity.

Based on these hypotheses, a battery of symbolic tests was developed under contract with the Matrix Research Company of Falls Church VA. A companion graphic symbolic test was developed for each of the job activities for which a criterion referenced JTPT had previously been developed. Based on two limited validations, all of the graphic symbolic tests, with the exception of the symbolic test for soldering, indicated sufficient promise to justify further consideration and refinement. Table 3 indicates the correlations obtained from these validations. Due to a shortage of available subjects, the number of pairs of subject was extremely small. All of these promising graphic symbolic tests, therefore, must be given more extensive validations using larger numbers of experienced subjects.

The validation of any such symbolic test requires the administration of a companion JTPT as a validation criterion. As a result, a validation is an expensive process in terms of equipment and experienced manpower. The troubleshooting symbolic tests require the most extensive refinement. Several suggestions are made for improving their empirical validity. A complete description of these symbolic test efforts can be found in AFHRL-TR-74-57(III) (Shriver and Foley, 1974b). An attempt, also, was made to develop video symbolic substitute tests, but this effort produced no promising results. (Shriver et al., 1974).

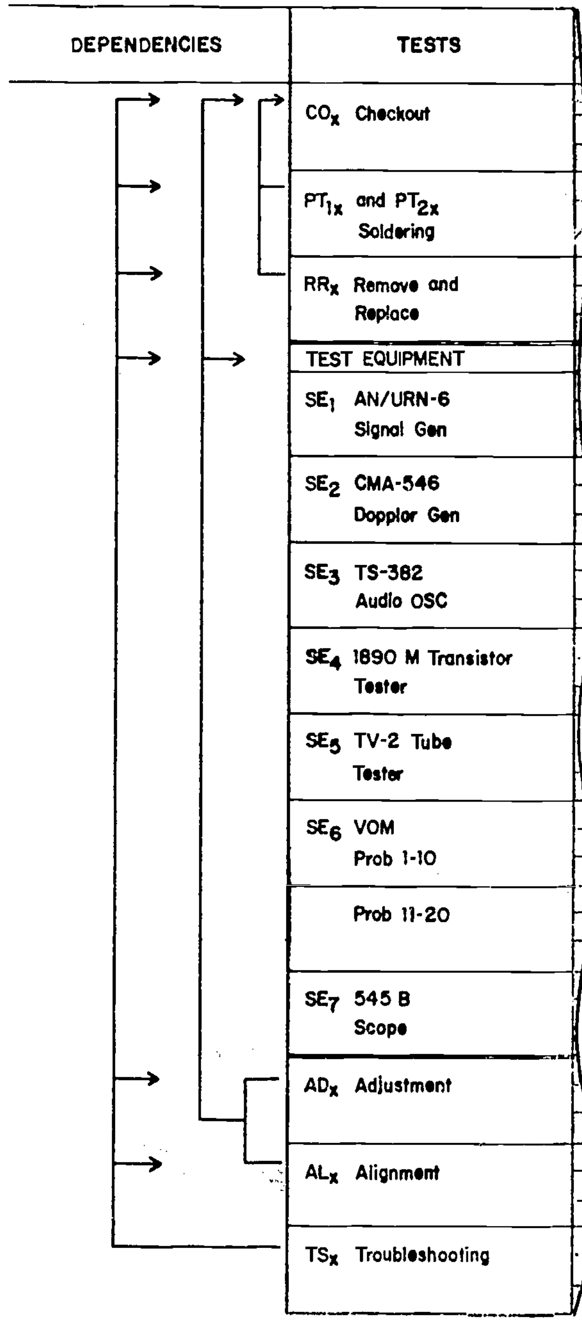

Figure 2. **The left hand portion of the profile displayed in Figure 1. This portion indicates the hierarchy relationships among the various tests of the JTPT battery.**

Table 3. **Indicates the Number of Pairs Used as Well as the X^2 and the Correlations Obtained during Two Small Validations of Symbolic Tests**

Test Area	N Pairs	X^2	ϕ	r_t
Novice Subjects (Altus)				
Checkout	4	4.00	1.00	–
Remove & Replace	14	2.57	.43	–
Soldering Tests	4	0	0	–
General Test Equip	6	2.67	.67	–
Special Test Equip	6	.67	.33	–
Alignment/Adjustment	19	6.37	.58	–
Troubleshooting	9	1.00	–.33[a]	–
Experienced Subjects (TAC)				
Overall Troubleshooti.	30	6.53	.47	.68
Chassis (Black box)				
Isolation	30	16.33	.73	.81
Stage Isolation	30	3.33	.33	.46
Piece/Part Isolation	15	.07	.07	.16

[a]This negative correlation was probably due to a number of deficiencies such as (1) deficiencies in the Fully Proceduralized Job Performance Aids provided the subjects, (2) deficiencies in the sequencing of the troubleshooting JTPT in relatinn to the sub-tests in the JTPT battery, (3) maintenance difficulties with the AN/APN-147 AN/ASN-35 system, and (4) difficulties with the content and administration of test equipment pictorials provided in the original troubleshooting symbolic tests.

Even if graphic symbolic substitutes of high empirical validity can be produced, the use of symbolic substitutes will never, in my opinion, dispense with the requirement for the liberal administration of actual JTPT to maintenance personnel. We can never include all aspects of an actual performance of a task in a paper-and-pencil symbolic representation of that task, but our work indicates that we can come much closer than has been done in the past.

APPLICATION OF JTPT IN MAINTENANCE SUBSYSTEMS

AFHRL-TR-74-57(II), Part I, mentioned earlier contains many suggestions for the application of criterion referenced JTPT in maintenance squadrons and in training (pages 54-56). I will not attempt to discuss all of these here. Today I am going to limit my suggestions for the application of this JTPT technology for the improvement of the efficiency of maintenance subsystems. Any actions taken to improve maintenance efficiency, should give due consideration to the following points.

1. At present there is nothing in the structure of each maintenance subsystem to ensure that its personnel have the ability to correctly perform the maintenance tasks dictated by its maintenance concept. Any general weaknesses in such ability decrease the *efficiency* of the maintenance subsystem.

2. The inability to perform such tasks need not reflect on the appearance of maintenance effectiveness provided that unlimited materiel resources are available to compensate for such personnel weaknesses.

3. Current systems effectiveness measures consider and ensure acceptable effectiveness, but not efficiency, of the maintenance subsystem.

4. The limited amount of available hard data indicates that maintenance personnel generally are not able to perform many of their critical tasks correctly.

5. During their careers, maintenance personnel take many paper-and-pencil job knowledge tests but few if any JTPT. Scores obtained from such job knowledge tests tell us almost nothing about how well people can perform tasks of their maintenance jobs. (I am convinced that the current unquestioned usage of such paper-and-pencil tests in field and training situations would be unforgivable if the people involved really understood how invalid their current testing practices really are. No matter how cheaply paper-and-pencil job-knowledge tests can be prepared or how easily they can be administered, such tests are not a bargain. Their results are almost meaningless in terms of ability to perform maintenance tasks AFHRL-TR-74-57(I) contains a detailed discussion of their criteria and validity).

Considering the above five points, effective action should be taken to modify the structure of each maintenance subsystem so as to reward not only effective performance but also efficient performance. Maintenance personnel should be provided supervised practice in performing key tasks of their jobs. I am suggesting that a battery of JTPT should be developed for each new maintenance subsystem, and that each member of the maintenance subsystem should be tested periodically on his ability to perform key tasks of his job. Special attention should be given to his ability to use test equipment accurately. Sampling techniques can be applied for individuals and groups to reduce the time required for such tests.

I am sure that the first efforts for applying these measures will meet with the traditional objections – especially that such procedures would cost too much in time and money. My answer to such objections is that we can no longer afford the luxury of unlimited materiel support required by our current maintenance subsystem operations. We can no longer afford the luxury of personnel in our maintenance subsystems who can obtain high scores on paper-and-pencil job-knowledge tests but who are unable to perform the critical tasks of their maintenance jobs. So – we can't afford not to include JTPT in the structure of our maintenance subsystems. The model battery which I described briefly today, together with the discussions found in AFHRL-TR-74-57(I) (Foley, 1974) and AFHRL-TR-74-57(II) Part I (Shriver and Foley, 1974a) provide the bases for such action.

REFERENCES

Anderson, A.V. *Training, Utilization on Proficiency of Navy Electronics Technicians: III. Proficiency in the Use of Test Equipment.* Navy Technical Bulletin 62-14, AD-294 330. San Diego: U.S. Navy Personnel Research Activity, 1962.

Brown, G.H., Zaynor, W.C., Bernstein, A.H., & Shoemaker, H.A. *Development and Evaluation of an Improved Field Radio Repair Course.* Technical Report 58, Project Repair. Human Resources Research Office, Washington: The George Washington University, 1959.

Crowder, N., Morrison, E.J., & Demaree, R.G. *Proficiency of 0-24 Radar Mechanics: VI. Analysis of Intercorrelations of Measures.* AFPTRC-TR-54-127, AD-62 115. Lackland AFB, Tex.: Air Force Personnel and Training Research Center, 1954.

Evans, R.N., & Smith, L.J. *A Study of Performance Measures of Troubleshooting Ability on Electronic Equipment.* Illinois: College of Education, University of Illinois, October 1953.

Foley, J.P., Jr. *Evaluating Maintenance Performance: An Analysis.* AFHRL-TR-74-57(I), AD-A004 761. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, October 1974.

Foley, J.P., Jr. *Factors to Consider in Developing New Test and Evaluation Techniques.* "Proceedings of the Human Factors Testing Conference, 1-2 October 1968." Synder, M.T. (Chm.), Kincaid, J.P., & Potempa, K.W. (Ed.), AFHRL-TR-69-6, AD-866 485. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, October 1969.

Mackie, R.R., Wilson, C.L., & Buckner, D.N. *Practical Performance Test Batteries for Electricians Mates and Radiomen Developed in Conjunction with a Manual for use in the Preparation and Administration of Practical Performance Tests.* AD-98 239. Los Angeles, Calif.: Management and Marketing Research Corporation, June 1953.

Saupe, J.L. *An Analysis of Troubleshooting Behavior oJ Radio Mechanic Trainees.* AFPTRC-TN-55-47, AD-99 361. Lackland AFB, Tex.: Air Force Personnel and Training Center, November 1955.

Shriver, E.L., & Foley, J.P., Jr. *Evaluating Maintenance Performance: The Development and Tryout of Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(II), Part I, AD-A004 845. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, September 1974. (a)

Shriver, E.L., & Foley, J.P., Jr. *Evaluating Maintenance Performance: The Development of Graphic Symbolic Substitutes for Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(III), AD-A005 296. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, November 1974. (b)

Shriver, E.L., Hayes, J.F., & Hufhand, W.R. *Evaluating Maintenance Performance: Test Administrator's Manual and Test Subject's Instructions for Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(II), Part II, AD-A005 785. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, January 1975.

Shriver, E.L., Hayes, J.F., & Hufhand, W.R. *Evaluating Maintenance Performance: A Video Approach to Symbolic Testing of Electronics Maintenance Tasks.* AFHRL-TR-74-57(IV), AD-A005 297. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, July 1974.

Wallace, S.R. *The Relationship of Psychological Evaluation to Needs of the Department of Defense* in "Proceedings of 7th Annual Military Testing Association Conference." San Antonio, Tex.: October 1965, 1-10.

Williams, W.L., Jr., & Whitmore, P.G., Jr. *The Development and Use of a Performance Test as a Basis for Comparing Technicians with and without Field Experience. The NIKE AJAX AFC Maintenance Technician.* Technical Report 52, AD-212 663. Human Resources Research Office, Washington, D.C.: The George Washington University, January 1959.

APPENDIX A

LIMITATIONS OF SYSTEM EFFECTIVENESS MEASURES

A critical examination of systems effectiveness measures supports the contention that such measures demand maintenance effectiveness but not necessarily maintenance efficiency. Typical systems effectiveness measures are indicated in column 1 of Table A-1. Column 2 indicates those measures which require operator effectiveness. Column 3 indicates those measures which require maintenance effectiveness.

Table A-1. **Indicating the Levels of Performance Demanded of the Operator and Maintenance Portions of a Human Subsystem by Traditional Systems Effectiveness Measures**

System Effectiveness Measures	Operators	Maintenance Personnel
Efficiency of System to Perform Mission (Accuracy, Time, Etc.)	$E_1 + E_2$	E_2
Reliability of System	$E_1 + E_2$	E_2
Turnaround Time	–	E_2
Down Time	–	E_2
Spare Parts Consumption	–	E_2
Acquisition Cost	–	–
Life Cycle Cost	C	C

E_1 = Efficiency E_2 = Effectiveness C = Contribute to

The table indicates that turn around time, downtime, and spare parts consumption are systems measures that reflect on maintenance effectiveness. But all of these measures are based on a predetermined level of what is considered to be normal based on past history.

Available evidence indicates that such measures are probably masking some expensive defects and weaknesses in the performance of the human portion of the maintenance subsystem. High estimates for these measures, which have been made during a system's design, probably cover up these weaknesses. For example, levels of consumption of spare parts, maintenance actions, down times, and numbers of maintenance technicians which are truly excessive are frequently accepted as "normal." Such excesses would only become apparent when the predetermined levels were not sufficient to cover them. As a result, the inefficiencies of maintenance personnel may not become immediately apparent during system evaluation, if ever. The overall weaknesses of the human maintenance subsystems may not be attributed to the maintenance technicians but to something else. (The evidence to support these statements were presented by the author in a previous paper, Foley 1969. Available data indicate that the craftsmenship of a large number of maintenance men is poor, that they cannot use test equipment effectively, and that they cannot troubleshoot efficiently). Any such inefficiency on the part of maintenance men will of course increase the life cycle of the man-machine system.

The above analysis of systems effectiveness measures indicates that a man-machine system demands efficient, as well as effective performance of the operator portion of its human subsystem, but only demands effective performance of the maintenance portion of its human subsystem. Available data indicates that the actual performance of the maintenance subsystem generally fits this pattern. In my opinion, this acceptance of effectiveness (without efficiency) is an extremely important causative factor, but not the only factor contributing to maintenance subsystem inefficiency.

☆U.S. GOVERNMENT PRINTING OFFICE: 1975- 671-602/8002

Table 3. Indicates the Number of Pairs Used as Well as the X^2 and the Correlations Obtained during Two Small Validations of Symbolic Tests

Test Area	N Pairs	X^2	ϕ	r_t
Novice Subjects (Altus)				
Checkout	4	4.00	1.00	–
Remove & Replace	14	2.57	.43	–
Soldering Tests	4	0	0	–
General Test Equip	6	2.67	.67	–
Special Test Equip	6	.67	.33	–
Alignment/Adjustment	19	6.37	.58	–
Troubleshooting	9	1.00	–.33[a]	–
Experienced Subjects (TAC)				
Overall Troubleshooting	30	6.53	.47	.68
Chassis (Black box) Isolation	30	16.33	.73	.81
Stage Isolation	30	3.33	.33	.46
Piece/Part Isolation	15	.07	.07	.16

[a]This negative correlation was probably due to a number of deficiencies such as (1) deficiencies in the Fully Proceduralized Job Performance Aids provided the subjects, (2) deficiencies in the sequencing of the troubleshooting JTPT in relation to the sub-tests in the JTPT battery, (3) maintenance difficulties with the AN/APN-147 AN/ASN-35 system, and (4) difficulties with the content and administration of test equipment pictorials provided in the original troubleshooting symbolic tests.

Even if graphic symbolic substitutes of high empirical validity can be produced, the use of symbolic substitutes will never, in my opinion, dispense with the requirement for the liberal administration of actual JTPT to maintenance personnel. We can never include all aspects of an actual performance of a task in a paper-and-pencil symbolic representation of that task, but our work indicates that we can come much closer than has been done in the past.

APPLICATION OF JTPT IN MAINTENANCE SUBSYSTEMS

AFHRL-TR-74-57(II), Part I, mentioned earlier contains many suggestions for the application of criterion referenced JTPT in maintenance squadrons and in training (pages 54-56). I will not attempt to discuss all of these here. Today I am going to limit my suggestions for the application of this JTPT technology for the improvement of the efficiency of maintenance subsystems. Any actions taken to improve maintenance efficiency, should give due consideration to the following points.

1. At present there is nothing in the structure of each maintenance subsystem to ensure that its personnel have the ability to correctly perform the maintenance tasks dictated by its maintenance concept. Any general weaknesses in such ability decrease the *efficiency* of the maintenance subsystem.

2. The inability to perform such tasks need not reflect on the appearance of maintenance effectiveness provided that unlimited materiel resources are available to compensate for such personnel weaknesses.

3. Current systems effectiveness measures consider and ensure acceptable effectiveness, but not efficiency, of the maintenance subsystem.

4. The limited amount of available hard data indicates that maintenance personnel generally are not able to perform many of their critical tasks correctly.

5. During their careers, maintenance personnel take many paper-and-pencil job knowledge tests but few if any JTPT. Scores obtained from such job knowledge tests tell us almost nothing about how well people can perform tasks of their maintenance jobs. (I am convinced that the current unquestioned usage of such paper-and-pencil tests in field and training situations would be unforgivable if the people involved really understood how invalid their current testing practices really are. No matter how cheaply paper-and-pencil job-knowledge tests can be prepared or how easily they can be administered, such tests are not a bargain. Their results are almost meaningless in terms of ability to perform maintenance tasks. AFHRL-TR-74-57(I) contains a detailed discussion of their criteria and validity).

Considering the above five points, effective action should be taken to modify the structure of each maintenance subsystem so as to reward not only effective performance but also efficient performance. Maintenance personnel should be provided supervised practice in performing key tasks of their jobs. I am suggesting that a battery of JTPT should be developed for each new maintenance subsystem, and that each member of the maintenance subsystem should be tested periodically on his ability to perform key tasks of his job. Special attention should be given to his ability to use test equipment accurately. Sampling techniques can be applied for individuals and groups to reduce the time required for such tests.

I am sure that the first efforts for applying these measures will meet with the traditional objections – especially that such procedures would cost too much in time and money. My answer to such objections is that we can no longer afford the luxury of unlimited materiel support required by our current maintenance subsystem operations. We can no longer afford the luxury of personnel in our maintenance subsystems who can obtain high scores on paper-and-pencil job-knowledge tests but who are unable to perform the critical tasks of their maintenance jobs. So – we can't afford not to include JTPT in the structure of our maintenance subsystems. The model battery which I described briefly today, together with the discussions found in AFHRL-TR-74-57(I) (Foley, 1974) and AFHRL-TR-74-57(II) Part I (Shriver and Foley, 1974a) provide the bases for such action.

REFERENCES

Anderson, A.V. *Training, Utilization on Proficiency of Navy Electronics Technicians: III. Proficiency in the Use of Test Equipment.* Navy Technical Bulletin 62-14, AD-294 330. San Diego: U.S. Navy Personnel Research Activity, 1962.

Brown, G.H., Zaynor, W.C., Bernstein, A.H., & Shoemaker, H.A. *Development and Evaluation of an Improved Field Radio Repair Course.* Technical Report 58, Project Repair. Human Resources Research Office, Washington: The George Washington University, 1959.

Crowder, N., Morrison, E.J., & Demaree, R.G. *Proficiency of 0-24 Radar Mechanics: VI. Analysis of Intercorrelations of Measures.* AFPTRC-TR-54-127, AD-62 115. Lackland AFB, Tex.: Air Force Personnel and Training Research Center, 1954.

Evans, R.N., & Smith, L.J. *A Study of Performance Measures of Troubleshooting Ability on Electronic Equipment.* Illinois: College of Education, University of Illinois, October 1953.

Foley, J.P., Jr. *Evaluating Maintenance Performance: An Analysis.* AFHRL-TR-74-57(I), AD-A004 761. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, October 1974.

Foley, J.P., Jr. *Factors to Consider in Developing New Test and Evaluation Techniques.* "Proceedings of the Human Factors Testing Conference, 1-2 October 1968." Synder, M.T. (Chm.), Kincaid, J.P., & Potempa, K.W. (Ed.), AFHRL-TR-69-6, AD-866 485. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, October 1969.

Mackie, R.R., Wilson, C.L., & Buckner, D.N. *Practical Performance Test Batteries for Electricians Mates and Radiomen Developed in Conjunction with a Manual for use in the Preparation and Administration of Practical Performance Tests.* AD-98 239. Los Angeles, Calif.: Management and Marketing Research Corporation, June 1953.

Saupe, J.L. *An Analysis of Troubleshooting Behavior of Radio Mechanic Trainees.* AFPTRC-TN-55-47, AD-99 361. Lackland AFB, Tex.: Air Force Personnel and Training Center, November 1955.

Shriver, E.L., & Foley, J.P., Jr. *Evaluating Maintenance Performance: The Development and Tryout of Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(II), Part I, AD-A004 845. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, September 1974. (a)

Shriver, E.L., & Foley, J.P., Jr. *Evaluating Maintenance Performance: The Development of Graphic Symbolic Substitutes for Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(III), AD-A005 296. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, November 1974. (b)

Shriver, E.L., Hayes, J.F., & Hufhand, W.R. *Evaluating Maintenance Performance: Test Administrator's Manual and Test Subject's Instructions for Criterion Referenced Job Task Performance Tests for Electronic Maintenance.* AFHRL-TR-74-57(II), Part II, AD-A005 785. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, January 1975.

Shriver, E.L., Hayes, J.F., & Hufhand, W.R. *Evaluating Maintenance Performance: A Video Approach to Symbolic Testing of Electronics Maintenance Tasks.* AFHRL-TR-74-57(IV), AD-A005 297. Wright-Patterson AFB, OH.: Advanced Systems Division, Air Force Human Resources Laboratory, July 1974.

Wallace, S.R. *The Relationship of Psychological Evaluation to Needs of the Department of Defense* in "Proceedings of 7th Annual Military Testing Association Conference." San Antonio, Tex.: October 1965, 1-10.

Williams, W.L., Jr., & Whitmore, P.G., Jr. *The Development and Use of a Performance Test as a Basis for Comparing Technicians with and without Field Experience. The NIKE AJAX AFC Maintenance Technician.* Technical Report 52, AD-212 663. Human Resources Research Office, Washington, D.C.: The George Washington University, January 1959.

APPENDIX A

LIMITATIONS OF SYSTEM EFFECTIVENESS MEASURES

A critical examination of systems effectiveness measures supports the contention that such measures demand maintenance effectiveness but not necessarily maintenance efficiency. Typical systems effectiveness measures are indicated in column 1 of Table A-1. Column 2 indicates those measures which require operator effectiveness. Column 3 indicates those measures which require maintenance effectiveness.

Table A-1. **Indicating the Levels of Performance Demanded of the Operator and Maintenance Portions of a Human Subsystem by Traditional Systems Effectiveness Measures**

System Effectiveness Measures	Operators	Maintenance Personnel
Efficiency of System to Perform Mission (Accuracy, Time, Etc.)	$E_1 + E_2$	E_2
Reliability of System	$E_1 + E_2$	E_2
Turnaround Time	–	E_2
Down Time	–	E_2
Spare Parts Consumption	–	E_2
Acquisition Cost	–	–
Life Cycle Cost	C	C

E_1 = Efficiency E_2 = Effectiveness C = Contribute to

The table indicates that turn around time, downtime, and spare parts consumption are systems measures that reflect on maintenance effectiveness. But all of these measures are based on a predetermined level of what is considered to be normal based on past history.

Available evidence indicates that such measures are probably masking some expensive defects and weaknesses in the performance of the human portion of the maintenance subsystem. High estimates for these measures, which have been made during a system's design, probably cover up these weaknesses. For example, levels of consumption of spare parts, maintenance actions, down times, and numbers of maintenance technicians which are truly excessive are frequently accepted as "normal." Such excesses would only become apparent when the predetermined levels were not sufficient to cover them. As a result, the inefficiencies of maintenance personnel may not become immediately apparent during system evaluation, if ever. The overall weaknesses of the human maintenance subsystem may not be attributed to the maintenance technicians but to something else. (The evidence to support these statements were presented by the author in a previous paper, Foley 1969. Available data indicate that the craftsmanship of a large number of maintenance men is poor, that they cannot use test equipment effectively, and that they cannot troubleshoot efficiently). Any such inefficiency on the part of maintenance men will of course increase the life cycle of the man-machine system.

The above analysis of systems effectiveness measures indicates that a man-machine system demands efficient, as well as effective performance of the operator portion of its human subsystem, but only demands effective performance of the maintenance portion of its human subsystem. Available data indicates that the actual performance of the maintenance subsystem generally fits this pattern. In my opinion, this acceptance of effectiveness (without efficiency) is an extremely important causative factor, but not the only factor contributing to maintenance subsystem inefficiency.

☆U.S. GOVERNMENT PRINTING OFFICE: 1975– 671-602/800