Report No. FAA-RD-81-30

A SYSTEMS ENGINEERING EVALUATION METHOD FOR PILOTED AIRCRAFT AND OTHER MAN-OPERATED VEHICLES AND MACHINES

WITH HYPOTHETICAL EXAMPLE OF A SYSTEMS EVALUATION AND QUANTIFIED SYSTEM PERFORMANCE-WORKLOAD RATING SCALES

Thomas H. Higgins

March 1981
Final Report

Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Systems Research & Development Service
Washington, D.C. 20590

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-RD-81-30	AD-A099 196	

4. Title and Subtitle	5. Report Date
A SYSTEMS ENGINEERING EVALUATION METHOD FOR PILOTED AIRCRAFT AND OTHER MAN-OPERATED VEHICLES AND MACHINES WITH HYPOTHETICAL EXAMPLE OF A SYSTEMS EVALUATION AND QUANTIFIED SYSTEM PERFORMANCE-WORKLOAD RATING SCALES	March 1981
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Thomas H. Higgins	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington, D.C. 20591	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	
Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington, D.C. 20591	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

A system evaluation method is presented which systematizes and quantifies both PRP pilot rating procedures and ECP engineering calculation procedure measures of system performance on a logarithmic ratio basis of test aircraft configurations compared to a known selected standard aircraft (vehicle) configuration. The logarithmic units 10 log (ECPtest/ECPstd) and 10 log PRPtest/PRPstd) used in this system evaluation method are termed "decivals, dV" as they are 10 times the log base 10 of the ratio of the ECP and PRP values obtained during tests for the test aircraft configuration compared to the chosen standard aircraft configuration.

The system evaluation is for chosen time periods of selected flight operations which are critical to flight safety, such as may occur during takeoff, or approach to landing and may include emergency engine failure, flight control or instrument malfunction conditions. System equations are presented which answer the question as to how good is the test configuration in relation to the known standard configuration during these same flight conditions. Potential ECP measures are discussed and their correlation with PRP pilot ratings obtained during flight test or flight simulator test determines their retention as effective system performance and evaluation measures. The non-dimensional logarithmic nature of the retained ECP system performance descriptors allows their combination by logarithmic summation and their correlation with the PRP pilot ratings is determined. The combination of ECP measures having the highest correlation with pilot ratings is retained for final system evaluation.

17. Key Words	18. Distribution Statement
Systems Systems Evaluation Piloted Aircraft Man-Operated Vehicles Man-Operated Machines	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22151.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED	57	

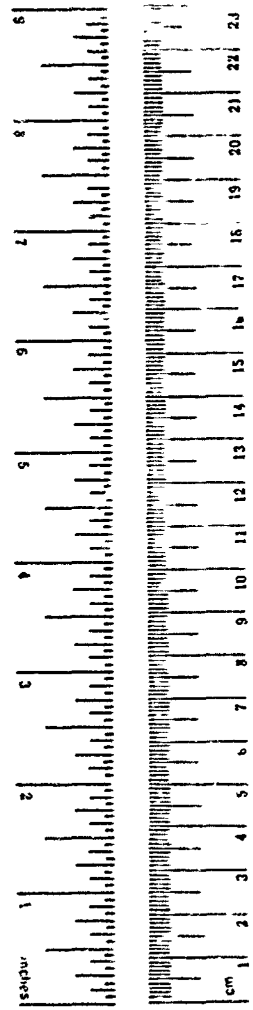
Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

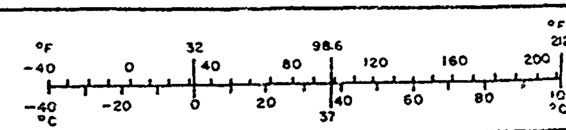
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



SUMMARY

A complete systems engineering method is presented which provides a missing frame of reference and a procedure including new measurement units which quantifies and simplifies the evaluation of piloted aircraft systems tested in flight or the systems evaluation of any man operated vehicle or machine tested in its operational environment.

New pilot rating procedures, PRP, and new engineering calculation procedures, ECP, are presented comparing the test aircraft configuration ratings and measures with a chosen standard aircraft configuration using non-dimensional logarithmic units termed decivals, dV.

A system evaluation equation is presented which quantifies and describes the relationship between the flight test pilot ratings and the measured flight test data. This method may be used for evaluation purposes during system design, test, certification of airworthiness and flight safety evaluations to determine the aircraft in flight man-machine pilot-aircraft (operator vehicle) operating system relationship:

$$\frac{PRP_{test}}{PRP_{std}} = k_1 \left(\frac{ECP_{test}}{ECP_{std}}\right)^n \qquad (1)$$

The answer to[1] how good the aircraft or vehicle is, regarding the operationally defined system evaluation goal, given an engineering calculation procedure ECP measure and[2] given a

pilot or operator rating procedure PRP rating what should the ECP measured value be is provided by the above equation in logarithmic form as follows:

$$10 \log\left(\frac{PRP_{test}}{PRP_{std}}\right) = k_2 + n \quad 10 \log\left(\frac{ECP_{test}}{ECP_{std}}\right) \qquad (2)$$

The effectiveness of the completed system evaluation may be determined by the product moment correlation coefficient, $r_{PRP\,.\,ECP}$ obtained and $r^2_{PRP\,.\,ECP}$ showing the amount of common variance between pilot ratings and flight test data.

INTRODUCTION

A system is defined as any machine or tool including all modes of transportation vehicles, or their roadways, aerospace-ways, and seaways, or signs and signals used by man to attain some goal. All systems include the effects of the changing man-made or natural environment in which they are used and operated in real life and real time.

There are therefore a great many extraordinarily different kinds of systems, all of these systems may be evaluated to determine exactly how useful they are in attaining the various goals that have been set for them. It can be seen that this depends on the experience or ability of the user or operator and the characteristics of the machine or tool or vehicle or roadway or sign or signal and the changing man-made and natural environment in which it is used or operated in real life and real time.

There is therefore a great need for an objective and reliable method of evaluating systems. There is at the present time no complete unifying method which provides a frame of reference, and which defines, quantifies and describes the relationship between the criteria for evaluation and the observable and measureable data regarding the system performance in attaining its set goal.

Today there are many ways to describe and evaluate the separate and component parts of the system. For example there are mathematical models of the human operator behavior which are summarized and explained in Reference 1 and in the numerous references cited therein.

There are many excellent rating scales such as the Cooper-Harper pilot rating of aircraft handling qualities described in Reference 2 and its references lead the reader to other methods available for obtaining pilot ratings[2].

A summary of the definition and measurement of perceptual and mental workload is contained in Reference 3 which points out that the common method of assessing handling qualities relies on the subjective opinion of experienced test pilots[3].

Reference 1 emphasizes that these subjective pilot opinions have proved to be quite reliable but that they are sometimes difficult to use for design purposes. It also stresses that the ultimate goal in building systems engineering models of operator behavior is to use these models in the evaluation of aircraft handling qualities essential for flight operations. Of utmost importance is the acknowledgement that these handling qualities are vital for flight safety[1].

But there is a severe limitation pointed out in Reference 1 regarding these engineering models which have been postulated to date, quote:

"A rather severe limitation of both engineering models lies in the fact that they lack a direct connection with the pilot's subjective opinion of the workload involved in a given control situation. In the end it is this expert opinion, which decides on the acceptability, or otherwise, of an aspect of the aircraft's handling qualities. There exists, as yet, no generally applicable method to derive the pilot's workload from the parameters of the engineering model. There are serious doubts whether such a method will ever be developed.

An explanation of this very real shortcoming of most existing engineering models is considered to lie in the fact that such models portray primarily the control aspect of the human operators activities, thereby ignoring or bypassing the equally important mental activities of data processing and decision making[4]. It seems, therefore, that there is a need for new developments, explicitly combining into a single mathematical model the internal mental processes of data handling and decision making going on in the human brain, with the more overt control activities previously discussed.

It turns out to be a truly interdisciplinary study for which an accepted frame of reference has not yet been established, and iteration between rather general hypothesis, test of methods, and detailed analysis is necessary."

This report is intended to present ideas which provide the missing frame of reference for the total system description, analysis and evaluation. This is accomplished by quantifying the component parts of a system as defined in the first paragraph of the introduction and the degree of relationship between the test user, operator, pilot, driver ratings, and the measured test system input and output performance data.

This method allows the system evaluator to first operationally define his system evaluation goals and then conduct a systems evaluation which determines how good are the measured test data as well as how good are the test user, operator, pilot driver ratings. Of utmost importance is the fact that the method allows the evaluator to determine by quantifying it exactly how good is the final system evaluation.

PROCEDURE

The criteria presented for determining the effectiveness of a completed piloted aircraft system evaluation is the degree of correlation of flight test pilot ratings PRP with engineering calculations procedure ECP measures $r_{PRP \cdot ECP}$. The higher the product-moment correlation $r_{PRP \cdot ECP}$, the better the engineering calculating procedures ECP, selected are in predicting the combined or selected system characteristics being evaluated. In like manner, the higher the correlation the more confidence that may be placed in the validity and reliability of the flight test pilot (operator) ratings.

The new flight test pilot rating procedure, PRP, presented is based on the average of the logarithms of the pilot (operator) magnitude estimation ratings of total mental and physical workload due to the handling qualities and performance characteristics of the aircraft (vehicle) during a time segment of a selected or required critical flight (road, sea, space) operation. This rating, PRP_{test} of the test configuration is reference a given standard aircraft configuration with known measured handling qualities and performance characteristics, i.e., measured and critical flight operation which is given a, PRP_{std} pilot mental and physical workload rating of 10. If the test aircraft configuration requires twice the amount of the pilots mental and physical workload, in comparison to the standard configuration, the pilot would rate it as having a PRP of 20. If one-half, the pilot would rate it as having a PRP_{test} of 5, etc. If less than one-half, ratings of 1 through 4 are available for describing the test configuration workload. In similar fashion ratings of 21 through 40 describe workload increasing from double to four times the selected standard workload.

The new pilot rating procedure is therefore stated mathematically as follows:

$$PRP = 10 \log (PRP_{test}/PRP_{std}). \text{ decivals, dV.} \quad (3)$$

This method provides a pilot rating continuum from zero or a rating of the test configuration as being identical with the standard configuration to an infinite number of ratings. Ratings in a positive direction indicate a doubling of the pilot-aircraft system workload for each three decival, dV increase. The negative direction indicates a halving of the pilot-aircraft system workload for each three decival, dV decrease.

The new engineering calculation procedures presented may be stated mathematically as follows:

$$ECP = 10 \log (ECP_{test}/ECP_{std}) \quad \text{decivals, dV.} \quad (4)$$

The logarithmic units, 10 log (ECP_{test}/ECP_{std}) and 10 log (PRP_{test}/PRP_{std}) used in this system evaluation method are termed "decivals, dV" as they are 10 times the log base 10 of the ratio of the selected ECP and PRP system evaluation units for the test aircraft configuration compared to the chosen standard aircraft configuration.

In this way a doubling or halving of the ECP or PRP of the test configuration ratio with the chosen standard configuration results in a ±3 decival, 3 dV, increase or decrease as 10 times log 2 = 10 x 0.3 = 3 dV and 10 log (1/2) = 10 x -0.3 = -3 dV.

Each ECP is selected to quantify a known outstanding critical characteristic of the operating system such as integration of the pound-seconds of force exerted by the pilot on the flight controls about the pitch, roll, and yaw axis of the aircraft during the time segment of a selected flight operation which may later be shown to correlate highly with the pilots ratings of physical workload. In like manner, the integration of the number of bits of information transmitted per second by the primary flight instruments during the time segment of the selected flight operation may later be shown to correlate with the pilots rating of mental workload.

Each ECP is therefore a non-dimensional logarithmic ratio of the selected critical system characteristic ECP_{test} to the same critical system characteristic of the selected standard aircraft configuration ECP_{std}. This is a very important part of this new system evaluation procedure as it provides a means of converting all critical systems characteristics to non-dimensional logarithmic ratios of selected critical test configuration characteristics compared to the same critical standard configuration characteristics under identical test conditions. Note well that this non-dimensional nature of all the system ECPs permits the combination of all selected critical system ECPs into a single integrated ECP for the time segment of the selected flight operation. This is accomplished by summing them logarithmically:

$$ECP = 10 \log (\text{antilog } ECP_1/10 + \text{antilog } ECP_2/10 \ldots + \text{antilog } ECP_n/10) \quad (5)$$

It is important to note that pilot ratings PRP are also non-dimensional as they are 10 times the log of the ratio of the test aircraft configuration to the selected standard configuration, PRP_{test}/PRP_{std}. Pilot ratings obtained during the same time segment of the selected flight operation for both the test and the standard aircraft configurations, therefore, provide a summary PRP test pilot rating which corresponds to the summary ECP flight test data calculations. It is the degree of correlation $r_{PRP \cdot ECP}$ of these values obtained from a suitable number of pilots which determines how effective the selected system evaluation ECP measures are in describing the characteristics of the operating system as rated by the test pilots.

The test pilot ratings obtained during the chosen time period of the selected flight operation are the accepted criteria for determining the actual real-life, real-time characteristics of the test aircraft in its flight environment.

As the PRP and ECP measures are a log X log Y relationship, it can be assumed and proven subsequently by inspection that their relationship is rectilinear and that a straight regression line equation established by the method of least squares through the plotted values obtained in each test flight (or flight similator test) provides the answer to how good is the aircraft given an ECP:

$$10 \log \left(PRP_{test}/PRP_{std} \right) = k + n\ 10 \log \left(ECP_{test}/ECP_{std} \right) \quad (6)$$

The number and type of ECP measures chosen for an evaluation may vary from one measureable characteristic of the piloted-aircraft system made during the time period of the selected flight operation to several which may be examined individually, in pairs and all possible combinations. The effectiveness of any individual ECP measureable piloted-aircraft system characteristic may be determined by its degree of correlation with the PRP flight test pilot ratings obtained. Those with low correlations may be rejected while those which correlate highly may be retained for further system evaluation.

The goal is to obtain through flight and simulator test experience those ECP measures which provide the highest correlation with the flight test pilot ratings.

There are two significant types of ECP measures available for use by the systems engineer. One is a peak measure which is the maximum value obtained by integration over the complete time period of the values sampled approximately every 0.5 or 1.0 second.

Previous psychoacoustic test experience has shown that the time integrated measures of cockpit sound pressure and frequency provide higher correlation with pilot ratings of cockpit perceived sound level than maximum peak measures.[5] The pilot may, therefore, be asked to rate and the engineer measure the total changing situation over a chosen time period of a selected critical flight operation as well as the peak value occuring during that time period.

Some of the possible measureable ECP characteristics of the piloted-aircraft system made during the chosen time period of the selected flight operation include pilot-aircraft flight control induced physical workload; pilot-aircraft primary flight instrument display induced mental workload; total in flight system pilot-aircraft induced mental and physical workload; pilot-ATC system-aircraft navigation communication display and control induced mental and physical workload; pilot-aircraft flight operational environment induced workload (turbulence, wind shear, visibility, IFR, VFR, Noise, Vibration, Cockpit lighting) and pilot-aircraft induced workload due to the aircraft handling quality-performance descriptors such as: dutch-roll, adverse yaw, longitudinal stability, lift over drag ratios L/D in different flap, gear, slot, trim configurations, etc., thrust to weight ratios T/W, etc. As they ultimately become available with the operation of all systems, the aircraft accident, flight safety statistics over time for the various critical flight operations such as takeoff and approach to landing, etc., will provide the most significant potential ECP.

The degree of correlation of any of the above ECP measures with the PRP pilot ratings obtained during flight tests or flight simulator tests for a time period of a selected flight operation determines their suitability for retention and consideration in the final ECP measures included in the final system evaluation.

All PRP pilot ratings and ECP calculation procedures are logarithmic ratios between the test aircraft configuration and the chosen standard of reference for all ratings and measures, a given standard aircraft configuration, during a time period of a selected flight operation flown under identical or as constant as practicable test conditions.

All flight test pilots will have experience and sufficient training in the selected standard aircraft configuration during the time period of the selected flight operation and constant flight test conditions to assure high reliability and confidence in their ratings.

All system evaluation ECP and PRP measures are calculated in logarithmic form and so may be integrated over time for levels sampled at discrete time intervals of for example (0.5) one-half second. The working expression becomes:

$$\Sigma PRP = 10 \log \left[\sum_{k=0}^{k=\frac{d}{\Delta t}} 10^{PRP_{(k)}/10} \right] + 10 \log \Delta t \qquad (7)$$

$$\Sigma ECP = 10 \log \left[\sum_{k=0}^{k=\frac{d}{\Delta t}} 10^{ECP_{(k)}/10} \right] + 10 \log \Delta t \qquad (8)$$

where d is the time duration in seconds during which ECP (PRP) is measured and Δt is the time interval between the ECP samples. Note well that $ECP_{(k)}$ is as defined by Equation Number 4, i.e., $10 \log (ECP_{test}/ECP_{std})$.

The psychological set of the flight test pilots, that is the operational definition of the system evaluation goals determines the outcome of the evaluation. The psychological set of the pilots is based on the operational definitions used in the system evaluation which define precisely what is being evaluated. The set adopted should be maintained throughout for rating both the standard and the test aircraft configurations. The possible sets include but are not limited to the following:

(1) The whole or total operationally defined pilot-aircraft system workload induced by the aircraft configuration during a chosen time period of a selected flight operation such as an ILS or MLS approach to Dulles International Airport.

(2) The operationally defined physical workload only associated with movement of the flight controls and integration over time of the pound-seconds of force about the pitch, roll and yaw axis.

(3) The operationally defined mental workload only associated with the primary flight instruments and integration over time of the bits of information transmitted per second by these instruments.

(4) The operationally defined mental and physical workload of (3) and (4) above; combined in decival, dV system evaluation logarithmic units.

(5) Other psychological sets operationally defined according to the system evaluation goals. For example: System evaluation of any subsystem change such as: new active flight controls; new primary flight instruments such as electronic altitude director indicator changes; ATC Communications-Navigation changes; new collision avoidance or data link changes, etc., which change pilot mental and physical maximum peak workload or workload over time.

The degree of effectiveness of the final system evaluation resulting from the selection of appropriate system engineering calculation procedure ECP measures in evaluating the piloted aircraft or any human operated system's performance can be determined by the ECP correlation with the pilot ratings, i.e. by the value of the positive or negative product moment correlation coefficient $r_{PRP \cdot ECP}$ as follows:

TABLE 1. INTERPRETATION OF SYSTEM EVALUATION RESULTS

$r_{PRP \cdot ECP}$	SYSTEM EVALUATION RESULTS	$r^2_{PRP \cdot ECP}$*
0.9 $\pm$ to 1.0	Very High	81 to 100%
0.78 to 0.89	High	61 to 80%
0.64 to 0.77	Moderate	41 to 60%
0.46 to 0.63	Low	21 to 40%
0.00 to 0.45	Very Low	0 to 20%

*The amount of variance in test pilot ratings PRP that is common to or that is in agreement with the variance in the chosen engineering calculation procedure ECP test data.

A detailed explanation of the product-moment correlation coefficient and the interpretation of correlation upon which the above Table 1 is based is contained in Reference 6, Pages 79-106.

An example of system evaluation results and the conclusions that may be drawn are presented in Figure 1.

An example of the major steps to be taken in conducting a systems evaluation is contained in Appendix B.

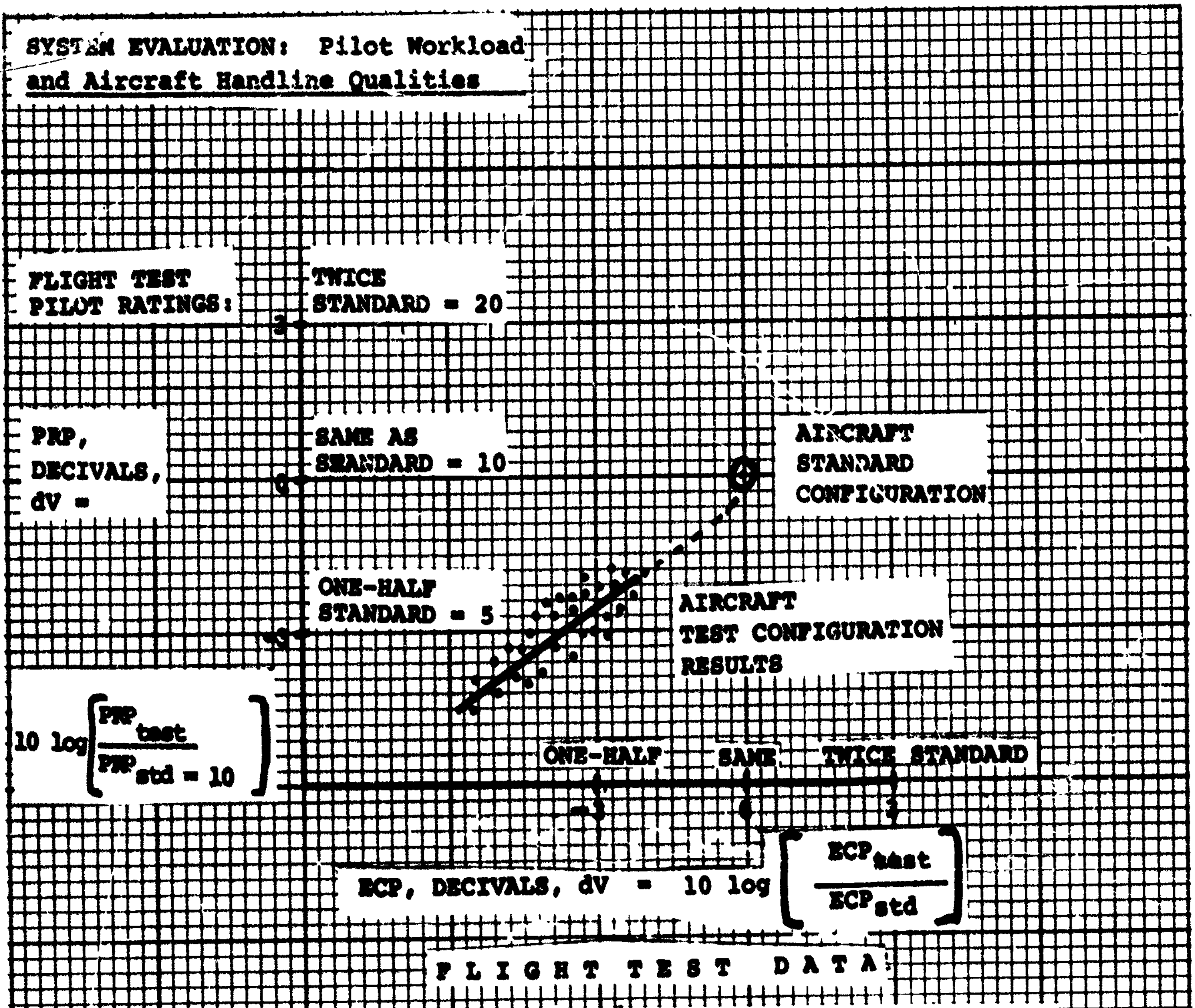


FIGURE 1. AN EXAMPLE OF SYSTEM EVALUATION RESULTS

Figure 1 is an example of system evaluation results and the following is an example of conclusions which may be drawn:

The expert flight test pilot ratings, PRP correlated highly, $r_{PRP \cdot ECP} = 0.90$ with the measured flight data ECP.

This result shows that over 80% of the variance in pilot ratings PRP is associated with the variance in the flight test data ECP which indicates a very high qualitative system evaluation. The system evaluation places the test configuration as being twice as good or requiring only one-half the pilot-aircraft system workload as the comparison standard configuration during instrument landing approaches under identical flight conditions.

DISCUSSION

The chosen or 'given' standard aircraft configuration is selected for comparison purposes of both the pilot or operators ratings and the flight test measures obtained for the test configuration. If the comparison or standard aircraft configuration is one with which the test pilots are already familiar and readily available for flights, this simplifies both the choice of a standard and the problem of learning its performance characteristics.

The standard configuration is not presented as anything other than a baseline for comparison with the test aircraft configuration for both pilot rating PRP and engineering calculation procedure ECP purposes. It provides the potential means by which various flight test data as different as apples and oranges are turned into non-dimensional logarithmic ratio units termed decivals dV and thereby these at first seemingly different types of measures may be combined by logarithmic summation.

This provides the opportunity to ultimately discover by repeated tests and exercise of this system evaluation method the most important ECP measures. These measures will be the sum of the flight test measured data parts which make up the whole pilot rating of the aircraft performance. These most important flight test measures when combined have the highest correlation with the test pilot ratings.

Pilot ratings obtained during flight tests may be ratings of maximum peak workload which occurs during the selected flight segment such as an instrument landing system ILS approach. It may also be necessary to obtain the pilot ratings of workload obtained at different times on different flight segments. In this case a pilot rating of total workload integrated over time may be computed by using Equation 7.

The correlation of the PRP ratings with Cooper-Harper pilot ratings may also be determined. It will be possible to then compare the Cooper-Harper ratings obtained on the comparison or standard aircraft with those for the test aircraft. Using the logarithm of the Cooper-Harper ratings will permit the checking of the flight test results and the possibility that the present 1 through 10 rating scale is somewhat logarithmic. A regression line established by the method of least squares will quantify the log linear relationship between the two scales. Preliminary paper and pencil tests (7) show that their is a logarithmic linear relationship between the pilots or rater's judgement of workload doubling and the Cooper-Harper scale especially in the mid-scale 4, 5, and 6 ratings. There is presently not enough workload described by ratings near the good end (present ratings of 1, 2, and 3) and possibly too much at the bad end (7, 8, 9, and 10) to meet a perfect logarithmic linear relationship with the raters judgement of the rate at which workload is doubling as presently described by the 1 through 10 descriptive language associated with each numerical rating. (Test (7) results are contained in Appendix A.)

This could be achieved by psychometric scaling of the descriptive language and should be considered in future modifications of these types of pilot ratings which use descriptive language. The descriptive language in each of the rating choices 1 through 10 would be 0.1 one-tenth logarithmic distance, i.e. relative rate of increasing workload as follows:

log R:	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
R : (Relative Rate of Increasing Workload)	1.	1.26	1.58	2.	2.5	3.16	4.0	5.0	6.31	8.00	10.0

Note that in this scale the workload doubles every third descriptive rating of workload and aircraft handling qualities.

It is interesting to note that when workload has increased 10 times over some minimum or ideal workload situation that it is a disaster and the aircraft is uncontrollable. Workload must be kept at levels ideally below the doubling and certainly the quadrupling of the minimum or ideal workload level where R = 1.0 and log R = 0.

This emphasizes the need to increase the number of descriptive ratings presently available to the test pilots at the good or low workload end of the scale and the development of a truly logarithmic descriptive language scaling of workload and handling qualities of aircraft.

The ideas regarding system evaluation and pilot ratings presented here are intended to supplement rather than replace the present Cooper-Harper system of rating aircraft and both types of pilot ratings would be obtained as well as other methods not discussed herein which then may be compared.

In fact to test this method, the following may be carried out not only for future tests but in the analysis of past recorded flight test results which include both pilot ratings and measures which are available for two or more aircraft. The decival values would be computed for the pilot ratings and flight test measures using one aircraft chosen arbitrarily as the comparison or standard and the other as the test aircraft. The decival value of $10 \log \left[\frac{C\text{-}H_{test}}{C\text{-}H_{standard}}\right]$ is determined for each test pilot and compared with the decival value of $10 \log \left[\frac{ECP_{test}}{ECP_{standard}}\right]$ for appropriate measured flight test data. The degree of correlation of the decival pilot ratings with the decival flight test measures will aid in the choice of future flight test measurements and procedures.

Where changing environmental variables are of concern as to their effects on the system including pilot or operator workload and the handling qualities of the aircraft or vehicle then the test variable becomes the logarithmic ratio of an environmental change with time compared to a chosen standard condition.

The decival value of the environmental change ENV = $10 \log \left[\frac{ENV_{test}}{ENV_{std}}\right]$ may be computed and compared with the PRP decival pilot ratings by computing their correlation coefficient, $r_{PRP \cdot ENV}$, obtained for these same environmental test and standard conditions. The environmental conditions may be man made or natural for example: changing airport, runway, approach lighting methods, or the amount of turbulence or wind-shear for repeated ILS approaches by the same piloted-aircraft. For road vehicles, one may test the differences in roadways and strive to learn from driver ratings and handling quality measures why there are more accidents for left-hand highway curves compared to right-hand curves. Other examples of conditions which might be tested in this way include the effects of different kind of roadway signs and signal methods as well as the effects of weather such as rain, snow, ice on the road safety of the national highway system.

Regarding future system models, it is interesting to note that the relationship between pilot response and system output measures having high correlation with each other based on the proposed system evaluation method may be used in developing and improving system models.

In the crossover, optimum control and other system models pilot response is equal to $10 \log \left[\frac{PRP_{test}}{PRP_{std}}\right]$ decivals and the system output is expressed as $10 \log \left[\frac{ECP_{test}}{ECP_{std}}\right]$ decivals. The nature of the units of pilot response and system output flowing through the system model are non-dimensional logarithmic ratios of pilot response and system output of the system being modeled, i.e., the test system compared to a "given" or chosen standard system. This represents how the pilot or operator responds to the displays of system output in the feedback loop. He is comparing how the new or test aircraft, vehicle, machine or tool responds compared to a known or learned standard system with which he has prior experience. This is suggestive of an approach to also modeling how the operator learns based on his experience with one system compared to another.

The degree of proficiency of the operators of systems being evaluated is usually a problem. In the present method the test subjects may have varying degrees of proficiency as they will be rating their workload on a test system relative to a standard system. Their ratings are, therefore, relative to the same degree of proficiency used to perform the selected operation evaluated for both the standard system and the test system. All test operator ratings and test measures are therefore equally useful for system evaluation purposes as long as the degree of proficiency is such that the test operators can perform the selected operations being evaluated from beginning to end on the test and standard systems.

Another advantage of this system evaluation method is that it can be used to assure that new evolving systems are improvements over previous systems as to their handling qualities and operator workload. These are factors which have a direct bearing on safety. The improvement or departure of new systems from old systems with known records of safety can easily be established early in the design stage. Continued use of this method will identify those engineering calculation procedures or measures which have correlated the best with past system performance ratings. These will be the best predictors of the future systems performance. Later in the design process the method may be used in simulation tests and continued with tests of the prototype and first operational vehicles.

AN EXAMPLE OF A SYSTEM EVALUATION

The following example is one made-up by the author to demonstrate the system evaluation procedure more clearly than mere words. Each step is outlined and pilot ratings and measured data were selected to demonstrate an evaluation in which there is a valid relationship between the pilot ratings and the test data, pound-seconds of force. This is an example only and not to be interpreted as a valid real-world relationship.

It is interesting to note that the first run of pilot ratings versus the test data yielded a -0.39 correlation. This demonstrated that the first selection of paired ratings and test data was indeed a random sample and there was no significant correlation coefficient or relationship between them. This may be the outcome of a real-world test of this method. But it is equally as important to determine that there is not a relationship between pilot ratings of what is being evaluated and the selected test data. It eliminates that measured system variable and the investigator must search for those measured system variables which do correlate with the pilots ratings. As these significant relationships are determined they may be used by everyone and system evaluation will become increasingly simpler and more realistic than it is today.

AN EXAMPLE OF A SYSTEM ENGINEERING EVALUATION

The systems engineering evaluation is of the physical workload/handling qualities of a simulator versus the physical workload/handling qualities of the aircraft simulated.

Background -- Because of undesirable pilot induced oscillations found in the first simulator model the control forces required in the simulator during an approach were increased by the designer over those found in the aircraft. This evaluation is to determine first whether this difference in the simulator workload/handling qualities and the simulated aircraft is a significant variable in pilot transition training which is detected by the pilots. Of equal importance is to determine if pound-seconds of force is a useful measure which correlates highly with the pilots ratings and which therefore, may be used as a prediction in future aircraft design and evaluation by the Federal Aviation Administration (FAA) and the Industry.

Procedure --

Step 1: The first step is to define operationally what is being evaluated which sets the system engineering evaluation goals. System engineering evaluation goal: Do the control forces in the simulator affect pilot transition training to the simulated aircraft? Specifically are the increased control forces, measured in pound-seconds integrated over time, required in the simulator to avoid pilot induced oscillations during a standard instrument approach noticed and rated as such by the pilots after they have successfully transitioned to the aircraft and obtained actual flight experience? Operationally defined the question is as follows: What are the pilot ratings of physical workload in the simulator during a standard instrument approach compared to a given physical workload rating of 10 in the aircraft? And most important: Is there a correlation between the chosen measure, total pound-seconds of force integrated over time, and the pilot ratings of physical workload?

Step 2: Select the aircraft standard configuration--This is the real aircraft configuration in flight during a standard instrument approach. The standard is selected for comparison purposes only. It is the configuration which has been incorporated in the simulator and it is the one which the pilots are most familiar with and obtained flight experienced in and is readily available for flight.

Step 3: Select the flight test engineering calculation procedure ECP, measures that will be recorded: The pound-seconds of force required about the pitch, roll and yaw axes will be integrated over time to obtain objective measures of the forces actually required: (1) To fly the selected aircraft standard configuration in actual flight and (2) in the simulator from the outer marker to touchdown and roll-out during a standard instrument approach to the local airport.

Step 4: Train the pilots so that they understand what is being rated as defined in Step 1: What are the magnitude estimation pilot ratings of physical workload in the simulator during a standard instrument approach compared to a given physical workload rating of ten in the aircraft standard configuration selected in Step 2 and flown under similar conditions on a standard instrument approach? Pilots are instructed to rate the simulator physical workload as five if one-half that of the aircraft physical workload and 20 if doubled with any rating from zero to infinity being available to quantify their rating of simulator physical workload in relation to the selected standard aircraft configuration physical workload given a rating of ten.

Step 5: Ten pilots fly the selected standard aircraft configuration and learn during three instrument approaches that the operationally defined physical workload required to fly the standard instrument approach from the outer marker to touchdown and roll out is given a magnitude estimation value of ten.

Step 6: The same ten pilots fly three identical instrument approaches in the simulator which is being evaluated. Each pilot provides his magnitude estimation judgement of the physical workload required to fly the flight simulator with its known simulated aircraft characteristics from the outer marker to touchdown and rollout compared to the physical workload required to fly the actual aircraft under identical conditions given a physical workload rating of ten. If in the pilot's judgement the physical workload in the simulator is double that required by him to fly the actual approaches his rating is 20. If one-half his rating is five. Note well: The pilot may use any rating from zero to infinity to describe and most important quantify their ratings on a continuous scale which is anchored to a known and quantified standard physical workload.

Step 7: Record the selected test data during the actual flights and the simulated approaches so that they may be quantified (ECP) and compared to the pilot ratings (PRP) and their exact relationship to each other determined, i.e., λ_{PRP} . ECP will equal some number between - 1.0 and 1.0 which quantifies the degree of relationship to each other determined between the selected and recorded test data (ECP) and the pilot judgements (PRP).

The selected test data recorded on all approaches in the aircraft and simulator are the total pound-seconds of force, measured by electrical transducers, about the pitch, roll and yaw axes integrated over time resulting from the pilots control movements from the outer marker to touchdown and rollout.

Step 8: Compute for each flight and each simulated approach the total pound-seconds of force measured about the three axes from the outer marker to touchdown and rollout. Determine (1) the total pound-seconds of force for each pilot during the three actual flights and (2) the total for each pilot for the three flights in the simulator, (3) the average total pound-seconds of force for each pilot and the ratios of the force required in the simulator compared to that required in the aircraft.

5

$$\left[\frac{\text{Force (pound-seconds)}_{\text{Simulator}}}{\text{Force (pound-seconds)}_{\text{Aircraft}}}\right]$$

(4) Calculate the logarithm of this ratio and multiply by ten to obtain the ECP value.

$$ECP = 10 \log_{10} \left[\frac{\text{Force (pound-seconds)}_{\text{Simulator}}}{\text{Force (pound-seconds)}_{\text{Aircraft}}}\right]$$

The ECP is the Engineering Calculation Procedure used to obtain a logarithmic continuous non-dimensional measure. (See example, Table 1)

Step 9: In a similar fashion to Step 8 (1) Record for each simulator test flight, the pilot's magnitude estimation of the physical workload required to fly the simulator (2) determine for each pilot the ratio of these magnitude estimation ratings compared to the aircraft's given magnitude estimation value of 10.

$$\left[\frac{\text{Magnitude Estimation Rating (0 to } \infty)_{\text{Simulator}}}{\text{Given Magnitude Estimation Rating (10)}_{\text{Aircraft}}}\right]$$

(3) Calculate the logarithm of this ratio and multiply by 10. This is the pilot rating, PRP value entered in the last column of Table 2 for each pilot A through J.

$$PRP = 10 \log_{10} (\text{M.E. Rating}_{\text{Simulator}} / \text{Given M.E. Rating Aircraft})$$

6

AN EXAMPLE OF SYSTEM ENGINEERING EVALUATION TEST DATA AND RESULTS

Engineering Calculation Procedures, ECP -- The following data continued in Table 1 are the measured test data obtained from electrical transducers which measured the total pound-seconds of force required about the pitch, roll and yaw axes recorded during simulator and aircraft flights from the outer marker to touchdown and roll out.

Table 1 shows how the measured data are averaged in both the simulator and aircraft and their ratios obtained, simulator: aircraft, and converted into 10 log form to obtain ECP values.

TABLE 1. PHYSICAL WORKLOAD - MEASURED TEST DATA

	TOTAL POUND-SECONDS OF FORCE									
	SIMULATOR FLIGHTS				AIRCRAFT FLIGHTS				AVERAGE RATIO	10 (LOG AVERAGE RATIO)
PILOTS	1	2	3	AVERAGE	1	2	3	AVERAGE	SIMULATOR: AIRCRAFT	ECP
A	800	775	825	800	425	400	375	400	2:1	3.01
B	950	900	850	900	450	525	500	500	1.8:1	2.553
C	750	800	850	800	350	375	400	375	2.133:1	3.29
D	825	800	775	800	425	375	400	400	2:1	3.01
E	850	900	950	900	500	525	550	525	1.714:1	2.34
F	750	850	800	850	375	400	350	375	2.267:1	3.554
G	775	825	800	825	400	425	375	400	2.063:1	3.144
H	900	950	850	900	400	450	500	450	2:1	3.01
I	800	750	850	800	350	375	400	375	2.133:1	3.29
J	825	800	775	800	400	425	450	425	1.88:1	2.75

NOTE: The measured data are obtained and calculated individually for each of the ten pilots A through J.

7

Pilot Rating Procedures, PRP -- The following data contained in Table 2. are the pilot magnitude estimation ratings of pilot physical workload compared to aircraft physical workload given a magnitude estimation rating of ten. Ratings are obtained from the same ten pilots listed in Table 1 as pilots A through J.

The ratio of the average magnitude estimation of physical workload in the simulator for each pilot compared to the magnitude estimation of physical workload in the aircraft given a value of ten is calculated for each pilot. Ten times the logarithm of this average ratio [10 log (average ratio)] gives the pilot rating procedure, PRP, value for each pilot A through J.

TABLE 2. PILOT MAGNITUDE ESTIMATION RATINGS OF SIMULATOR PHYSICAL WORKLOAD COMPARED TO AIRCRAFT WORKLOAD GIVEN A MAGNITUDE ESTIMATION RATING OF TEN

PILOTS	SIMULATOR FLIGHTS				RATIO OF AVERAGE ME RATING SIMULATOR: GIVEN ME RATING AIRCRAFT = 10	10 Log (AVERAGE RATIO) PRP
	1	2	3	AVERAGE		
A	15	18	21	18	1.8	2.553
B	16	17	18	17	1.7	2.3
C	22	18	24	22	2.2	3.424
D	15	18	21	18	1.8	2.553
E	15	19	17	17	1.7	2.3
F	24	20	22	22	2.2	3.424
G	16	21	26	21	2.1	3.22
H	23	20	17	20	2.0	3.01
I	26	21	16	21	2.1	3.22
J	18	17	16	16	1.7	2.3

An alternative quantified descriptive system performance/operator, pilot or controller workload (total mental and physical effort) rating procedure is presented in Appendix C.

8

The relationship or correlation between pilot ratings, PRP and the engineering measures, ECP, may now be obtained. From this quantified relationship, two equations based on the least squares fit of the PRP and ECP data may be obtained so as to predict one of these measures based on the other.

The magnitude of the relationship may be compared to a table of correlation coefficients, such as Table D in Reference 6. According to the number of degrees of freedom, N pairs-1 = nine in the present case, correlations larger than 0.602 and 0.735 could occur by chance and chance alone only five and one time in a hundred respectively.

Table 3 presents an example of this process.

TABLE 3. THE TEST OF RELATIONSHIP OF MEASURED TEST DATA, ECP WITH PILOT RATINGS, PRP

NUMBER OF AVERAGED PAIRS	PRP	ECP	RESULTS
1	2.553	3.01	$r_{PRP \cdot ECP} = 0.89$*
2	2.3	2.553	The correlation of pilot ratings with measured flight test data.
3	3.424	3.29	Eq.1. $PRP = 0.647 + 1.161\ (ECP)$
4	2.553	3.01	Used to predict pilot ratings, PRP, based on measured flight test data, ECP.
5	2.3	2.34	Eq.2. $ECP = 1.07 + 0.68\ (PRP)$
6	3.424	3.554	Use to predict measured flight test data, ECP based on pilot ratings, PRP.
7	3.22	3.144	$r^2 = 0.89^2 = 80\%$
8	3.01	3.01	Common variance in ratings and flight test data.
9	3.22	3.29	
10	2.3	2.75	* Statistical tables such as Table D of Reference 6 show that for 10 pairs of data the correlation must be greater than 0.735 to be 99 percent confident that the correlation is real and did not occur by chance and chance alone. Because of the 0.89 correlation the system evaluation is valid and Equation 1 will be useful to predict pilot ratings of physical workload based on measured or calculated total pound-seconds of force values.

CONCLUSIONS

A system evaluation method is presented which systematizes and quantifies both PRP pilot rating procedures and ECP engineering calculation procedure measures of system performance on a logarithmic ratio basis of test aircraft configurations compared to a known selected standard aircraft (vehicle) configuration. The logarithmic units 10 log (ECP_{test}/ECP_{std}) and 10 log (PRP_{test}/PRP_{std}) used in this system evaluation method are termed "decivals, dV" as they are 10 times the log base 10 of the ratio of the ECP and PRP values obtained during tests for the test aircraft configuration compared to the chosen standard aircraft configuration.

The system evaluation is for chosen time periods of selected flight operations which are critical to flight safety, such as may occur during takeoff, or approach to landing and may include emergency engine failure, flight control or instrument malfunction conditions. System equations are presented which answer the question as to how good is the test configuration in relation to the known standard configuration during these same flight conditions. Potential ECP measures are discussed and their correlation with PRP pilot ratings obtained during flight test or flight simulator test determines their retention as effective system performance and evaluation measures. The non-dimensional logarithmic nature of the retained ECP system performance descriptors allows their combination by logarithmic summation and their correlation with the PRP pilot ratings is determined. The combination of ECP measures having the highest correlation with pilot ratings is retained for final system evaluation.

REFERENCES

1. Gerlack, O. H. "DEVELOPMENTS IN MATHEMATICAL MODELS OF HUMAN PILOT BEHAVIOR", The Aeronautical Journal, July 1977, pp. 293-305.

2. Cooper, C.E., Harper, R.P., "THE USE OF PILOT RATING IN THE EVALUATION OF AIRCRAFT HANDLING QUALITIES", NASA-TN-D-5153, April 1969.

3. Gerathewohl, S.J. "DEFINITION AND MEASUREMENT OF PERCEPTUAL AND MENTAL WORKLOAD IN AIRCREWS AND OPERATORS OF AIR FORCE WEAPON SYSTEMS: A STATUS REPORT IN HIGHER, MENTAL FUNCTIONING IN OPERATIONAL ENVIRONMENTS", AGARD-CP-181 Neuilly, 1975.

4. Shindon, T.B., Ferrell, W.B., "MAN-MACHINE SYSTEMS: INFORMATION, CONTROL, AND DECISION MODELS OF HUMAN PERFORMANCE", The MIT Press, Cambridge, Massachusetts, 1974.

5. Higgins, T.H., "A UNIFYING THEORY FOR DETERMINING HUMAN RESPONSE TO SOUND," Community Noise, ASTM STP 692, R.J. Peppin and C.W. Rodman, Eds., American Society for Testing and Materials, 1979, pp. 144-157.

REFERENCES

6. Edwards, Allen L., "STATISTICAL ANALYSIS FOR STUDENTS IN PSYCHOLOGY AND EDUCATION", pp.79-106, Rinehart and Company, Inc., 1946

7. Higgins, T. H., Unpublished test of pilots and non-pilots ratings of workload associated with the Cooper-Harper pilot rating scale.

APPENDIX

Two tests were conducted where 32 subjects, 11 pilots, and 21 non-pilots rated the increasing workload based on the Cooper-Harper descriptive language regarding aircraft characteristics, demands on the pilot, and adequacy of the aircraft for selected task or required operation. The required operation was limited to takeoff, landing or cruise under visual flight rules and non-turbulent weather.

In test one, the 32 subjects estimated the increasing workload for each Cooper-Harper pilot rating one through 10 reference a standard aircraft with a pilot rating of one, i.e., a C-H rating of one. This C-H rating number one was given a standard workload rating of 10. Where the aircraft characteristics and demands on the pilot are doubled the subjects gave that C-H number a rating of 20. When that workload doubled they entered 40 and 80 for the next doubling, etc.

In test two, conducted 60-days later, the 32 subjects rated the increasing and decreasing workload but Cooper-Harper pilot rating 6 was given a workload standard rating of 10. The test subjects rated as in test one but also where workload was one-half they entered 5 and one-half of that workload they entered 2.5, etc.

APPENDIX (Cont'd)

The hypothesis tested and the results of the tests were as follows:

C-H Rating:	1	2	3	4	5	6	7	8	9	10
Hypothesis Tested (Log Workload Ratio):	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Test 1 (Log Workload Ratio):	0 Standard	0.04	0.176	0.3	0.43	0.57	0.76	0.92	1.1	1.25
Test 2 (Log Workload Ratio):	-0.04	-0.16	0.11	0.3	0.462	0.57 Standard	0.84	1.0	1.28	1.55
Hypothesis Tested (Workload Ratio):	1	1.25	1.58	2	2.5	3.16	4.	5.	6.5	8.0
Test 1 (Workload Ratio):	1 Standard	1.1	1.5	2	2.7	3.7	5.7	8.3	12.7	18.0
Test 2 (Workload Ratio):	0.37	0.65	1.3	2	2.9	3.7 Standard	6.9	10.4	19.0	35.5

The rating of the pilots was not significantly different from the non-pilots.

APPENDIX B

Steps In Systems Evaluation

1. Set the system evaluation goals: Operationally define what is being rated. (See Page 14, Section (1) through Section (5) for examples.)
2. Select the aircraft standard configuration for comparison purposes only. It is one with which the raters are most familiar and is readily available for flight. (See Page 18 for discussion.)
3. Train the raters regarding what is being rated and what the system evaluation goal is using operational definitions and the magnitude estimation technique reference a standard workload (or what is being evaluated as defined in Step 1) of 10 for the comparison system. (See Page 14 for discussion.)
4. Select the flight test data ECP measures that will be monitored and recorded regarding test and standard system performance. (See Page 9 for discussion of ECP examples.)
5. Raters fly the selected standard aircraft on the chosen time segment of a required flight operation such as an ILS approach and learn that the operationally defined workload (or what is being evaluated as defined in Step 1) is equal to 10.
6. Raters fly the test aircraft on the same time segment of the chosen flight operation such as an ILS approach and obtain pilot ratings of the operationally defined workload (or

what is being rated) set in Step 1. Use magnitude estimation to judge how workload compares to the arbitrarily chosen but learned standard. If double, rate it 20, if one-half rate it 5, etc. (See Page 7 for discussion.)

7. Record selected flight test data, ECP measures, during the operation of both the test and standard aircraft for which pilot ratings, PRP magnitude estimates of workload, are obtained and a standard operationally defined workload (or what is being rated as set in Step 1) of 10 set for the standard aircraft.

8. Compute pilot magnitude estimates of workload:

   $$PRP = 10 \log \left[\frac{ME_{test}}{10_{std}}\right] \text{ in decivals, dV.}$$

9. Compute system flight test data measures:

   $$ECP = 10 \log \left[\frac{ECP_{test}}{ECP_{std}}\right] \text{ in decivals, dV for each selected}$$

   important system variable selected in Step 4.

10. Determine the quantified relationship ranging from zero to plus or minus 1, i.e. the product-moment correlation coefficient r PRP . ECP between all the pilot ratings determined in Step 8 with the corresponding flight test data measures selected in Step 9. (See Reference 6 for detailed explanation of the correlation coefficient.)

11. Retain those measures which correlate highly with the pilot ratings. Discard those which lack correlation. Combine by logarithmic summation those measures which correlated best with pilot ratings. (Independent ECP measures)

12. Compute the correlation between the combined ECP measures retained and pilot ratings where $\rho_{PRP \cdot ECP}$ equals ± 1.0 is perfect correlation and zero is none. (See Page 16 for discussion of interpretation of $\rho_{PRP \cdot ECP}$ and system evaluation results.)

This is the degree of effectiveness of the system evaluation and provides a check on the predictive value of the flight test data as well as the reliability of the pilot ratings, PRP. As more system evaluations using the method are completed there will be a growing identification of which flight test data measures, ECP, are the best predictors of the operationally defined system evaluation goals set in Step 1.

APPENDIX C

Quantified System Performance/Operator, Pilot, and Controller Workload Rating Scales

The addition of mathematical, operational definitions of workload and system performance language to quantify precisely and denote exactly the increasing and decreasing rate associated with the present (revised) Cooper-Harper (1)* rating scale language is suggested as a way to aid the test pilot in rating aircraft. This will also aid the aviation community, users, designers, systems analysts, scientists, engineers, etc., in better understanding of what the rating means.

It is suggested, based on tests and ideas contained in reference (2), that by definition, the first Cooper-Harper rating of 1.0 be assigned a workload (standard) ratio of 1.0. The workload, again by definition, associated with the Cooper-Harper rating of 4.0 is now assigned a workload ratio (compared to the standard) of 2.0 or a doubling of workload due to the language presently assigned to the Cooper-Harper rating. Again by definition, the workload associated with a Cooper-Harper rating of 7.0 is assigned a workload ratio (compared to the standard) of 4.0. This is a doubling of the workload associated with the change in language from a Cooper-Harper rating of 4.0 to one of 7.0. In similar fashion by definition the workload associated with the language of a Cooper-Harper rating of 10 is assigned a workload ratio (compared to the standard) of 8.0. This is a doubling of the workload associated with a change in the language from a Cooper-Harper rating of 7.0 to one of 10.

This procedure has by definition fixed the rate of doubling of workload/system performance every third block starting from the first Cooper-Harper rating of 1.0. This gives us at this point the following exact

*Denotes reference number

definition of workload and its rate of increase/decrease associated with the present language of the Cooper-Harper rating scale:

Cooper-Harper Rating:	1	2	3	4	5	6	7	8	9	10
Workload Ratio:	1.0			2.0			4.0			8.0
The Log Workload Ratio:	0			0.3			0.6			0.9

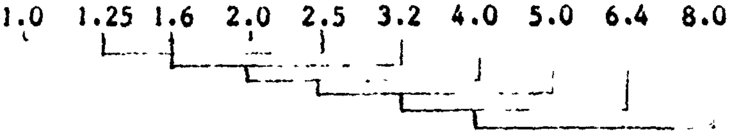
It only remains at this point to perceive that this defines and assigns quantified workload values to the present Cooper-Harper rating scale in an easily understood fashion, the doubling of workload every third rating. This by happy accident (serendipity) is a logarithmic scale and, therefore, the intervening blocks are also defined to be logarithmic which establishes for the complete Cooper-Harper rating scale the rate at which pilot mental and physical workload varies with the language associated with each rating.

The workload for the Cooper-Harper scale is, therefore, defined as follows:

Cooper-Harper Rating:	1	2	3	4	5	6	7	8	9	10
Workload Ratio:	1.0	1.25	1.6	2.0	2.5	3.2	4.0	5.0	6.4	8.0
The Log Workload Ratio:	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

The defined workload ratio associated with each rating is seen to double with every third rating as follows:

Cooper-Harper Rating:	1	2	3	4	5	6	7	8	9	10
Workload Ratio:	1.0	1.25	1.6	2.0	2.5	3.2	4.0	5.0	6.4	8.0



This quantified definition of workload, therefore, simplifies the rating task for the test pilot raters, i.e., the doubling or halving of workload in relation to the language contained in each rating block every third block. As a result, the increasing aircraft system evaluation rating scale is logarithmic. The distances between the ratings are now, by definition, equal to one-tenth (0.1) logarithmic units apart.

The workload ratios from rating block to rating block can, therefore, be precisely defined. The workload ratios are, of course, the antilogs or the numbers whose logarithm defines the amount of workload associated with each rating block.

This quantifying of the workload associated with each pilot rating not only simplifies the rating task for the pilot but perhaps most importantly helps to quantify and therefore simplify the analysis of the data associated with many pilots and many flights, etc.

Revised quantified rating forms for aircraft, air traffic control, and systems evaluation in general are contained in the Appendix.

Workload for systems in general is defined as the total work associated with all mental and physical tasks per unit time period required to successfully complete the selected task or required operation under a given system configuration and operating conditions.

In each case, definition of the required operation involves designation of the system phase and/or subphase with accompanying conditions.

The pilot, air traffic controller, or system operator ratings are arbitrarily called decivals (dV). A decival (dV) is equal to 10 $\log_{10}$ (quantified workload ratio). The decival is selected rather than decibel (dB) to clearly distinguish it from the reference level used in electrical engineering practice. The decival may also be used in system(s) evaluation. The reference level is defined by the language associated with the system in the first rating block which has a defined value of 1.0.

Workload Ratio:	1.0	1.25	1.6	2.0	2.5	3.2	4.0	5.0	6.4	8.0
Log	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
dV	0	1	2	3	4	5	6	7	8	9

DISCUSSION

The quantified system performance/pilot, controller, operator workload rating scale provides a continuum of possible ratings ranging from 0 to 9 decivals. This is a continuum ranging in terms of system performance/workload ratios from the baseline condition of 1/1 or 1 to 8/1, 8 times or 3 doublings of the defined baseline condition. The pilots, air traffic controllers, or operators of any system may make magnitude estimate judgements of the system performance/workload anchored not only at the baseline but at a total of 10 points by the language of reference (1) already in wide use to rate the handling qualities of aircraft.

The quantified continuous nature of the scale provides both the rater and the scientist, engineer, systems analyst, etc., with an infinite number of possible ratings. These ratings may not only be located at the 10 points defined by the language describing the system performance/workload but at any point in between these points.

For example, the rater and system analyst finally have a way of determining where a system is precisely which may be between the third and fourth rating choices in system performance/workload. The rater is not limited, for example, to a judgement of either 1.6 or 2.0 times the quantified workload. The rater may select these values which are defined by the language of reference (1) or may interpolate to any value in between to describe his exact magnitude estimate rating of the system performance/workload.

CONCLUSIONS

Three quantified continuous system performance/workload rating scales are presented for use in obtaining pilot, air traffic controller ratings of aircraft and air traffic control systems, and user/operator ratings of any system.

The addition of mathematical, operational definitions of system performance/workload language is suggested to quantify precisely the exact ratio associated with the present (revised) Cooper-Harper rating scale language as a way to aid test pilots, air traffic controllers, etc., in rating aircraft and air traffic control systems. This approach simplifies analyses and improves the understanding of the meaning of averages of pooled ratings of system operators/users.

REFERENCES

1. Cooper, George E. and Harper, Robert P. Jr., "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities." NASA TND 5153, April 1969.

2. Higgins, Thomas H., "A System Engineering Evaluation Method for Piloted Aircraft and Other Man-Operated Vehicles and Machines." Final Report No. FAA-78-78, August 15, 1978.

APPENDIX

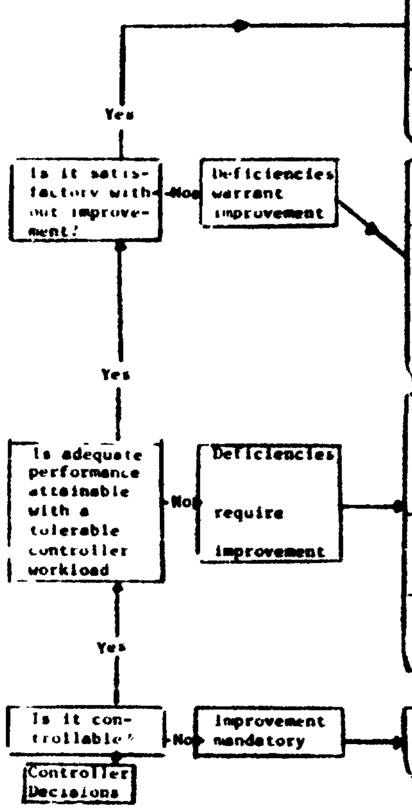
The Three Forms

1 ATC

2 Aircraft

3 General

Air Traffic Control System Engineering Evaluation, Decivel (dV), Rating Scale*****



Air Traffic Control System Characteristics	Demands on the Controller in Selected Task or Required Operation*	Quantified Workload Ratio**	Quantified Logarithmic Workload Ratio***	Controller Rating Decivels, dV****
Excellent Highly desirable	Controller compensation not a factor for desired performance	1.0	0	0
Good Negligible deficiencies	Controller compensation not a factor for desired performance	1.25	0.1	1
Fair - Some mildly unpleasant deficiencies	Minimal controller compensation required for desired performance	1.6	0.2	2
Minor but annoying deficiencies	Desired performance requires moderate controller compensation	2.0	0.3	3
Moderately objectionable deficiencies	Adequate performance requires considerable controller compensation	2.5	0.4	4
Very objectionable but tolerable deficiencies	Adequate performance requires extensive controller compensation	3.2	0.5	5
Major deficiencies	Adequate performance not attainable with maximum tolerable controller compensation. Controllability not in question	4.0	0.6	6
Major deficiencies	Considerable controller compensation is required for control	5.0	0.7	7
Major deficiencies	Intense controller compensation is required for control	6.4	0.8	8
Major deficiencies	Control will be lost during some portion of required operation	8.0	0.9	9

* Definition of required operation involves designation of the air traffic control system phase and/or subphases with accompanying conditions.

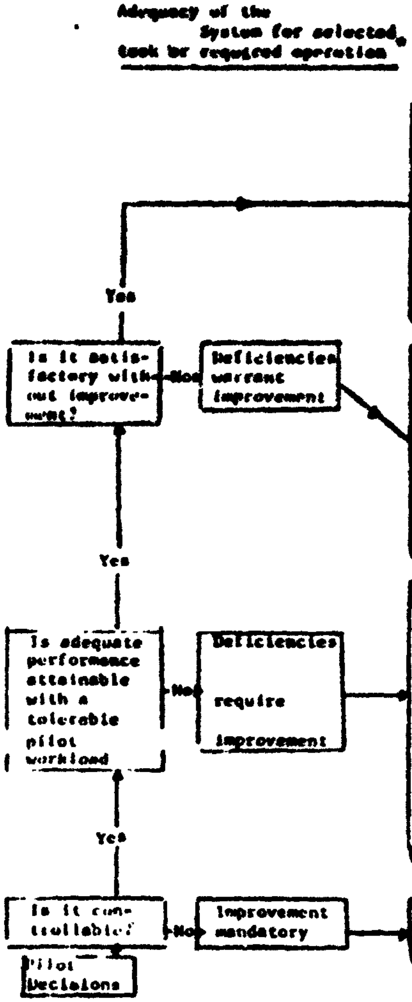
** Controller workload is defined as the total work associated with all mental and physical tasks per unit time period required to successfully complete the selected task or required operation.

*** For ease in controller rating and in understanding the meaning of each rating, the workload imposed by the ATC system configuration and control operations doubles by definition every third rating. As a result, the increasing system engineering evaluation scale is logarithmic.

**** A decivel (dV) is equal to 10 $\log_{10}$ (Quantified workload ratio).

***** Higgins, Thomas H. Draft Report No. November 1979.

Aircraft System Engineering Evaluation, Decivel (dV), Rating Scale*****



System Characteristics	Demands on the Pilot in Selected Task or Required Operation*	Quantified Workload Ratio**	Quantified Logarithmic Workload Ratio***	Pilot Rating Decivels, dV****
Excellent Highly desirable	Pilot compensation not a factor for desired performance	1.0	0	0
Good Negligible deficiencies	Pilot compensation not a factor for desired performance	1.25	0.1	1
Fair - Some mildly unpleasant deficiencies	Minimal pilot compensation required for desired performance	1.6	0.2	2
Minor but annoying deficiencies	Desired performance requires moderate pilot compensation	2.0	0.3	3
Moderately objectionable deficiencies	Adequate performance requires considerable pilot compensation	2.5	0.4	4
Very objectionable but tolerable deficiencies	Adequate performance requires extensive pilot compensation	3.2	0.5	5
Major deficiencies	Adequate performance not attainable with maximum tolerable pilot compensation. Controllability not in question	4.0	0.6	6
Major deficiencies	Considerable pilot compensation is required for control	5.0	0.7	7
Major deficiencies	Intense pilot compensation is required for control	6.4	0.8	8
Major deficiencies	Control will be lost during some portion of required operation	8.0	0.9	9

* Definition of required operation involves designation of the flight phase and/or subphases with accompanying conditions.

** Pilot workload is defined as the total work associated with all mental and physical tasks per unit time period required to successfully complete the selected task or required operation.

*** For ease in pilot rating and in understanding the meaning of each rating, the workload imposed by the aircraft configuration and control operations doubles by definition every third rating. As a result, the increasing system engineering evaluation scale is logarithmic.

**** A decivel (dV) is equal to 10 $\log_{10}$ (Quantified workload ratio).

***** Higgins, Thomas H. Draft Report No. November 1979.

System Engineering Evaluation, Decival (dV), Rating Scale****

Adequacy of the System for selected task or required operation*	System Characteristics	Demands on the Operator in Selected Task or Required Operation**	Quantified Workload Ratio	Quantified Logarithmic Workload Ratio***	Operator Rating Decivals, dV****
Is it satisfactory without improvement? Yes	Excellent Highly desirable	Operator compensation not a factor for desired performance	1.0	0	0
	Good Negligible deficiencies	Operator compensation not a factor for desired performance	1.25	0.1	1
	Fair - Some mildly unpleasant deficiencies	Minimal operator compensation required for desired performance	1.6	0.2	2
No → Deficiencies warrant improvement	Minor but annoying deficiencies	Desired performance requires moderate operator compensation	2.0	0.3	3
	Moderately objectionable deficiencies	Adequate performance requires considerable operator compensation	2.5	0.4	4
	Very objectionable but tolerable deficiencies	Adequate performance requires extensive operator compensation	3.2	0.5	5
Is adequate performance attainable with a tolerable operator workload? Yes / No → Deficiencies require improvement	Major deficiencies	Adequate performance not attainable with maximum tolerable operator compensation. Controllability not in question	4.0	0.6	6
	Major deficiencies	Considerable operator compensation is required for control	5.0	0.7	7
	Major deficiencies	Intense operator compensation is required for control	6.4	0.8	8
Is it controllable? Yes / No → Improvement mandatory (Operator Decisions)	Major deficiencies	Control will be lost during some portion of required operation	8.0	0.9	9

* Definition of required operation involves designation of the system phase and/or subphases with accompanying conditions.

** Operator workload is defined as the total work associated with all mental and physical tasks per unit time period required to successfully complete the selected task or required operation.

*** For ease in operator rating and in understanding the meaning of each rating, the workload imposed by the system configuration and operations doubles by definition every third rating. As a result, the increasing system engineering evaluation scale is logarithmic.

**** A decival (dV) is equal to 10 $\log_{10}$ (Quantified workload ratio).

***** Higgins, Thomas H. Draft Report No. November 1979.