Flight Testing and Evaluation of Airborne Multisensor Display Systems

FRANCIS BELINNE* AND DARYL E. TONINI†

Air Force Missile Development Center, Holloman Air Force Base, N. Mex.

This paper presents the test and analysis techniques developed for a multisensor display evaluation conducted at the Air Force Missile Development Center, Holloman Air Force Base, New Mexico. Four display systems were tested in the laboratory and in-flight during the period from July 12, 1967 to July 11, 1968. The over-all objectives of the program were to evaluate the relative performance of each display system in the TV mode using a series of targets designed to provide a variety of optical conditions. The displays were also tested in the radar mode; however, this paper treats only the electro-optical phase.

Nomenclature

C = range at which a target is identified on the cockpit display
$\bar{C}$ = average contact range
C/V = ratio of display contact-to-visual range
$\overline{C/V}$ = $\Sigma(C/V)/n$
$\bar{C}/\bar{V}$ = ratio of average display contact range to average visual range
G = gray shades measured with RETMA test pattern
n = number of events
R_D = TV limiting resolution of display
R_O = over-all resolution of sensor and display
R_S = TV limiting resolution of sensor
S = frame to frame signal storage
V = range at which a target is identified out-the-window
$\bar{V}$ = average visual range
X_D = product of normalized values of G, R_D, and S
X_O = product of normalized values of G, R_O, and S

I Introduction

DEVELOPMENT of certain multisensor weapon guidance systems has produced a new generation of aircraft cockpit displays. Included among these sensors are electro-optical, infrared, laser, and improved resolution radar. Space and operational factors in the F4 Phantom II have led to a requirement for a multimode display suitable for use with a variety of sensors. These range from direct view storage tubes to a first generation airborne scan-converter display system.

In an effort to establish the relative technical merit of various contractor proposals, a competitive test was directed by the Aeronautical Systems Division, Air Force Systems Command, United States Air Force. As a result of these tests, conducted by the Air Force Missile Development Center (AFMDC), a method of testing and analysis has been evolved which establishes a quantitative approach to the testing of cockpit display systems and which provides a new tool for predicting display performance in terms of easily measured laboratory parameters.

Table 1 Normalized laboratory results

	System			
	A	B	C	D
Parameter				
Nonlinearity				
Horizontal	0.94	[a]	0.71	1.36
Vertical	0.37	[a]	1.09	1.54
Geometric	1.39	[a]	1.16	0.45
Minimum horizontal and vertical sync	1.18	[a]	0.71	1.18
Aspect ratio	1.000	[a]	1.005	0.959
Brightness	0.74	[a]	1.05	1.21
Gray shades	0.92	0.99	0.99	1.11
Resolution	0.73	0.56	1.10	1.62
Signal storage	0.73	1.13	1.03	1.13

[a] Not measured.

II System Laboratory Evaluation

General

Four display systems were evaluated in the laboratory using inputs from standard signal sources. The following parameters were measured: 1) horizontal, vertical, and geometric nonlinearity, 2) minimum horizontal and vertical sync signals, 3) aspect ratio, 4) brightness, 5) gray shades, 6) resolution, and 7) signal storage.

With the exception of signal storage, these parameters were measured using conventional inputs and techniques. Normalized values of the performance by system are presented in Table 1.

Technique for Measuring Signal Storage

The conventional technique for measuring signal storage does not provide a basis of comparison for different types of displays since signal retention is normally measured as a function of the duty cycle [pulse repetition frequency (PRF), pulse width, and amplitude] using arbitrarily chosen light levels.

In order to provide the required basis of comparison between the systems, a photometric technique was developed for this test. The equipment used consisted of the AN/ASM-184 Aircraft Test Set, a Spectra Pritchard Photometer, and an oscilloscope.

With the AN/ASM-184 test pattern input to the display system, the intensity of the lighted block in the central horizontal row of blocks was adjusted to the intensity obtained during the brightness test. With the photometer viewing the brightest whole block in the lowest horizontal row of blocks, the output of the photometer with its indicator disconnected

Presented as Paper 69-317 at the AIAA 3rd Flight Test, Simulation, and Support Conference, Houston, Texas, March 10–12, 1969; submitted March 25, 1969; revision received Aug. 28, 1969.

* Major, U.S. Air Force, Display Project Officer.

† Major, U.S. Air Force, Display Project Officer. Associate Member AIAA.

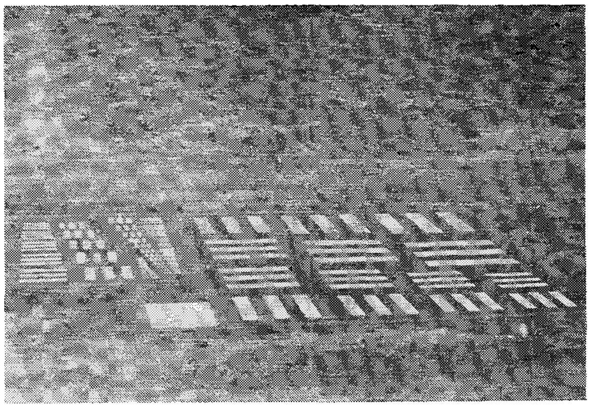
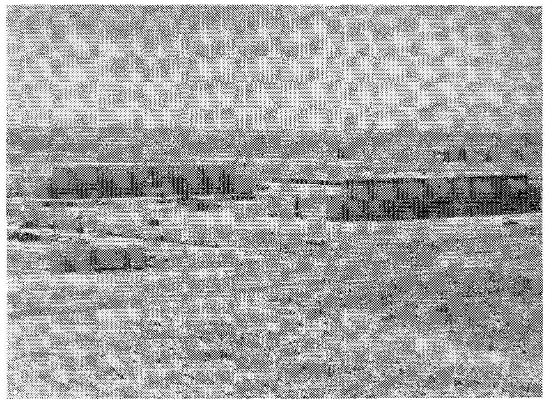
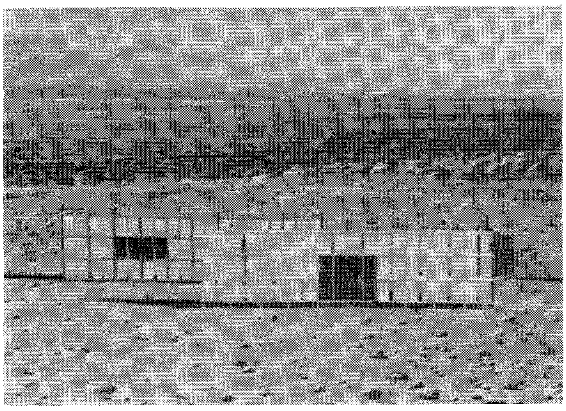


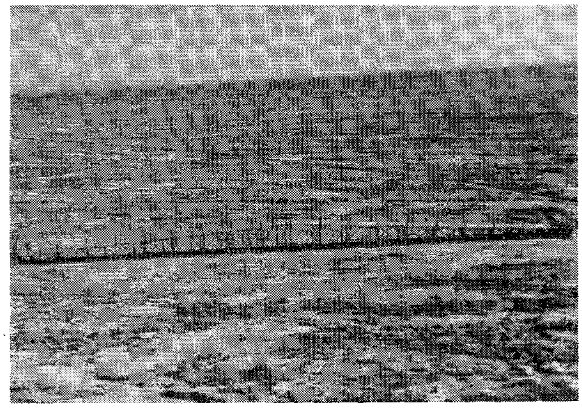
Photo Resolution Target



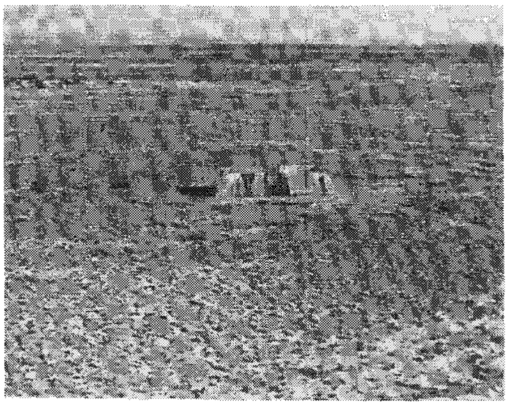
Holloman Building Target



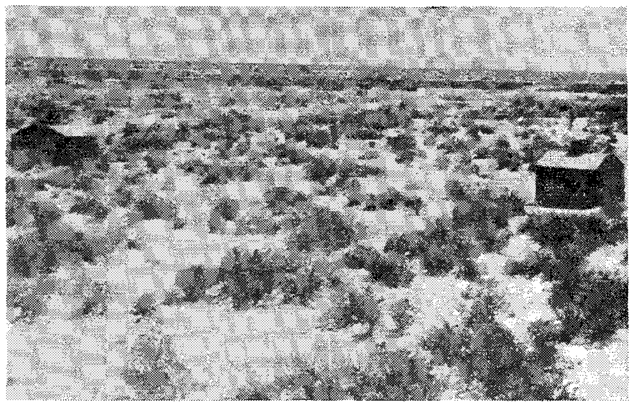
Simulated Building (Waldo) Target



Simulated Bridge Target



Simulated SAM Site Target



Simulated Tanks Target

Fig. 1 Targets.

was displayed on the oscilloscope. To allow the photo-multiplier/oscilloscope combination to respond more rapidly, the photometer sensitivity scale was set at its least sensitive position in order to reduce the resistive capacitive (RC) time constant of the output circuit of the photo-multiplier tube to a minimum. By adjusting the filter density on the photometer to prevent saturation of the photo-multiplier tube, the maximum and minimum intensities for one field were then read from the oscilloscope and recorded.

A zero value of the display tube output in the presence of ambient light was obtained by setting the display brightness control at its minimum setting and recording the photometer output on the oscilloscope. By dividing the maximum intensity by the minimum intensity values, the percentage storage was calculated.

III System Flight Evaluation

Test Procedures

The targets used for the evaluation were designed to produce a wide range of lighting conditions as well as to test the display systems under critical performance conditions. They consisted of a photo resolution target, black with white stripes, 100 ft × 130 ft; a large building on Holloman AFB, 134 ft × 363 ft × 49 ft; a simulated building structure (Waldo) on White Sands Missile Range, 220 ft × 75 ft × 24 ft composed of four nonjoining panels; a simulated bridge, 300 ft × 20 ft × 15 ft high at midpoint with 1300 ft of black-top to each end; a simulated surface-to-air missile (SAM) site, 50 ft × 70 ft × 12 ft high; and simulated tanks 14 ft × 10 ft × 10.5 ft high (Fig. 1).

The range instrumentation system used for this evaluation consisted of an FPS-16 precision tracking radar with associated software for automatic data output. Time-space positions were obtained by a one station trigonometric solution using azimuth, elevation, and range of the target aircraft. The FPS-16 data output was converted to an analog format to provide a real-time plot board output which was used for a quick-look data analysis. Precision for the data system is summarized in Table 2.

The principal aircraft instrumentation consisted of a master panel for control of cockpit cameras, UHF tone and camera event marks, a 1 KHz tone which was transmitted to correlate mission events with the space position data, a C band transponder which aided in tracking the aircraft with the FPS-16 radar, a rear seat scope camera to obtain a cine record of the display in flight, an audio recorder for a record of interphone communications, and a commercial Sony TV modified for use as an airborne monitor. The TV monitors were used to directly evaluate the output of the electro-optical (EO) sensor and to isolate video difficulties to either the sensor or to the display system. A differential amplifier was designed and constructed to provide common-mode rejection of any noise present on the leads and to provide isolation of the sensor as well as the Sony monitor from the display system under test.

In-Flight Procedures

Since the tests were concerned with the range at which the target could be identified on the display, the mission profile was designed to this end. The aircraft was pointed toward the target area and the gunsight pipper placed near the target at a range at which the target could not be identified, either on the attack display or out the window. This required that the pilot hold the pipper in the vicinity of the target although he was unable to identify it. He transmitted a tone when the target could be identified on the display (defined as contact), and another was transmitted when the target could be identified out the window (defined as visual). Ranges at contact and visual were obtained from FPS-16 radar space position data.

In order to standardize the data (minimize the variables), the passes for each target were flown on specified headings, using a nominal dive angle and air speed. Since the purpose of the test was to evaluate display system capability, the EO sensor envelope was not considered as a test constraint.

Target identification criteria were standardized for each target to permit comparison of data between the individual systems as well as between individual crew members. Since the passes were flown with the front and rear crew members alternately performing the display contact identification, the necessity of having a repeatable criteria was particularly important.

All interphone communications were recorded and used to assess any parameters which may have affected any individual pass. Typical factors would include interfering clouds, heavy shadows obscuring the target area, too shallow dive angle, etc.

Analysis of Flight Data

In analyzing the flight test data, the data were first grouped by display system, target, and by the true course of the aircraft relative to the target. Then the average contact

Table 2 Radar precision for the data system

System	X axis, yd	Y axis, yd	Z axis, yd	Over-all, yd
A	9.2	10.13	11.29	10.24
B	8.38	10.35	11.75	10.28
C	5.42	6.28	7.72	6.54
D	5.42	6.28	7.72	6.54

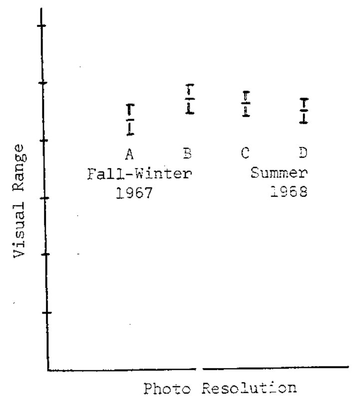


Fig. 2 Sun angle effects, PR target.

($\bar{C}$) and visual ($\bar{V}$) ranges were calculated for each group of passes on the same nominal heading. Next, the ratio of contact-to-visual range (C/V) was calculated for each pass; and from these ratios, a mean value of contact-to-visual ($\overline{C/V}$) ratio was computed. (Note: $\bar{C}/\bar{V} \neq \overline{C/V}$.) The sample was corrected for finite sample size, and a domain for the mean value based on a 0.90 to 0.95 probability was computed. The domain for the mean was then used as an indicator of the relative reproducibility and validity of the test data when compared with results on different headings or with different systems for the same target. In general, the domain served as a moderation factor for comparing mean values; since if there was an overlap in the domain of $\overline{C/V}$ for a given confidence level, it was not possible to state that the mean value of one condition was representative of an improvement, regardless of the relative difference in the mean values. (Of course, one may reduce the magnitude of the domain of the mean by reducing the probability that the mean will lie between given limits and thus eliminate the undesired overlap. In this manner, the mean values may be compared at a reduced level of confidence.)

Analysis of Test Variables

As the test developed, it soon became apparent that the effect of many of the flight test variables could not be independently evaluated. However, several important parameters did lend themselves to analysis. As an illustration of both cases, weather effects, pilot experience factors, front and rear cockpit identification criteria, and the effect of sun angles have been chosen as examples.

1) *Weather.* The problems encountered while attempting to quantify the degree of interference caused by weather conditions led to the elimination from the comparative analysis of any pass which had an interference from weather. That is, unless there was at least 40 miles visibility, no ceiling-and no interfering cloud cover, the data from the pass of mission were deleted.

2) *Pilot variables.* Although there were individual differences between the AFMDC pilots against particular targets, it was not possible to detect a significant trend attributable to either pilot ability or rate of learning. However, there was an unmistakable reflection of pilot visual acuity in one grouping of test data. This will be discussed in more detail later in this paper.

3) *Identification criteria.* Early in the program, considerable concern was expressed about the ability of the front and rear pilots to apply the same criteria for the display contact identification. It was shown that, in fact, the criteria chosen had been so defined that no measurable difference could be detected between the $\overline{C/V}$ values obtained from the front and rear cockpit.

4) *Sun angle effects.* Since all of the display systems were not tested at the same time of the year, the question arose as to whether or not the values of C/V for the various

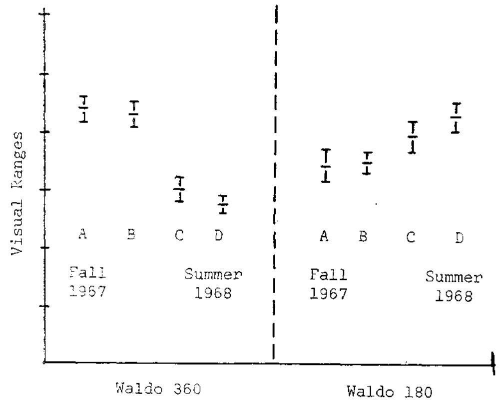


Fig. 3 Sun angle effects, Waldo.

systems could be affected to the detriment or advantage of one or more of the systems. The data were analyzed to determine how the visual range values were affected by sun angle. Two targets, Waldo and Photo Resolution, were considered. For the Photo Resolution target, it was found that there was no significant change in visual range with the time of year (Fig. 2). Since this target is a flat panel on the ground, this result was not unexpected. For the Waldo target, however, it was found that the visual ranges obtained from the south side of the target were much higher in the

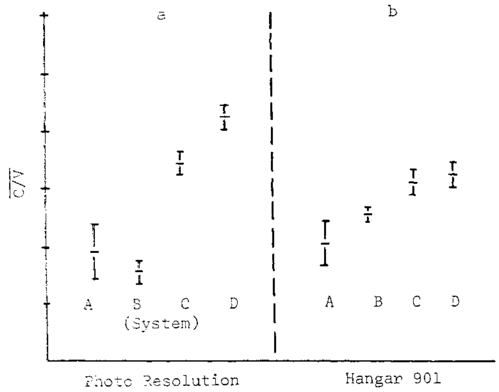


Fig. 4 Results, PR and Holloman building.

winter than in the spring. From the north side of the target, the opposite was true (Fig. 3). If one were to assume that the contact range was unaffected, the C/V for a system tested in the autumn would, as a result, be lower than if it were tested in the spring. However, since the C/V obtained on the reciprocal heading was affected inversely, the effect was considered self-compensating. Looking at the data for the Photo Resolution target, it can be seen that the improvement

Table 3 Normalized values of G, R_D, and S

System	G	R_D	S
A	0.92	0.73	0.725
B	0.99	0.56	1.125
C	0.99	1.10	1.025
D	1.11	1.62	1.125

in display performance was the result of the capability of the system and was not attributable to sun angle effects.

Results of Analysis

Comparative results of each system by target are presented in Figs. 4–7. Two data points did not follow the general trend and required further analysis.

1) On Fig. 4 (Photo Resolution), the performance of system B appears to be low when compared with the performance achieved against other targets. However, this target was chosen as a test of system resolution and, as can be seen from Table 3, system B did in fact have the lowest performance index of the systems tested. This illustrates the strong effect of resolution capability when other target variables are constrained.

2) On Fig. 5 (Waldo 180), the performance of system C was higher than would have been anticipated. A detailed analysis of the data for this target and heading revealed that the visual ranges obtained by the pilot responsible for 14 of the 27 passes were significantly lower than the average visual range obtained by other pilots against this target and heading. Since no unusual test conditions existed to invalidate these 14 passes, the data were included in the analysis. However, it is interesting to note that the visual acuity of this particular pilot (20/20) were measurably lower than the acuity of the other test pilots (20/10 to 20/15).

The over-all results for all targets by system are shown in Fig. 7. Since there was no overlap in the domains of the mean of the systems, within the limits of 95% confidence, it was concluded that there was a significant statistical difference in the flight performance capability of the four systems tested.

IV Correlation of Flight and Laboratory Results

Constraints

Since the measurements recorded in the laboratory concerned the display and the TV monitor only while the flight parameter (C/V) was a function of the total system, it was necessary to examine the inter-relationship between a display and a sensor. Such a relationship has been proposed by G. K. Slocum et al.[1] and was used in this analysis. This

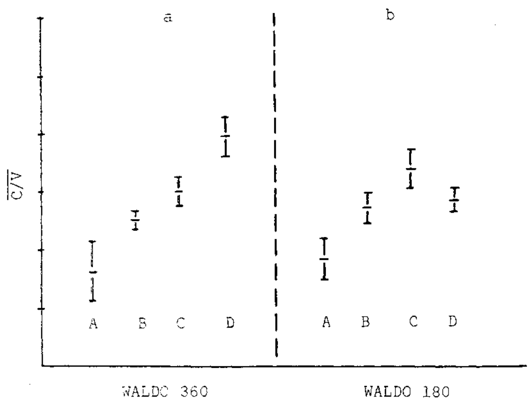


Fig. 5 Results, Waldo.

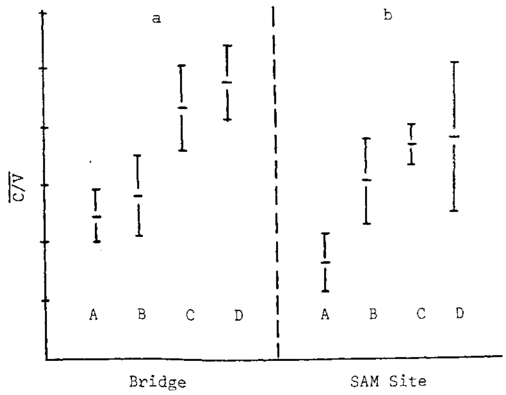


Fig. 6 Results, bridge, SAM site.

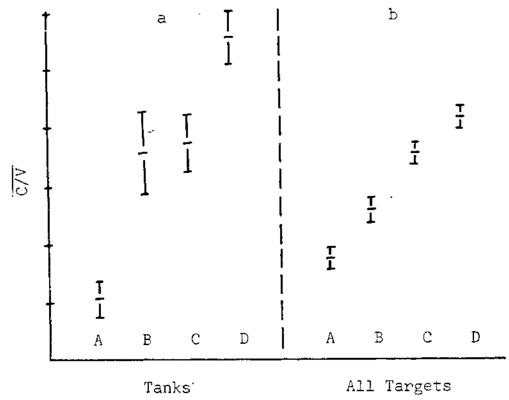


Fig. 7 Results, tanks, over-all.

relationship is nonlinear and is expressed by

$$R_O = R_S R_D/(R_S^2 + R_D^2)^{\frac{1}{2}}$$

The shape of this curve is as shown in Fig. 8.

Using this equation, the measured resolution of each display as seen through the GPL 1000 camera was adjusted for interface degradation. (The TV limiting resolution of the GPL 1000 is approximately 1890 lines.) The corrected values are shown in Table 3 and agree very closely with the resolution capabilities claimed by the tube vendors.

Normalizing of Laboratory Parameters

Four major display characteristics were judged to be significant enough to influence the scope contact to visual range ratio (C/V). These were brightness (B), gray shades (G), horizontal resolution (R_D), and field-to-field storage (S).

1) Brightness. During the inflight video tests, the front and rear display were equipped with different filters. As a result, the brightness of the front and rear displays was significantly different. Nevertheless, the C/V ratios achieved when using either cockpit display show no significant difference. Therefore, it was concluded that once a sufficient level of brightness is reached, any increase in that level will not directly affect C/V performance although it will increase the over-all quality of the presentation.

2) Gray shades, resolution, and storage. The mean values for G, R_D, and S were calculated and were normalized about the mean for each system. When normalizing S, 50% storage was taken as the zero level; i.e., smearing at 50% storage would render the system unusable. These values are as shown in Table 3.

Correlation of Results

Numerous combinations of these normalized parameters were plotted against the over-all mean C/V ratios of the systems. The combination which resulted in the most meaningful relationship was the product of the three; i.e.,

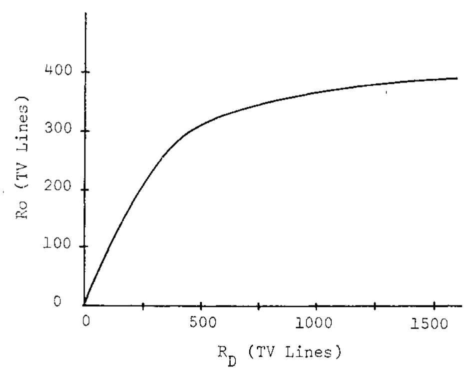


Fig. 8 Display/sensor resolution relationship (Slocum).

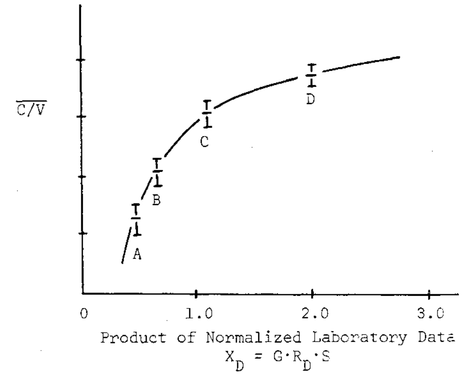


Fig. 9 Flight/lab. correlation.

$X_D = G \times R_D \times S$. X_D was then considered as the measure of technical capability of each system.

Figure 9 shows the curve considered to be the most representative of the relationships between the in-flight performance and the laboratory evaluation for the displays and sensor tested based on display resolution.

Using Slocum's equation for over-all system resolution, $X = C \times R_o \times S$ was computed and plotted against C/V. These points now describe a linear relationship. From this, it is reasonable to conclude that sensor display performance for any sensor or display improvements may be predicted (Fig. 10).

Conclusions

The results of this test program have demonstrated that it is feasible to correlate in-flight performance of multimode cockpit displays with three principal laboratory measured parameters, and within the limits of the test conducted it is possible to predict system performance on the basis of gray shades capability, resolution, and signal storage.

The validity of choice of the C/V ratio as the principle flight test parameter was established, thus permitting quantitative comparison of performance data collected under a relatively wide range of test conditions. In addition, a restatement of system performance specifications was obtained in terms of the parameters found to have significant impact on the in-flight performance. Also, by extrapolating the data, with some reservations, it is possible to predict performance of improvements to the sensor/display system.

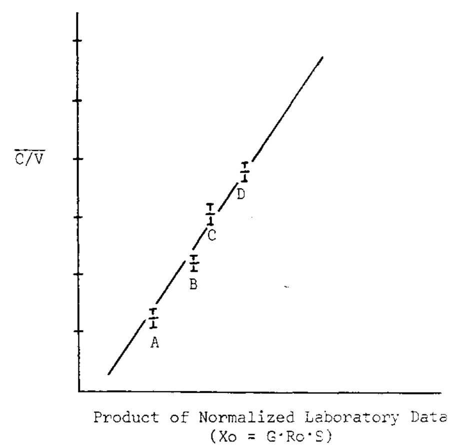


Fig. 10 Flight/system correlation.

Reference

[1] Slocum, G. K. et al., "Airborne Multi-Sensor Displays," Rept. TM 883 A, July 1967, Hughes Aircraft Co.