MARCH 1977

TEST EFFECTIVENESS STUDY REPORT

Contract No. NAS ●-W2949

(NASA-CR-153212) TEST EFFECTIVENESS STUDY REPORT: AN ANALYTICAL STUDY OF SYSTEM TEST EFFECTIVENESS AND RELIABILITY GROWTH OF THREE COMMERCIAL SPACECRAFT PROGRAMS Final Report (Hughes Aircraft Co.) 200 p

N77-25200

HC A09/MF A01

Unclas

G3/12 15008

HUGHES

HUGHES AIRCRAFT COMPANY
SPACE AND COMMUNICATIONS GROUP

Hughes Ref No. D8234 • SCG 70103R

MARCH 1977

TEST EFFECTIVENESS STUDY REPORT

Contract No. NAS [illegible]W2949

AN ANALYTICAL STUDY OF
SYSTEM TEST EFFECTIVENESS
AND RELIABILITY GROWTH
OF THREE COMMERCIAL
SPACECRAFT PROGRAMS

Prepared by: J F Feldstein
Senior Project Engineer
Principal Investigator

With assistance from:
E P Mahr
D R Thibodo
G A Engle
H D Marbach

Reviewed by: W D Brown
Systems Engineering Manager

HUGHES

HUGHES AIRCRAFT COMPANY
SPACE AND COMMUNICATIONS GROUP

Hughes Ref No. D8234 • SCG 70103R

ABSTRACT

Failure data from 16 commercial spacecraft were analyzed to evaluate failure trends, reliability growth, and effectiveness of tests. The study showed that the test programs were highly effective in ensuring a high level of in-orbit reliability. There was only a single catastrophic problem in 44 years of in-orbit operation on 12 spacecraft. The results also indicate that in-orbit failure rates are highly correlated with unit and systems test failure rates. The data suggest that test effectiveness estimates could be used to guide the content of a test program to ensure that in-orbit reliability goals are achieved. Cost considerations suggest that an aggressive corrective action program to achieve a near-zero failure rate for all testing should be implemented.

The unit and systems level qualification tests were found to be marginally effective in detecting design deficiencies. Systems level vibration is an effective qualification test. Systems level tests are more effective for qualification than for acceptance. Most vibration problems are detected during qualification tests or during unit acceptance. System acceptance thermal tests were marginally effective, probably because of deficiencies in test technique. Most thermal vacuum test failures are detected in the first 60 hours of testing.

Test effectiveness models can be used at unit and systems level to estimate the defective population and to predict in-orbit or systems test performance. The in-orbit cumulative failure rate has a constant slope which implies a constant test effectiveness. The in-orbit failure rates generally decrease with successive numbers of a spacecraft series, thus indicating reliability growth. The systems test and in-orbit cumulative failure rate data cannot be accurately combined into a single Duane plot because the slopes are generally different.

The consequences of an in-orbit critical failure do not permit reduced testing. The test program should be considered insurance. The most cost effective program is one in which a near-zero failure rate is achieved through aggressive corrective action. A continuing effort is warranted to measure and improve the screening effectiveness of each test.

CONTENTS

		Page
1.	OBJECTIVES	1-1
2.	APPROACH	2-1
3.	FAILURE DATA BASE	3-1
	3.1 Data Base Summary	3-1
	3.2 Failure Report Categories	3-7
	3.3 Test Screen Effectivity	3-11
	3.4 Defects Analysis	3-20
	3.5 Orbit FR Analysis	3-31
4.	TEST EFFECTIVENESS	4-1
	4.1 Introduction	4-1
	4.2 Data Base	4-1
	4.3 Specific Test Effectiveness	4-2
	4.3.1 Qualification Tests	4-2
	4.3.2 Acceptance Tests	4-7
	4.4 General Test Effectiveness	4-8
	4.4.1 Qualification Tests	4-11
	4.4.2 Acceptance Tests	4-11
	4.5 Test Effectiveness for Design Problems	4-12
	4.5.1 Qualification Tests	4-12
	4.5.2 Acceptance Tests	4-13
	4.6 Reliability Predictions Using a Test Effectiveness Model	4-15
5.	FAILURE RATES AS A FUNCTION OF TIME	5-1
6.	SCREENING EFFECTIVENESS OF SYSTEMS THERMAL-VACUUM TESTS	6-1
7.	COST EFFECTIVE TESTING	7-1
8.	CONCLUSIONS	8-1
	8.1 General	8-1
	8.2 Qualification Testing	8-2
	8.3 Acceptance Testing	8-3
	8.4 Reliability Growth	8-3
	8.5 Cost Effectiveness	8-4
	APPENDICES	
A.	Summary of Information Codes	A-1
B.	Analysis of Some Simple Failure Rate Models	B-1

1. OBJECTIVES

The primary purpose of this study was a detailed analysis of failure reports spanning three commercial spacecraft programs to assess both test screening effectiveness and reliability growth through all phases of spacecraft test and in-orbit operation. In particular, data were collected and analyzed for a total of 16 commercial spacecraft with the following specific objectives:

1) Segregation of all failures and anomalies by defect classification, cause of failure, and the test environment
2) Identification of failure groupings and trends related to learning curves and/or design maturity
3) Quantitative assessment of the defect screening effectiveness for each unit, system, and in-orbit test phase, using the System Test Effectiveness technique*
4) Evaluation of the Duane Reliability Growth Analysis technique to describe system test and in-orbit performance**
5) Development of recommendations for a cost effective test program

---

*A. F. Timmins, "A Study of Relationship between Performance in Systems Tests and Space," Proceedings of the Institute of Environmental Sciences, I.172, 1975

**J. T. Duane, TIS Report UF62MD00, General Electric Company DCM&G Department, Erie, Pennsylvania, February 1962

2. APPROACH

All failure reports for three commercial spacecraft programs (16 spacecraft) have been reviewed, categorized, and analyzed. These data include in-house Trouble and Failure Reports (TFRs), International Failure Reports (IFRs), and in-orbit problems as reported by customers. For simplicity, each event shall herein be referred to as a failure report, or FR, even though it should be noted that each FR does not necessarily represent a failure, since anomalies, problems, intermittents, wear-out, and failures were all reported on the FRs.

A total of 1991 FRs and 90 orbit reports were corrected and coded. Each FR was described on an 80 column format with 27 discrete pieces of information. Where missing or contradicting information existed, alternate sources (e.g., spacecraft logs) were utilized to correct and/or complete the data. The description of the information codes entered in each column is provided for reference in Appendix A.

Each of the 2081 data points was then evaluated in a bulk data analysis. Various sorting and cross sorting provided information about overall test performance and specific problem areas. The test time to failure was estimated for each of the systems test and in-orbit FRs. The thermal vacuum test time to failure data were obtained directly from the respective spacecraft log books.

The above data were used for the Duane analysis where the Systems Test and in-orbit Duane plots were calculated for each spacecraft. The linear regression coefficients were estimated, using the $\log_{10}$ of the cumulative failure rates and test times. The combined systems test and in-orbit cumulative failure rates were also plotted for individual spacecraft.

The test effectiveness for specific test phases of initial-ambient, vibration, and thermal vacuum, at both unit and systems test levels, was evaluated in accordance with the Timmins Test Effectiveness Equation. The test effectiveness of the entire unit and systems test program was also evaluated. In addition, the ability of the qualification and early production tests to detect design problems was analyzed. The use of test effectiveness as a reliability prediction technique was developed. The attributes of a cost effective test plan were developed.

3. FAILURE DATA BASE

3.1 DATA BASE SUMMARY

The failures and anomalies which comprise the data base are distributed through the various test and in-orbit phases of three commercial programs. The quantities of spacecraft for each program are summarized in Table 3-1. The relative scheduling of these programs is shown in Figure 3-1.

Program 1 consisted of two separate contracts. The first (A) was for a prototype and four flight spacecraft and the second (B) was for four additional flight spacecraft. Of the seven spacecraft successfully launched, six are still in service and have accumulated over 28.6 years of orbital life. The one orbital failure occurred in the payload subsystem of the second flight spacecraft (F-2), and resulted in a degradation of performance to unacceptable levels for commercial communications traffic. Program 2 was essentially a scaled down version of Program 1, both in payload and size. It utilized many of the same or similar units, although it introduced some new payload concepts. All three flight spacecraft were successfully launched and are in service. Program 3 was a new contract utilizing the

TABLE 3-1. PROGRAM SUMMARY

Program	Spacecraft Quantities	In-Orbit		Remarks
		Number	Σ Years	
1A	Prototype + four flight	4	18.16	Large communications payload
1B	Four flight	3*	10.52	Separate buy of same spacecraft
2	Prototype + three flight	3	9.99	Scaled down version of Program 1
3	Three flight + one payload shelf	2**	5.27	Follow-on contract to Program 2
Total for all programs	Two prototype Fourteen flight One payload shelf	12	43.94	Average of 3-2/3 years per spacecraft

*F-6 was launched but booster failed.

**F-6 has not been launched, and the shelf has not been integrated to provide a spacecraft.

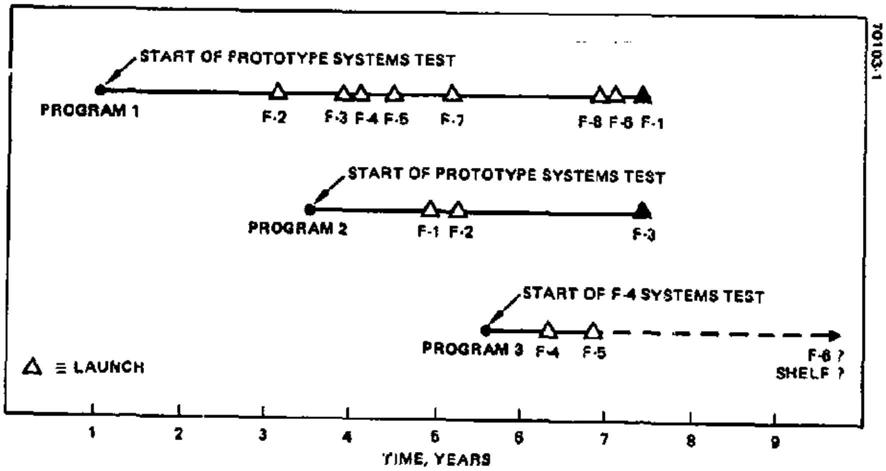


FIGURE 3-1. RELATIVE PROGRAM SCHEDULES FOR SYSTEMS TEST AND LAUNCH

same basic spacecraft as Program 2. Therefore, the three flight spacecraft were considered equivalent to a follow-on buy and were sequentially numbered, beginning with F-4. No prototype spacecraft was fabricated or tested for this program. A payload equipment shelf was additionally integrated with its respective units. Only two of the spacecraft have been launched, and the shelf has not been integrated into a spacecraft. Both of the launched spacecraft are in service.

The data base for this study includes all failure reports (FRs) generated on Program 1 from its contractual start date through the last launch (F-1). All FRs to date on Programs 2 and 3 are included. All orbit anomally reports to date (March 1, 1977) by the respective customers are also included. These reflect a total of almost 44 orbit years accumulated on the 12 spacecraft in orbit.

Because of the differences between programs in failure reporting methods at the lower levels of assembly, data were counted, but not analyzed, for FRs below unit level. For this study, a unit was defined as an identifiable (by serial number) piece, or pieces, of hardware which were formally acceptance tested either by a vendor or by Hughes Aircraft. These could and did include some major subassemblies which were delivered and then integrated to form a larger assembly (unit) and acceptance tested again prior to delivery to the spacecraft for integration. An example of this was the high power traveling wave tube amplifier (TWTA) unit. It consists of two major subassemblies: an electronics power controller and the traveling wave tube. Each of these was acceptance tested and delivered by the respective vendors and then integrated to form the TWTA and acceptance tested again. Each of these was considered to be a unit for this study. Serial numbers of each were tracked by the configuration management system. Each subassembly could be removed as a unit and reintegrated to form a different unit as test and performance results required.

To assess the relative complexity of each of the programs, Table 3-2 summarizes both the types and quantities of units which comprised a given flight spacecraft. The data are divided into eight different subsystems for each spacecraft and the number of mechanical units is also identified for each subsystem. A normalization factor was calculated, based on number of units per spacecraft, by using Program 1 as the reference. Another way to compare the complexity of programs is through the use of electronics parts count. Table 3-3 presents this data. It was concluded that values of 2.50 for Program 2 and 2.44 for Program 3 should be used when data required normalization for program to program comparisons. This was derived by use of a weighted average using 3 for electronics parts complexity, 2 for total units, and 1 for types of units.

All FRs were divided into five distinct groups by hardware development phase: subunit, unit, systems test, orbit, or not applicable. The distinction between subunits and units has been previously discussed. Systems test level was defined to begin at delivery of units to the spacecraft and continue through launch of each spacecraft, including those tests performed at launch base. After review of each FR, a small number were

TABLE 3-2. UNIT TYPES AND QUANTITIES

Unit Types	Program 1	Program 2	Program 3
Total unit types per spacecraft subsystem			
Communications	42	8(2)(1)	8(2)(1)
Antenna	12(11)	5(5)	5(1)
Telemetry and command	35(9)	3(1)	3(1)
Propulsion	10(4)	5(1)	5(1)
Power	13	6	6
Structure	22(6)	15(3)	17(3)
Attitude control	7(1)	8(3)	9(3)
Apogee kick motor	3(1)	2(1)	2(1)
Total	144(33)	52(16)	55(16)
Total units per spacecraft subsystem			
Communications	246	53(1)(2)	53(1)(2)
Antenna	18(17)	5(5)	5(5)
Telemetry and command	62(10)	4(1)	4(1)
Propulsion	75(18)(3)	10(2)	10(2)
Power	26	21	21
Structural	26(11)	43(3)(4)	46(3)(4)
Attitude control	2(3)	12(5)	13(5)
Apogee kick motor	3(1)	2(1)	2(1)
Total	468(60)	150(18)	154(18)
Normalization factor by unit types	1(reference)	2.77	2.62
Normalization factor by total units	1(reference)	3.12	3.04

Notes: ( ) = Number of mechanical units with no electronics or electrical hardware.

(1) Does not include 17 receiver unit parts that were included (tracked) on Program 1

(2) Because of (1), 25 receiver units were not included

(3) Includes 42 tracked heater elements

(4) Includes 28 tracked temperature sensors

TABLE 3-3. ELECTRONIC PARTS COUNT PER SPACECRAFT

	Program 1	Program 2	Program 3
Number of electronic parts	16,846	8,370	8,569
Normalization factor	1(reference)	2.01	1.97

excluded from the data base for one of the following reasons: 1) it was not written against spacecraft hardware (e.g., test equipment malfunction which had no effect on spacecraft hardware or the test program), 2) it was a duplicate of another FR documenting the same problem but originating in a different area, and/or 3) the original FR or copies of the FR could not be found in the files-to-enable-coding.

A summary of total data base is presented in Table 3-4. The FRs which form a basis for the analytical study are shown within the heavy line box. Table 3-5 presents the same data as a percentage of the whole, by program. Of interest is the increasing effectiveness of the subunit level tests on each program with the systems test level decreasing and the unit test level remaining relatively constant. This is probably the result of both varying FR reporting criteria on each program and a learning process whereby the subunit testing became more effective in screening problems. The fact that so few FRs were present on Program 3 at systems test level

TABLE 3-4. NUMERICAL SUMMARY OF BULK DATA

Program	Number of Spacecraft	Total FRs Reviewed	Level of FRs				Not* Applicable
			Subunit	Unit	Systems Test	In-Orbit	
1	8 + prototype	1,426	286	721	263	68	88
2	3 + prototype	355	104	177	42	13	19
3	3 + shelf	300	98	160	18	9	15
Total	14 flight + 2 prototype + 1 shelf	2,081	488	1,058	3^	90	122

*Includes duplicates, not spacecraft hardware problems, etc.

TABLE 3-5. PERCENTAGE SUMMARY OF BULK DATA

Program	Number of Spacecraft	Total FRs Reviewed	Level of FRs				Not* Applicable
			Subunit	Unit	Systems Test	In-Orbit	
1	8 + prototype	100	20.0	50.6	18.4	4.8	6.2
2	3 + prototype	100	29.3	49.9	11.8	3.7	5.4
3	3 + shelf	100	32.7	53.3	6.0	3.0	5.0
Average	14 flight +2 prototype + 1 shelf	100	23.5	50.8	15.5	4.3	5.9

*Includes duplicates, not spacecraft hardware problems, etc.

(18, and 6 of these were induced) requires further investigation. The higher percentage of systems test level FRs on Program 1 is also a result of smaller specification margins than those imposed on Programs 2 or 3.

As indicated in Table 3-4, there was a large difference in the number of FRs between programs, with Program 1 having about twice as many FRs per spacecraft as Program 3. However, normalized data based either on electronic parts count or on hardware differences is less disparate. The normalized data are summarized in Table 3-6. For the three spacecraft programs, approximately one FR was generated for each 127 electronic parts. Note that the apparent low FR average of Programs 2 and 3 becomes a high average when normalized on the basis of hardware differences. Normalized FR rates must be used to rationally compare any two spacecraft programs.

As discussed, many of the units utilized on each of the three programs were supplied to Highes Aircraft Company by vendors which were primarily international companies in France, Canada, Japan, England, Italy, etc. Unit data were separated into international and Hughes categories, based on the reporting activity. This was carried, as a subdivision of unit FRs, through part of the bulk data analysis to determine any significant differences in problem categories between the two methods of procurement. A summary of the division is presented in Table 3-7.

TABLE 3-6. FAILURE REPORT RATES COMPARISON

Program	Number of Spacecraft	Total Unit Systems, and In-Orbit FRs	Electronic Parts per FR	FR Average per Spacecraft	Normalization Factor	Normalized FR Average per Spacecraft
1	8 + prototype*	1,140	125.6	134.1	1 (reference)	134.1
2	3 + prototype*	251	116.7	71.7	2.50	179.25
3	3 + shelf*	202	148.4	57.7	2.44	140.79
Combined	~15.5	1,593	127.1	102.7	1.67	171.51

*Each of these ≅ ½ of a spacecraft

TABLE 3-7. HUGHES VERSUS INTERNATIONAL FAILURE REPORTS

Program	Unit Failure Reports*		Total
	Hughes	International	
1	348(48)	373(52)	721(100)
2	63(36)	114(64)	177(100)
3	16(10)	144(90)	160(100)
Total	427(40)	631(60)	1,058(100)

*Percentage shown in parentheses

3.2 FAILURE REPORT CATEGORIES

Two basic breakdowns of the FRs were required to effectively present and analyze the data. A distinction between FRs on the qualification (prototype) and acceptance (flight) spacecraft was made to distinguish between the different test programs. The second breakdown was one of "importance" for each of the two data sets. A significant number (about 20 percent) of all the FRs "integration and/or test" (I&T) induced. These were not the fault of the hardware, but were the result of human errors and required unit or system rework. Some (about 11 percent) of all the FRs were secondary in nature, in that although FRs were initiated, no resultant physical hardware action was taken. A third distinction (about 1 percent) was made for FRs initiated against hardware for an anomaly which could not be repeated. All remaining FRs (about 68 percent) were classified as primary and were included in the analysis data base. A summary of these several categories is presented in Tables 3-8, 3-9, and 3-10 for international units, Hughes produced units, and systems level test FRs, respectively, for both qualification and acceptance hardware.

Essentially, all qualification hardware was produced at Hughes for the two prototype spacecraft. Tables 3-8 and 3-9 show that a significantly higher number of secondary FRs were generated internationally; 64 percent were for out-of-specification conditions with no retest/rework required. This is probably a result of the formal contractual interface between Hughes and these subcontractors with regard to out-of-specification conditions.

Both Hughes and international unit test results showed that electrical test errors caused about 41 percent of the I&T induced FRs, and electrical test overstress caused another 24 percent. Hardware mishandling accounted for an additional 12 percent. It is probable that increased emphasis on test procedures and safety practices, both electrical and mechanical, would have been cost effective.

System test FRs are summarized in Table 3-10. A total of 31 percent of all FRs on Programs 1 and 2 were against the prototype hardware. The I&T induced failure rate remained about constant at 23 percent on all programs for prototype or flight hardware. The susceptibility of Program 1 to propulsion subsystem line heater damage is apparent. Hardware handling problems and electrical test errors were significant on all programs. The small number of primary FRs at the systems level on Programs 2 and 3 makes significant analysis difficult. However, the fact that the sample size is so small is a credit to the effectiveness of the subunit and unit test programs. The reduction in secondary out-of-specification and test error problems on Programs 2 and 3, relative to Program 1, probably reflects more mature systems test procedures and specifications requirements as a result of experience gained on the first program. It should be noted that the systems test teams were derived from the same organization on all three programs.

TABLE 3-8. FAILURE REPORT CATEGORIES - INTERNATIONAL UNITS

FR Type	Program* 1 Qualification	Program* 1 Acceptance	Program* 2 Qualification	Program* 2 Acceptance	Program* 3 Acceptance	Sum of all Programs
Totals	373		114		144	631
Qualification/acceptance breakdown	2(1)	371(99)	2(2)	112(98)	144	...
Less secondary	0	25(7)	0	35(31)	42(29)	102(16)
Less I&T induced	1(50)	62(17)	0	12(11)	25(17)	100(16)
Less anomalies	0	0	0	0	0	0
Total primary	1(50)	284(77)	2(100)	65(58)	77(54)	429(68)
Secondary reasons						
Out of spec/OK		20		17	28	65(64)
Meas tech/No retest		0		7	4	11(11)
Test equip/No retest		1		5	1	7(7)
Mfg error/multiples		1		0	0	1(1)
Test error/No rework		3		6	9	18(18)
Anomalies/No retest		0		0	0	–0
Misc/No retest		0		0	0	0
Total secondary	0	25	0	35	42	102(100)
I&T induced reasons						
Damaged wires, heaters, etc.		7		1	1	9(9)
Elec test overstress		17		2	7	26(26)
Elec test error		30		4	9	43(43)
Hardware handling	1	5		2	2	10(10)
Wrong environment imposed		1		~~3~~	6	10(10)
Bad part selection in test		2		0	0	2(2)
RCS test error		0		0	0	0
Test fixture induced		0		0	0	0
Total I&T induced	1	62	0	12	25	100(100)

*Percentage shown in parentheses; the data have not been normalized

TABLE 3-9. FAILURE REPORT CATEGORIES - HUGHES UNITS

FR Type	Program* 1 Qualification	Program* 1 Acceptance	Program* 2 Qualification	Program* 2 Acceptance	Program* 3 Acceptance	Sum of all Programs
Totals	348		63		16	427
Qualification/accept breakdown	41(12)	307(88)	29(46)	34(54)	16	---
Less secondary	0	5(2)	0	5(15)	8(50)	18(4)
Less I&T induced	9(22)	72(23)	8(28)	11(32)	4(25)	104(24)
Less anomalies	0	0	0	0	0	0
Total primary	32(78)	230(75)	21(72)	18(53)	4(25)	305(72)
Secondary reasons						
Out of spec/OK				3	1	4(22)
Meas. tech/No retest						0
Test equip/No retest						0
Mfg error/multiples					7	7(39)
Test error/No rework		1				1(6)
Anomalies/No retest		2		2		4(22)
Misc/No retest		2				2(11)
Total Secondary	0	5	0	5	8	18(100)
I&T induced reasons						
Damaged wires, heaters, etc.	4	8	0	5	0	17(16)
Elec test overstress	0	18	4	1	0	23(22)
Elec test error	1	34	1	1	3	40(39)
Hardware handling	4	6	2	3	1	16(15)
Wrong environment imposed	0	0	0	0	0	0
Bad part selection in test	0	0	0	0	0	0
RCS test error	0	6	1	0	0	7(7)
Test fixture induced	0	0	0	1	0	1(1)
Total I&T induced	9	72	8	11	4	104(100)

*Percentage shown in parentheses; the data have not been normalized

TABLE 3-10. FAILURE REPORT CATEGORIES - SYSTEM TEST

FR Type	Program* 1 Qualification	Program* 1 Acceptance	Program* 2 Qualification	Program* 2 Acceptance	Program* 3 Acceptance	Sum of all Programs
Totals	263		42		18	323
Qualification/accept breakdown	85(32)	178(68)	12(29)	30(71)	18	...
Less secondary	6(7)	20(11)	0	1(3)	2(11)	29(9)
Less I&T induced	19(22)	38(21)	3(25)	8(27)	6(33)	74(23)
Less anomalies	0	10(6)	0	0	0	10(3)
Total primary	60(71)	110(62)	9(75)	21(70)	10(56)	210(65)
Secondary reasons						
Out of spec/OK	0	11				11(38)
Meas. tech/No retest	0	0				0
Test equip/No retest	0	0				0
Mfg error/multiples	0	0			2	2(7)
Test error/No rework	2	8				10(34)
Anomalies/No retest	0	0		1		1(4)
Misc/No retest	4	1				5(17)
Total Secondary	6	20	0	1	2	29(100)
I&T induced reasons						
Damaged wires, heaters, etc.	12	13	0	1	2	28(38)
Elec test overstress	0	2	0	2	1	5(7)
Elec test error	4	9	2	1	2	18(24)
Hardware handling	3	13	1	2	1	20(27)
Wrong environment imposed	0	1	0	0	0	1(1)
Bad part selection in test	0	0	0	0	0	0
RCS test error	0	0	0	0	0	0
Test fixture induced	0	0	0	2	0	2(3)
Total I&T induced	19	38	3	8	6	74(100)

*Percentage shown in parentheses; the data have not been normalized

3.3 TEST SCREEN EFFECTIVITY

The primary data base was sorted by test activity to determine the origin of each FR. This sort reveals the screening effectivity of each test environment at each phase of development. When combined with the in-orbit results, these data become the basis for the evaluation of the two analytical methods presented in Sections 4 and 5 of this report.

Tables 3-11 and 3-12 present the respective numerical data sorts for qualification and acceptance test screens by hardware level and program. Table 3-11 shows a surprisingly high systems level FR rate (i.e., systems level FRs/all FRs) for Program 1 prototype spacecraft of 64.5 percent. The Program 2 prototype spacecraft FR rate of 28.1 percent was more reasonable, even though based on a small sample size. However, these FR rates are both high when contrasted to the systems level acceptance FR rates of 17.6, 20.2, and 11.0 percent for Programs 1, 2, and 3, respectively. The disproportionate systems level FR rate for prototype spacecraft may be the result of a higher screening effectiveness at the systems level for qualification (design?) problems.

Particular environments within a test phase are discernible from Tables 3-11 and 3-12. For all vibration tests, Z (longitudinal) and X (lateral) axes were more effective than Y (lateral) with 30, 26, and 12 FRs detected at each axis, respectively; but 32 percent of the total of 101 vibration test FRs did not indicate the axis of importance, most of these having been originated at Hughes. Both hot (22) and cold (16) temperatures were important to all the unit thermal vacuum (TV) tests, with eclipse simulation being particularly important to the spacecraft systems level, resulting in 14 of the 29 TV FRs (48 percent). Note that the cold phase of "despun compartment" or "electronic shelf" TV (DCTV/ESTV) simulates an extended eclipse condition and was therefore included. The unit thermal cycle test environment indicated no strong preference as to where FRs occurred.

The data of Tables 3-11 and 3-12 were grouped and distributed by spacecraft assignment to understand the significance of the test screen distribution and to look for trends. This is presented on Figures 3-2 and 3-3 for the unit and systems level test screens. The four groupings chosen for unit screens were: 1) initial assembly, inspection and performance tests, 2) vibration and post vibration performance test (SPT), 3) thermal vacuum and/or thermal cycling and the final inspection and performance test, and 4) performance tests which were not defined as to specifics and all reaction control subsystem (RCS) tests. For the systems level test screens, the four chosen groupings of FRs were: 1) initial assembly, integration, and integrated systems tests (IST), 2) vibration and SPT, 3) all TV tests (DCTV/ESTV and spacecraft TV) and the final IST and inspection, and, 4) all launch operations. Those units which were not assigned to a spacecraft were carried as a separate spacecraft grouping to enable analysis.

TABLE 3-11. QUALIFICATION TEST SCREEN – NUMERIC

Qualification Test Screen		Program 1			Program 2			
		International Units	Hughes Units	Systems Level	International Units	Hughes Units	Systems Level	Total
Initial inspection				7		2		9
Mechanical mfg.			3	3				6
Initial elec. perf. test		1	11	5	2	4	1	24
Perf. test – not identified			4			6		10
Final elec. perf. test						2		2
STV/ESTV	Transition			1				1
	Hot							
	Cold (eclipse)						1	1
Initial IST				11			1	12
Vibration	Setup			1				1
	Not defined		3	2		1		6
	X-axis		1	6		1		8
	Y-axis			2				2
	Z-axis		1	4		1		6
SPT (post vibration)			2	10		1	5	18
TV/STV	Hot		1					1
	Cold			5				5
	Eclipse simulation			1				1
	Ukn or transition			1				1
Thermal cycle	Unknown			-			-	
	Transition			-		3	-	3
	Hot			-			-	
	Cold		4	-			-	4
Final IST		-	-	1	-	-		1
Final inspection			1				1	2
RCS tests			1					1
Subtotal *		1 (1.0)	32 (34.4)	60 (64.5)	2 (6.3)	21 (65.6)	9 (28.1)	125
Total			93			32		125

*Percentage shown in parentheses; the data have not been normalized

TABLE 3-12. ACCEPTANCE TEST SCREEN - NUMERIC

Acceptance Test Screen		Program 1 International Units	Program 1 Hughes Units	Program 1 Systems Level	Program 2 International Units	Program 2 Hughes Units	Program 2 Systems Level	Program 3 International Units	Program 3 Hughes Units	Program 3 Systems Level	Total
Initial inspection		5	6	9	1	1		2	1	2	27
Mechanica mfg.		1	3	3		1	1	1		2	12
Initial elec. perf. test		113	85	21	34	4	3	31			291
Perf. test – not identified		68	23		11	8		20	1		131
Final elec. perf. test		4	5					1	1		11
DCTV/ESTV	Transition						2				2
	Hot			3							3
	Cold (eclipse)			2			4				6
Initial IST				27			4			1	32
Vibration Unit-random Spacecraft-sine and random	Setup not defined	1	24		1	1					27
	X-axis	10	6		2						18
	Y-axis	3	5		2						10
	Z-axis	4	14	1				4	1		24
SPT (Post vibration)		16	6	2			1			2	27
TV/STV	Hot	11	9	1		1					22
	Cold	4	1	2	6			5			18
	Eclipse simulation	–	–	5	–	–		–	–	2	7
	Unk. or transition	2	1	1	2			3			9
Thermal cycle	Unknown		1	–			–			–	1
	Transition	1	2	–			–	1		–	4
	Hot	14	11	–	2		–	1		–	28
	Cold	26	16	–	4		–	8		–	54
Final IST		–	–	10	–	–	4	–	–	1	15
Final inspection		1		4			2				7
RCS tests			12		–	2					14
Launch operations		–	–	19	–	–		–	–		19
Subtotals*		284 (45.5)	230 (36.9)	110 (17.6)	65 (62.5)	18 (17.3)	21 (20.2)	77 (84.6)	4 (4.4)	10 (11.0)	815
Total			624			104			91		819

*Percentage shown in parentheses, the data have not been normalized

Figures 3-2 and 3-3 show a number of significant features about the primary data base. Program 1 was generally more uniformly distributed within each test group than were Programs 2 or 3. This is probably the result of the larger sample size. At the unit level, at least 42 percent of all FRs were found in group 1 -- initial assembly and performance checkout. Some of group 4 (undefined performance test) undoubtedly belong in group 1; therefore, the initial assembly and performance test is the most important phase for test screening at the unit level. Vibration (group 2) is less effective (15 percent) than thermal (21 percent) tests (group 3) on all programs and about 21 percent of all unit FRs were in group 4 -- undefined performance. Within the unit groupings, the qualification data are not significantly different from acceptance. Qualification vibration was somewhat more effective than acceptance (21 percent of all qualification FRs versus 17.5 percent of all Program 1 FRs and 17 percent versus 8 percent for all Program 2). Thermal testing was somewhat less effective at qualification (18 percent versus 21 percent for Program 1). However, this difference was very pronounced at the systems level. Almost 42 percent (Program 1) and 55 percent (Program 2) of the respective prototype spacecraft FRs occurred in vibration. This contrasts to 16 and 20 percent overall and only 3 and 10 percent without prototype spacecraft FRs. Therefore, vibration and SPT were very important for systems level qualification (prototype spacecraft) but were significantly less effective screens for systems level acceptance (flight spacecraft). Systems level TV tests were more effective for the flight spacecraft than for the prototype spacecraft.

Figure 3-2 shows that as the unit tests screens mature with time, the effectiveness of the initial ambient tests remains constant at about 41 percent. Unit vibration generally trends downward with each successive spacecraft and program, and thermal test grouping trends upward in importance. The percentage of undefined unit performance tests (group 4) trends upward (19 percent on Program 1 to 27 percent on Program 3), indicating that attention to detail on the FRs waned with each successive program. It should also be noted from Table 3-7 that the percentage of foreign participation also increased with each successive program. The high percentage of group 4 for the F-4 and F-5 assigned units from Program 3 makes the Program 3 evaluation difficult. By comparison with Program 3 F-6 assigned units, where a high percentage is assigned to group 1 (59 percent) with a corresponding low group 4, one could conclude that the vast majority of group 4 Program 3 FRs actually occurred during initial checkout and the 44 percent overall group 1 value is very conservative.

At the systems level the initial assembly integration and IST (group 1) tests are again very effective: 50, 33, and 50 percent for Programs 1, 2, and 3, respectively, and 48 percent overall.

The systems level sine and random spacecraft vibration tests produced only 1 FR during acceptance (Z sine), and five FRs were additionally reported after the vibrational environment during the Systems Performance Test (SPT).

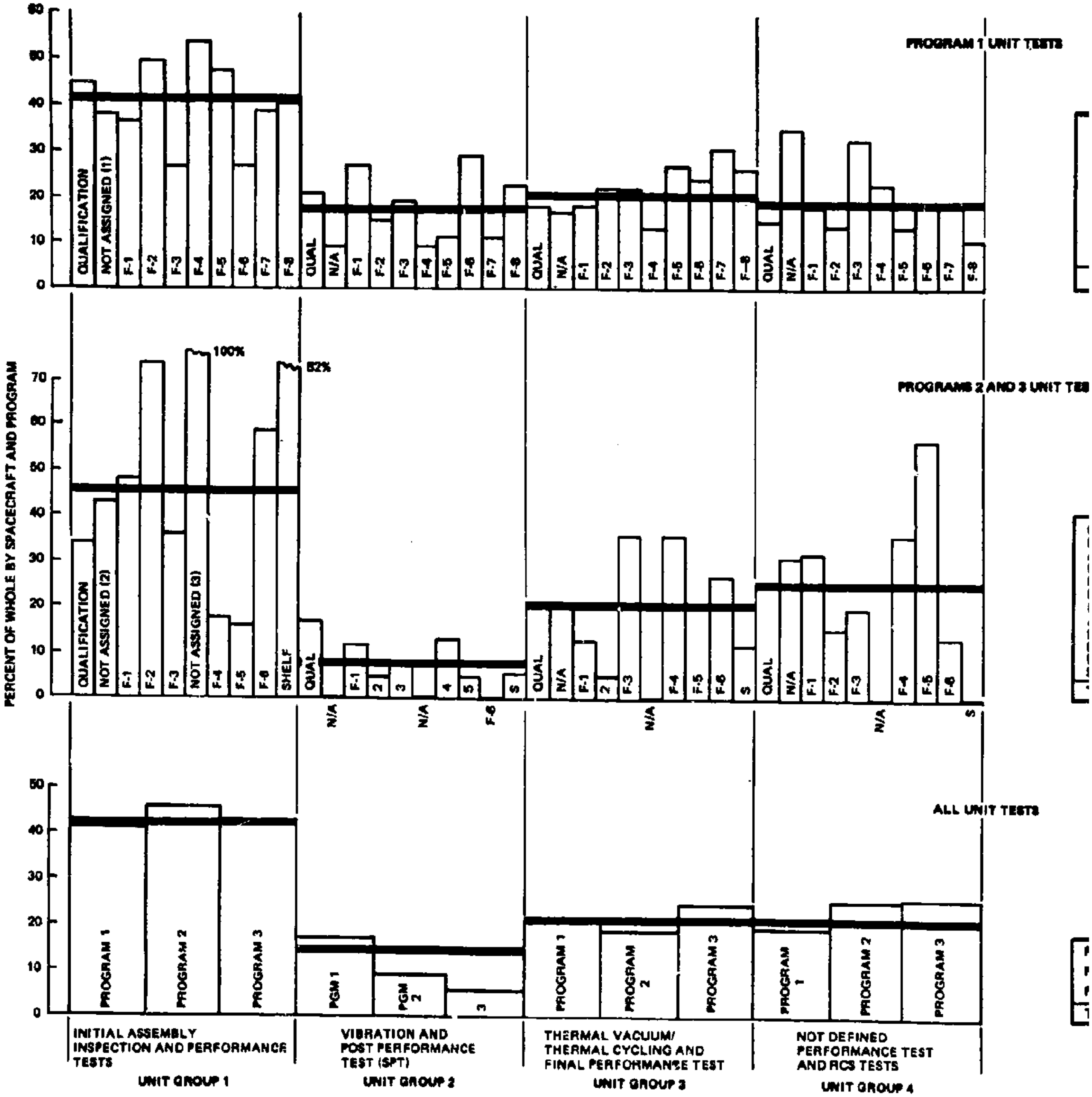
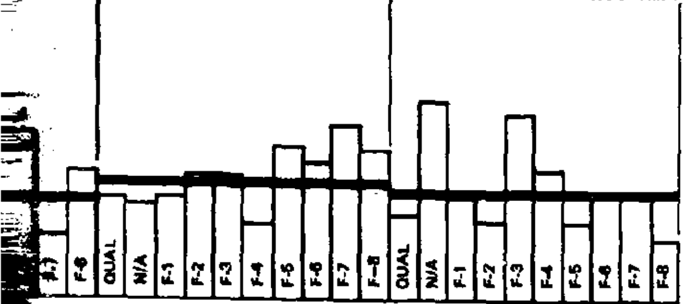


FIGURE 3-2. DISTRIBUTION OF UNIT TEST SCREENS

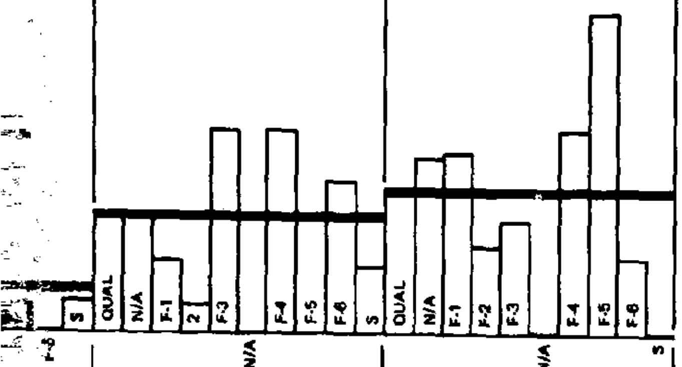
PROGRAM 1 UNIT TESTS



	NUMERICAL VALUES				
	INITIAL PERF	VIBRATION AND SPT	TV/TC + FINAL PERF	UNKN PERF + RCS	TOTALS
QUALIFICATION	15	7	6	5	33
NOT ASSIGNED (1)	15	4	7	16	42
F-1	28	21	14	14	77
F-2	37	11	16	10	74
F-3	14	10	11	17	52
F-4	45	8	11	19	83
F-5	25	6	14	7	52
F-6	11	12	10	8	41
F-7	14	4	11	7	36
F-8	23	13	16	5	57
TOTALS	228	96	115	108	547

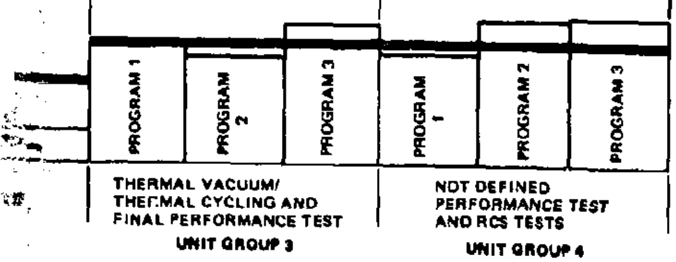
70103-2

PROGRAMS 2 AND 3 UNIT TESTS



QUALIFICATION	8	4	5	6	23
NOT ASSIGNED (2)	6	0	3	0	14
F-1	12	3	2	8	26
F-2	14	1	1	3	19
F-3	9	2	9	5	25
NOT ASSIGNED (3)	1	0	0	0	1
F-4	4	3	8	7	22
F-5	3	1	4	11	19
F-6	13	0	6	3	22
SHELF	14	1	2	0	17
TOTALS	84	16	40	48	187

ALL UNIT TESTS



PROGRAM 1	228	96	115	108	547
PROGRAM 2	49	10	20	27	106
PROGRAM 3	35	5	20	21	81
TOTALS	312	111	155	156	734

OF UNIT TEST SCREENS

FOLDOUT FRAME

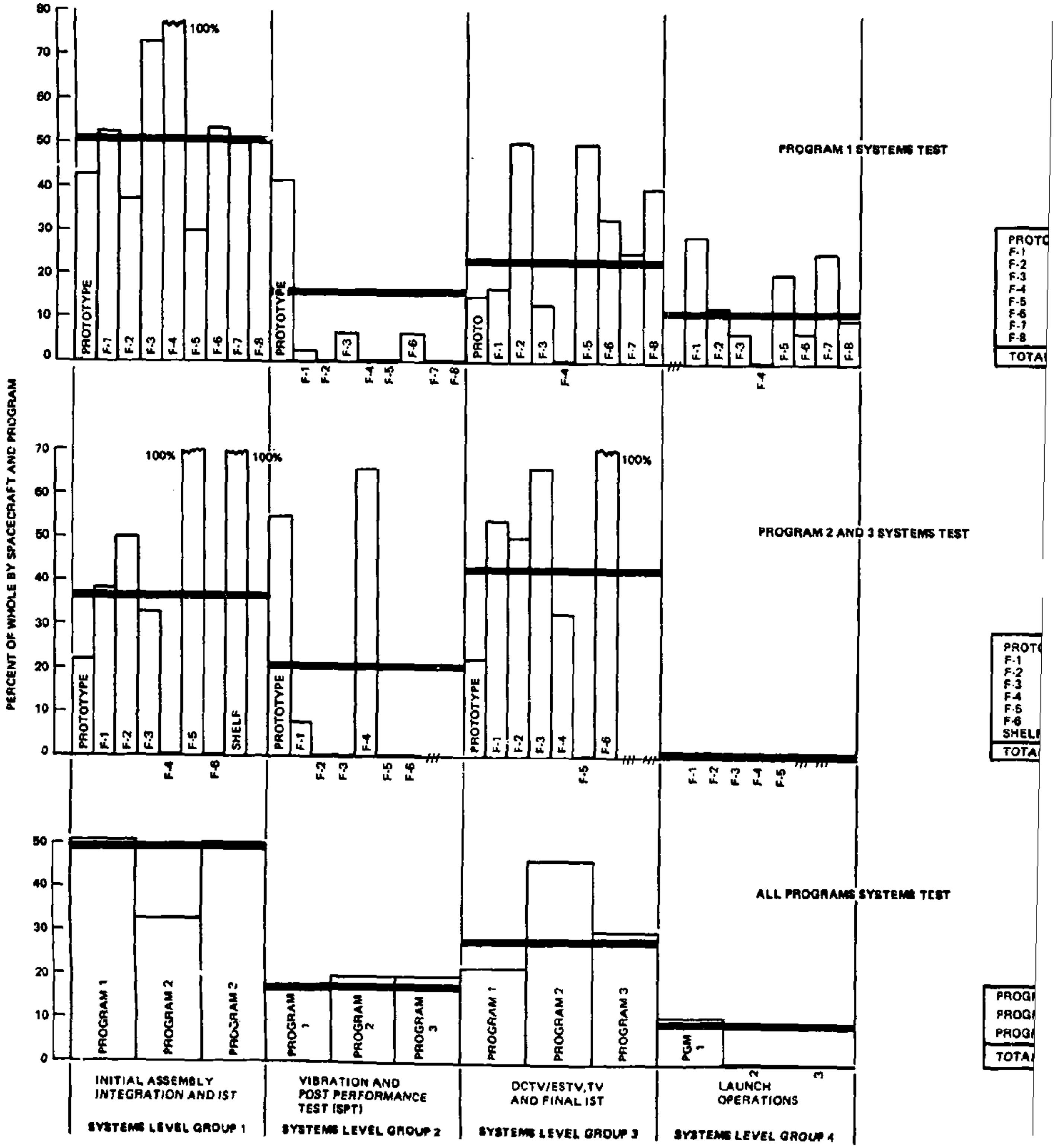
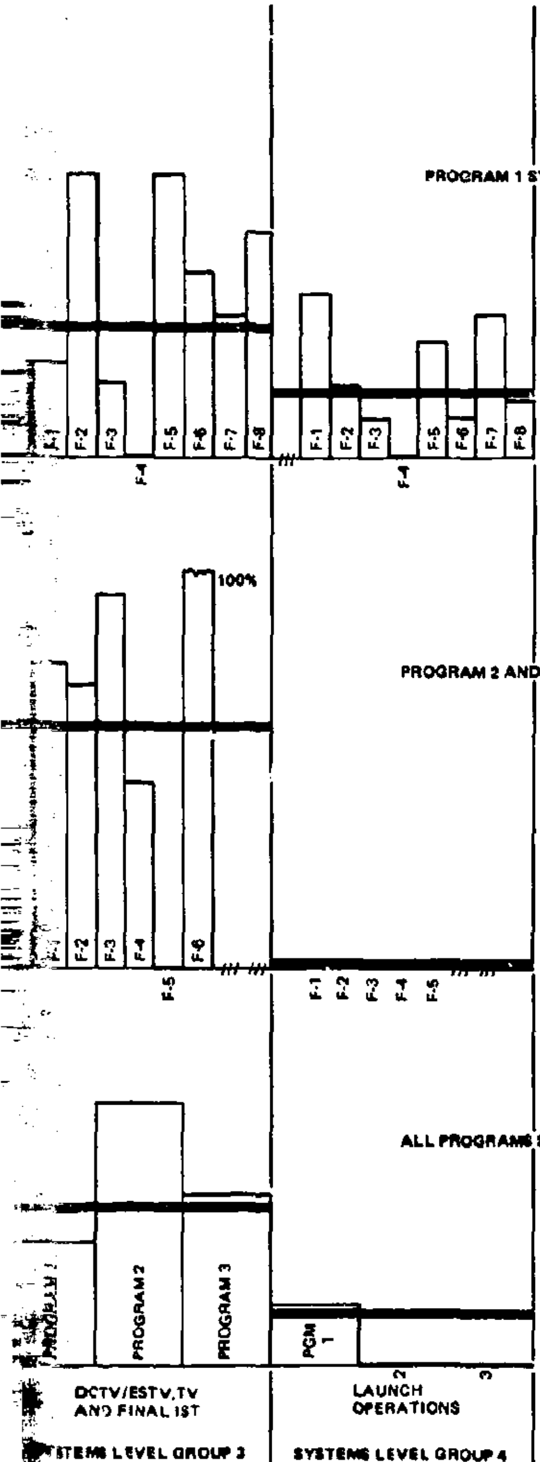


FIGURE 3-3. DISTRIBUTION OF SYSTEMS TEST SCREENS

PRECEDING PAGE BLANK NOT FILMED

70103-3



...EMS TEST SCREENS

PROGRAM 1 SYSTEMS TEST

	NUMERICAL VALUES				
	INITIAL ASSEMBLY AND IST	VIBRATION AND SPT	DCTV/ESTV, TV AND IST	ALL LAUNCH OPERATIONS	TOTALS
PROTOTYPE	26	25	9	N/A	60
F-1	20	1	6	11(F-2)	38
F-2	3	0	4	1(F-1)	8
F-3	11	1	2	1	15
F-4	6	0	0	0	6
F-5	3	0	5	2	10
F-6	8	1	5	1	15
F-7	4	0	2	2	8
F-8	5	0	4	1	10
TOTALS	86	28	37	19	170

PROGRAM 2 AND 3 SYSTEMS TEST

PROTOTYPE	2	5	2	N/A	9
F-1	5	1	7	0	13
F-2	1	0	1	0	2
F-3	2	0	4	0	6
F-4	0	2	1	0	3
F-5	3	0	0	0	3
F-6	0	0	2	N/A	2
SHELF	2	N/A	N/A	N/A	2
TOTALS	15	8	17	0	40

ALL PROGRAMS SYSTEMS TEST

PROGRAM 1	86	28	37	19	170
PROGRAM 2	10	6	14	0	30
PROGRAM 3	5	2	3	0	10
TOTALS	101	36	54	19	210

PRECEDING PAGE BLANK NOT FILMED

Figure 3-3 shows that an increase in Program 2 systems level TV occurred with a corresponding decrease in the initial (group 1) screening. The data shows that this may have been caused by both inadequate unit thermal testing (Figure 3-2 Group 3) and systems level initial performance testing (Group 1 of Figure 3-3).

Figure 3-3 also shows that all of the launch operations FRs occurred on Program 1. The first spacecraft launched on Program 1 (the F-2) had 11 FRs generated during launch operations. Although FRs continued to be generated against succeeding spacecraft, they accrued at an average of just over 1 per spacecraft. During the five launch operations of Programs 2 and 3 no FRs occurred. This is probably the result of a mature test program with proven procedures. It should be noted that the F-1 spacecraft of Program 1 was unique in that after the spacecraft systems level tests were completed it was stored and not launched. After all other spacecraft on Program 1 were tested, F-1 was returned to Hughes, upgraded with the latest configurations (a few design changes) and then again put through a systems level spacecraft test program prior to shipment to the launch base for launch.

Table 3-13 was prepared, using the data for the systems level acceptance test screens of Table 3-12. It distributes the FRs by spacecraft sets of all, initial (F-1), all remaining, and F-1 retest for Program 1 into three new groupings: 1) pre-spacecraft environment; 2) spacecraft environment (vibration and TV); and 3) launch operations. As can be seen from

TABLE 3-13. SPACECRAFT SYSTEMS LEVEL TEST SCREENS BY SPACECRAFT

		Program‡							
		1				2			3
Group	Systems Level Acceptance Test Screen by Spacecraft	Total, All Spacecraft	F-1 Spacecraft Original	F-2 through F-8 Spacecraft	F-1 Spacecraft Retest	Total, All Spacecraft	F-1 Spacecraft	F-2 and F-3 Spacecraft	Totals, All Spacecraft
1	Initial electrical and mechanical assembly and performance	33	9	21	3	4	4	0	2*
	DCTV/ESTV	5	0	5	0	6	3	3	0
	Initial IST	27	6	19	2	4	1	3	1
2	Vibration	1	0	1	0	0	0	0	0
	SPT	2	1	1	0	1	1	0	2
	TV/STV	9	0	8	1	0	0	0	2
	Final IST and inspection	14	3	9	2	6	4	2	1
3	Launch operations	19	11**	8**	N/A	0	0	0	0***
	Total	110	31	72	8	21	13	8	8
	Average per spacecraft		31	10.3	8		13	4.0	2.7

*Shelf data not included (two FRs)

**F-2 was 1st spacecraft launched, therefore, F-2 launch operations FRs were arbitrarily assigned to F-1 spacecraft and vice versa

***F-6 spacecraft has not as yet been tested.

‡Data have not been normalized.

such a grouping and distribution, there is a significant distinction in FR observations after the initial spacecraft of the two different test programs; note that Program 3 is considered as a follow-on to Program 2. Both F-1 spacecraft had three times as many problems as the average of succeeding spacecraft. The F-1 spacecraft retest problems of Program 1 provide unique evidence that the initial F-1 spacecraft test program was not a totally effective screen.

3.4 DEFECTS ANALYSIS

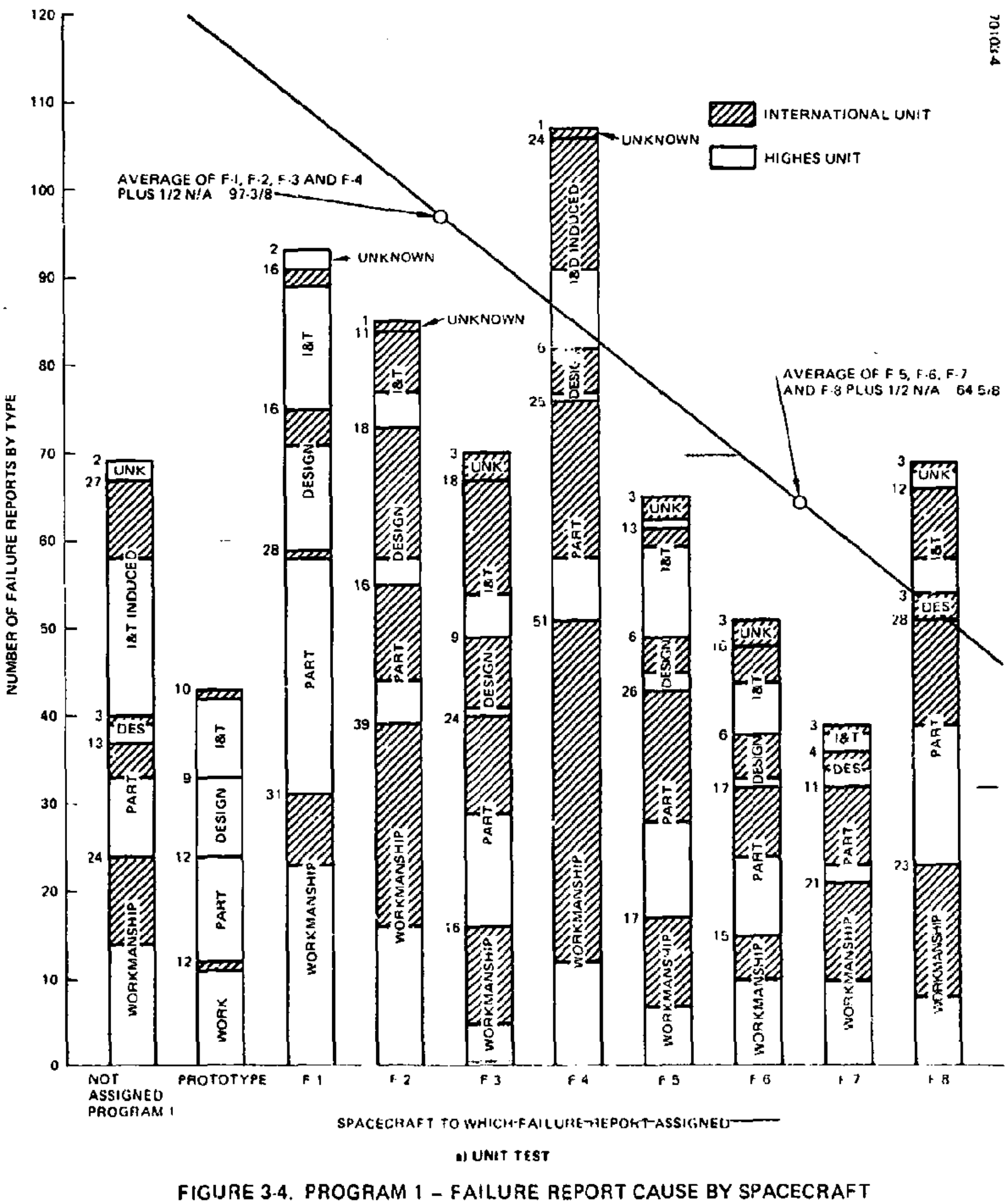
Each primary FR was divided into four categories according to the cause of failure: 1) workmanship, 2) part, 3) design, and 4) unknown. The results of this division by level of assembly and assigned spacecraft for each program are shown on Figures 3-4 and 3-5. The unit and systems level FRs are shown as a function of the assigned spacecraft. Where an FR occurred on a unit which was not eventually assigned to a spacecraft, it was classified as "not assigned." The relationship of international and Hughes unit FRs was maintained. Integration and test (I&T) FRs and the anomaly FRs were also added.

The figures show the downward trend from early to late spacecraft in the numbers of FRs occurring during a given spacecraft acceptance test program. An estimate of this trend is presented on each figure wherein the averages for the first set of spacecraft and the last set are calculated and a two point line is drawn. These averages included an apportioned number of the unassigned FRs.

With the few exceptions, each set of spacecraft exhibits the property that the last spacecraft units have a higher FR rate than the trend of the previous spacecraft units in that set. A probable explanation is that the problem units from each set of spacecraft are eventually assigned to the last spacecraft in that set. Units would be delayed in their production sequence to allow "repair" and would reenter the acceptance cycle later in time, resulting in late spacecraft assignment.

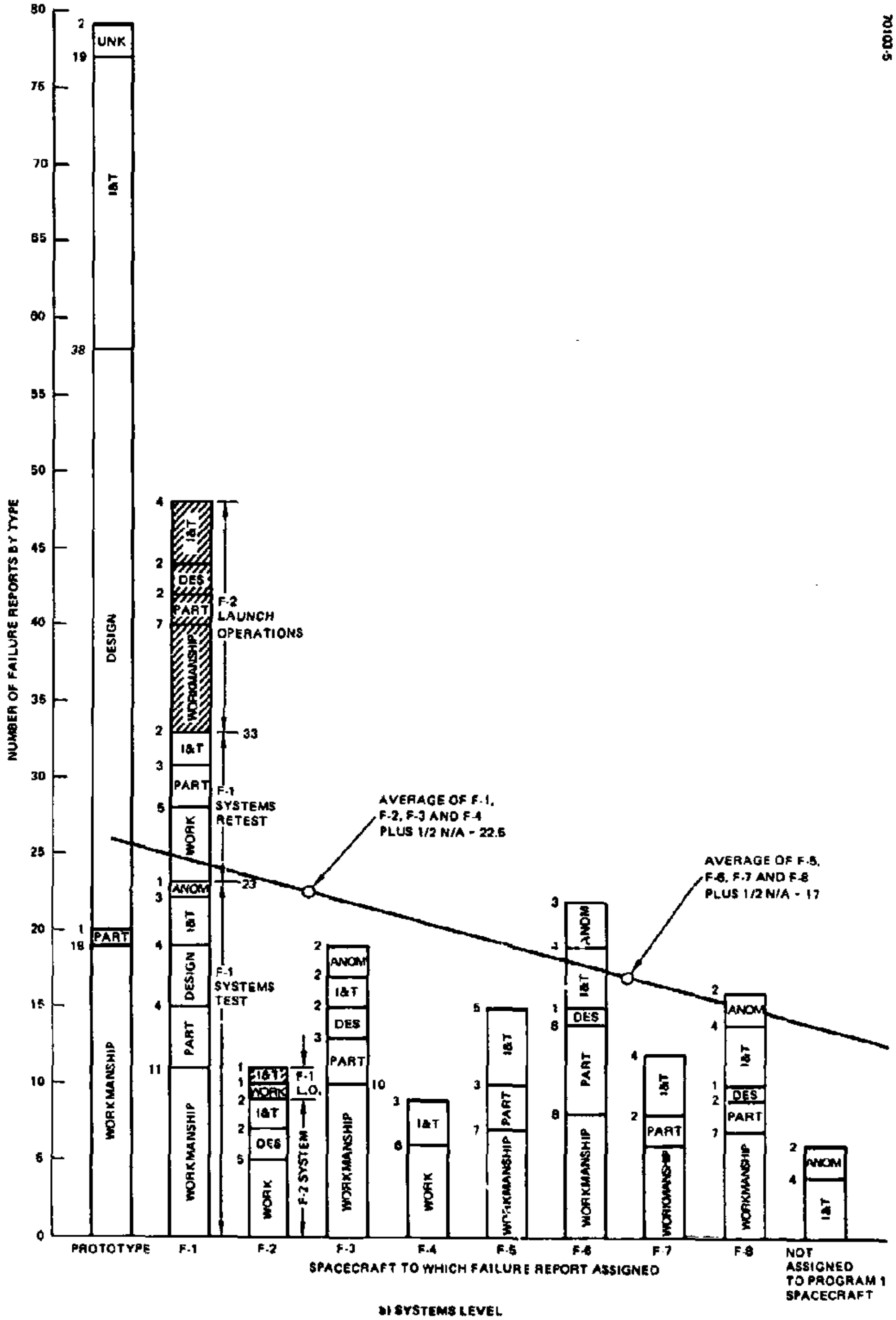
At the systems level, both F-2 spacecraft exhibited relatively low FR rates. In general, workmanship, part, and design FRs were all lower when compared to the preceding spacecraft (F-1) and the following spacecraft (F-3). No viable explanation is apparent. It should be noted that the systems level data of the F-1 spacecraft of Program 1 is a combination of FRs from two complete systems test sequences and the FRs which occurred on the F-2 spacecraft during launch operations (the F-1 spacecraft launch operations problems were arbitrarily assigned to the F-2 spacecraft). As stated earlier, the F-2 spacecraft was the first launched and had an abnormal amount of difficulty during the launch operations, most of which were because of the "first time" experience.

The data from Figures 3-4 and 3-5 have been redrawn on Figure 3-6 in the form of frequency charts. A clearer picture of the spacecraft to spacecraft relationships can be achieved for Program 1 and Programs 2 and 3, respectively. The not assigned FRs have been apportioned between the spacecraft sets.



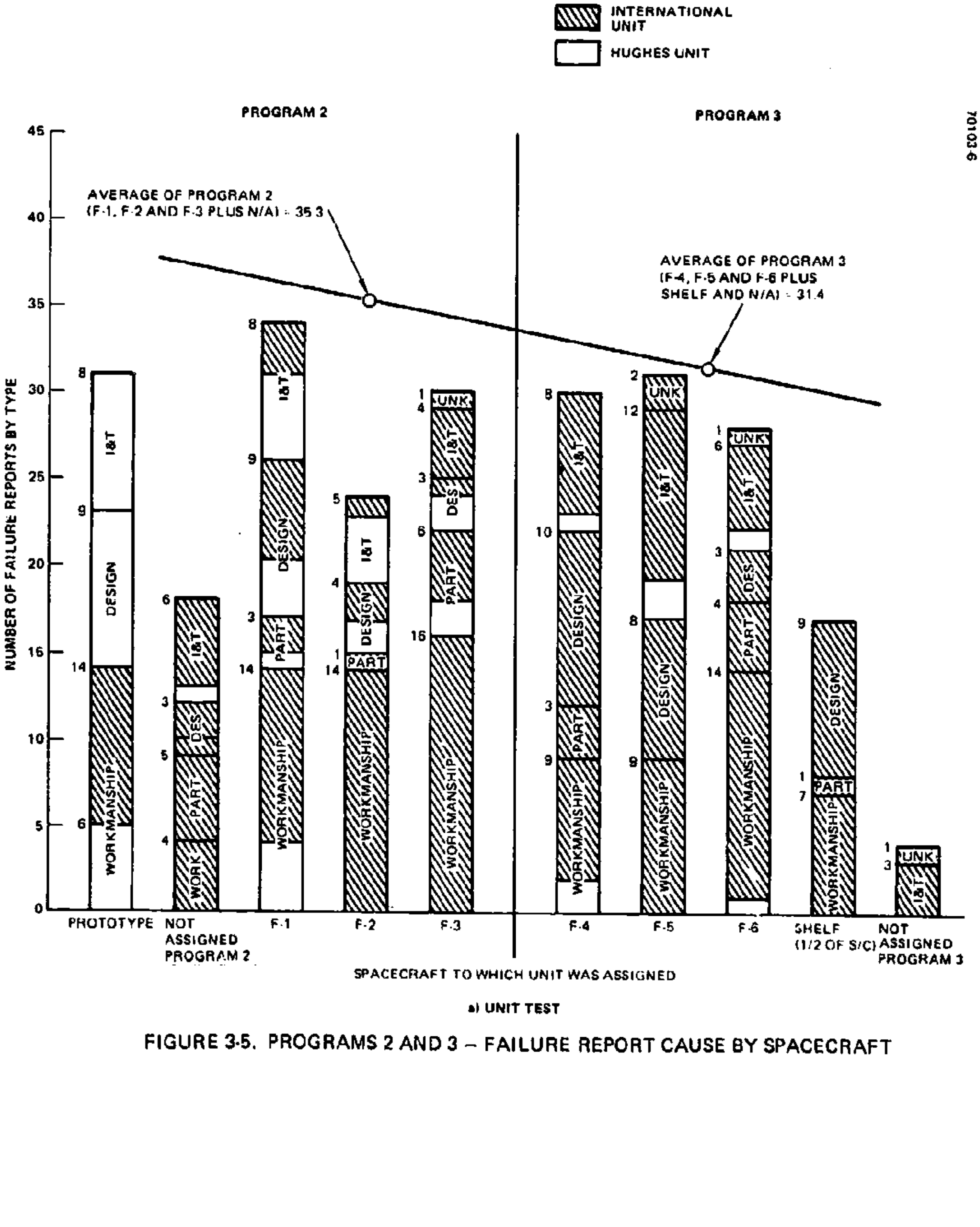
a) UNIT TEST

FIGURE 3-4. PROGRAM 1 – FAILURE REPORT CAUSE BY SPACECRAFT



b) SYSTEMS LEVEL

FIGURE 3-4 (CONTINUED). PROGRAM 1 – FAILURE REPORT CAUSE BY SPACECRAFT



a) UNIT TEST

FIGURE 3-5. PROGRAMS 2 AND 3 – FAILURE REPORT CAUSE BY SPACECRAFT

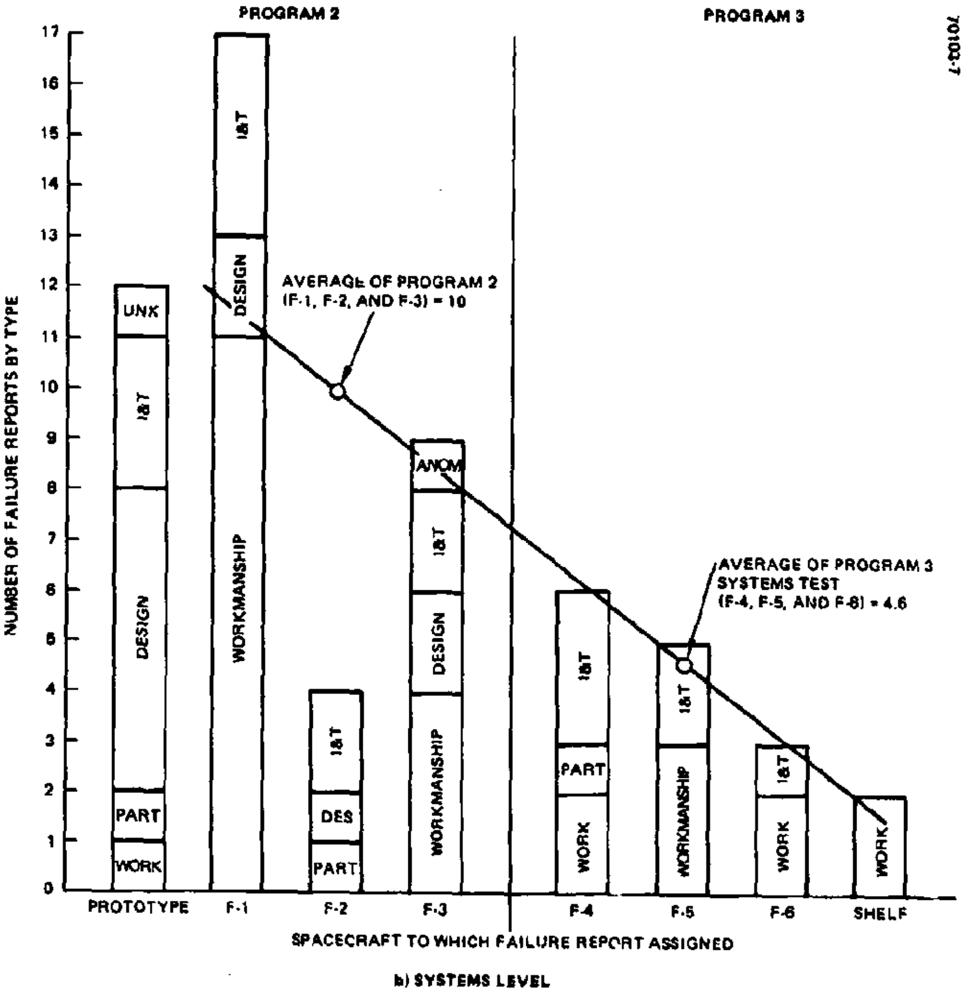


FIGURE 3-5 (CONTINUED). PROGRAMS 2 AND 3 – FAILURE REPORT CAUSE BY SPACECRAFT

FOLDOUT FRAME

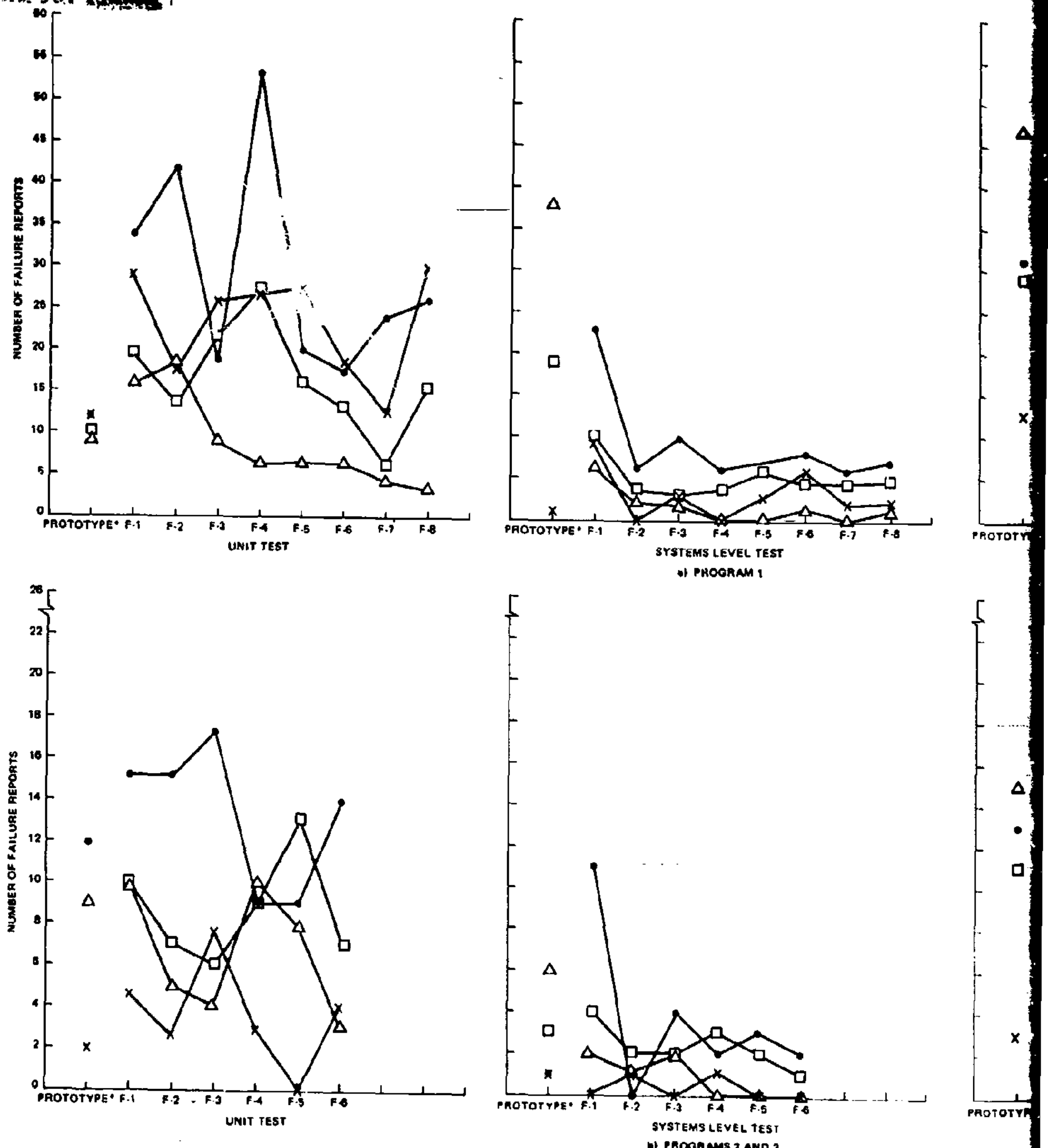
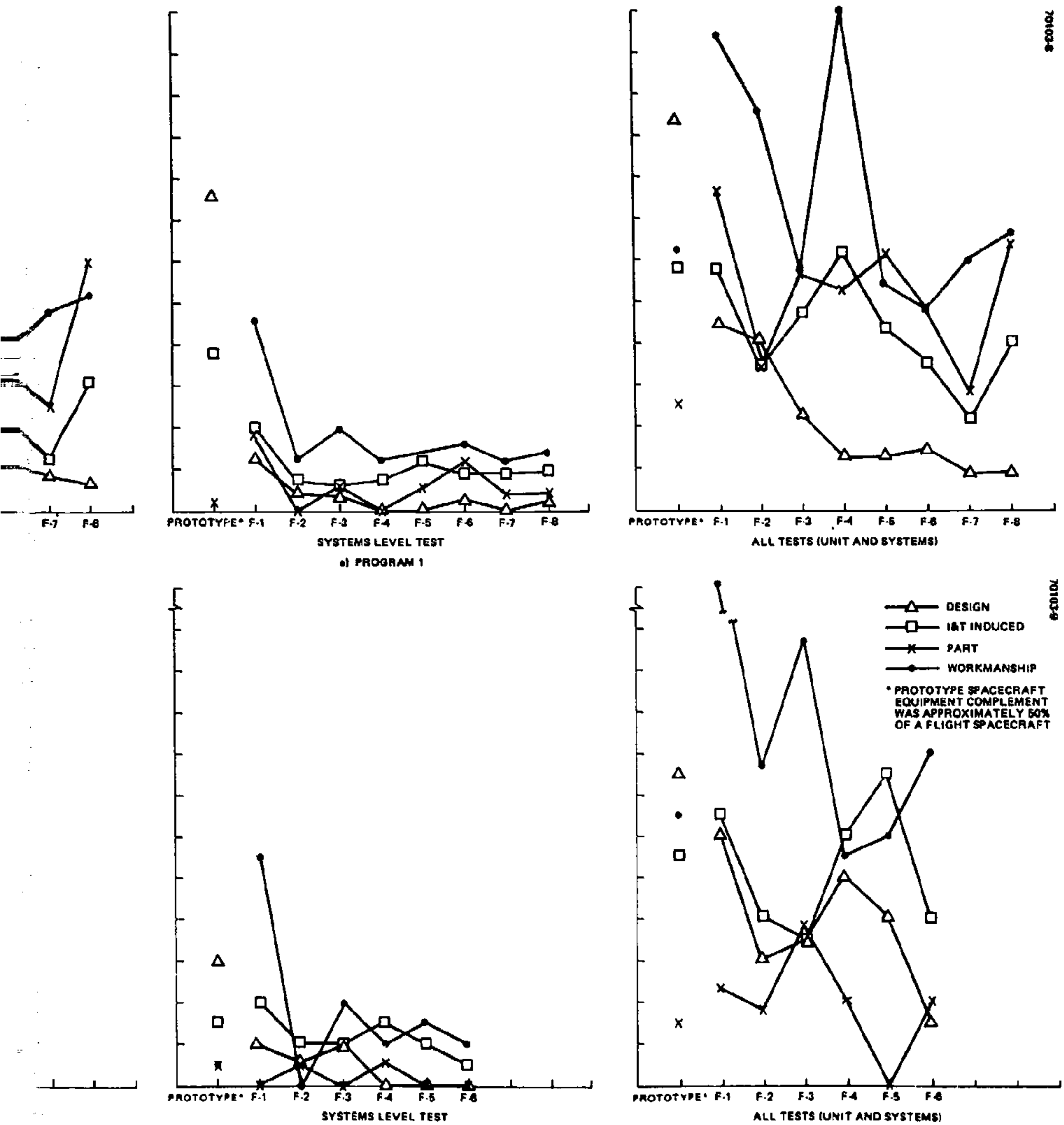


FIGURE 3-6. FAILURE REPORT CAUSE BY SPACECRAFT

FOLDOUT FRAME 2



...RE REPORT CAUSE BY SPACECRAFT

The apparent reduced effectiveness of the qualification test program is due in part to the fact that the equipment complement was only approximately 50 percent of a flight spacecraft. However, it is evident that the Program 1 systems test environment was significantly more effective as a qualification test screen than the unit test environment. The data also indicates that the design defects were not adequately screened during qualification environments, and significant numbers of design problems were revealed in subsequent flight spacecraft acceptance tests.

As might be expected, workmanship FRs are the largest single category in the data base at both test levels for all three programs and for almost every flight spacecraft. The shifting of problems to the last spacecraft in any set is clearly evident for workmanship FRs (see F-4 and F-8 spacecraft for Program 1, F-3 spacecraft for Program 2, and F-6 spacecraft for Program 3). Overall, the trend toward reduced numbers of FRs with each progressive spacecraft is consistent, with the exception of the workmanship FRs which trend upward for the last spacecraft in each set.

The second most important category of acceptance FRs was the I&T induced group. These FRs were significantly high for each spacecraft. For the later spacecraft at the systems level there were almost as many I&T induced failures and problems as the test program revealed in the spacecraft hardware.

Defective part problems occurred throughout the acceptance test programs and did not appear to be affected by spacecraft number or test maturity.

Design FRs followed the expected downward trend, but not to the degree one would expect. The initial spacecraft of each program revealed substantial numbers of design defects (despite the qualification program) which decayed to a relatively constant rate for Program 1 from the 4th to the 8th spacecraft. The F-4 spacecraft did involve some new designs on Program 3; this resulted in a new start of this design FR decay process.

The similarity of data from program to program was verified by the use of the normalization factors presented above. When the primary data is sorted and divided by the number of spacecraft within the set as appropriate, and then normalized, the results are as shown on Tables 3-14 and 3-15 for qualification and acceptance. Program 1 qualification FRs are similar to Program 2. The same is true of the FR per spacecraft average of the first flight spacecraft set for Programs 1 and 2. The last four spacecraft on Program 1 (set 1B) was almost identical in FRs per spacecraft to Program 3. In summary, the data, although derived from three different programs, is consistent when normalized for either qualification or acceptance. Therefore, it was concluded that program and level of development distinctions are not necessary for FR defect analysis purposes.

Table 3-16 shows a summary of all primary FRs by cause. Design FRs accounted for almost 50 percent of all qualification FRs, but only 16.7 percent of the acceptance group. Workmanship accounted for 51.4 percent of the acceptance FRs and 35.2 percent of the qualification group. Summaries segregated by the cause of failure are shown in Tables 3-17, 3-18, and 3-19 for workmanship, part, and design defect types, respectively.

TABLE 3-14. NORMALIZED PRIMARY QUALIFICATION DATA BASE BY SPACECRAFT

	Program	
	1	2
Normalization factor	1 (ref.)	2.50
Raw data:		
Unit	33	23
Systems level	60	9
Normalized data:		
Unit	33	57.5
Systems level	60	22.5
Total	93	79.0

TABLE 3-15. NORMALIZED PRIMARY ACCEPTANCE DATA BASE BY SPACECRAFT

	Program			
	1A	1B	2	3
Number of spacecraft	4	4	3	Unit: 3½* System: 3
Normalization factor	1 (ref)	1 (ref)	2.50	2.44
Raw data/spacecraft:				
Unit	76.5	52.0	27.67	23.14
Systems level	16.75	10.75	7.0	2.67
Normalized data/spacecraft:				
Unit	76.5	52.0	69.18	56.46
Systems level	16.75	10.75	17.50	6.51
Total	93.25	62.75	86.68	62.97

*Shelf units ≅ ½ of a spacecraft

TABLE 3-16. ALL PRIMARY FAILURE REPORTS BY CAUSE

FR Type	Qualification		Acceptance	
	Quantity	Percent of Total	Quantity	Percent of Total
Workmanship	44	35.2	421	51.4
Part	16	12.8	238	29.1
Design	62	49.6	137	16.7
Unknown	3	2.4	23	2.8
Total	125	–	819	–

Table 3-17 dramatically shows that 60 percent of all workmanship FRs occurred during installation and/or assembly for both qualification and acceptance tests. No "learning curve" improvements are apparent. Concentrated effort in this single area may have had a significant impact for both cost and schedule of the programs. Bad electrical connections (welds and solder joints) accounted for almost 16 percent of the acceptance workmanship problems.

Manufacturing process was the largest part FR problem; 43.8 percent for qualification and 30.7 percent for acceptance (possibly higher because of the unknowns 39.9 percent). Part internal contamination was also important (17.2 percent of part FRs) for acceptance. In general, the process problem was significant throughout all three programs.

The summary of design FRs by cause is shown in Table 3-19. Five different types of problems were almost equally important to the qualification program. Circuit design and mechanical installation were followed by manufacturing process errors and innate physical property problems, and then specification errors. Four of these five were still important during the acceptance test program with innate physical property problems becoming dominant. The qualification program appears to have been only marginally effective in screening these physical property problems, along with manufacturing process errors, circuit design errors, and specification errors (in order of importance to the acceptance program).

TABLE 3-17. WORKMANSHIP PRIMARY FAILURE REPORT DEFECT TYPES

Workmanship FR Defect Type	Qualification (35.2 percent)		Acceptance (51.4 percent)	
	Quantity	Percent of Total	Quantity	Percent of Total
Test error	1	2.3	15	3.6
Paper error	5	11.4	19	4.5
Bad electrical connection	1	2.3	66	15.7
Accidential handling	1	2.3	0	0
Misuse	1	2.3	1	0.2
Wrong part used	0	0	16	3.8
Damaged wires/coax	2	4.5	2	0.5
Contamination	5	11.4	26	6.2
Bad alignment	2	4.5	21	5.0
Installation/assembly	26	59.1	255	60.6
Total	44	–	421	–

Note: Usually, accidental handling and damaged wires/coax were I&T induced, not "workmanship."

The primary and I&T acceptance FR data at both unit and systems levels for all programs was classified by type and subsystem to study the variances between subsystems. These data are presented in Table 3-20. Generally, 40 percent of all FRs were workmanship, 22.5 percent part, 13.0 percent design, and significantly, 22.5 percent were I&T induced. By subsystem, both power and propulsion had high numbers of I&T type FRs when compared to the other subsystems (36 and 47.4 percent, respectively). The propulsion subsystem had a very low (5.3 percent) proportion of part problems (no electronics), and the power subsystem had very few design (4 percent) problems. The telemetry and command (T&C) digital subsystems, and the attitude control subsystems, as expected because of a high concentration of electronics, had the highest concentration of part problems -- 33.6 and 34.2 percent, respectively.

TABLE 3-18. PARTS PRIMARY FAILURE REPORT DEFECT TYPES

Part FR Defect Type	Qualification (12.8 Percent)		Acceptance (29.1 Percent)	
	Quantity	Percent of Total	Quantity	Percent of Total
Workmanship	3	18.8	16	6.7
Contamination	3	18.8	41	17.2
Mfg. process	7	43.8	73	30.7
Design	3	18.8	13	5.5
Unknown	0	0	95	39.9
Total	16	–	238	–

TABLE 3-19. DESIGN PRIMARY FAILURE REPORT DEFECT TYPES

Design FR Defect Type	Qualification (49.6 Percent)		Acceptance (16.7 Percent)	
	Quantity	Percent of Total	Quantity	Percent of Total
Wrong part	0	0	1	0.7
EMI	0	0	1	0.7
Physical properties	11	17.7	48	35.0
Specification error	8	12.9	20	14.6
Process error	11	17.7	31	22.6
Circuit design	16	25.8	24	17.5
Mechanical installation	16	25.8	10	7.3
Analytical error	0	0	2	1.5
Total	62	–	137	–

TABLE 3-20. PRIMARY AND I&T ACCEPTANCE FAILURE REPORT DATA BY TYPE AND SUBSYSTEM

Subsystem	Failure Report Type						Approximate Allocation of All Units - o/o
	Unknown	Workmanship	Part	Design	I&T	Totals	
Communications	16(4.0)	163(40.3)	97(24)	45(11.1)	83(20.5)	404(38.2)	49.4
T&C RF	2(1.5)	50(37.9)	24(18.2)	36(27.3)	20(15.2)	132(12.5)	11.0 (Combined)
T&C digital	2(1.3)	61(40.1)	51(33.6)	11(7.2)	27(17.8)	152(14.4)	
Attitude control	2(1.8)	39(35.1)	38(34.2)	12(10.8)	20(18.0)	111(10.5)	3.6
Power	0(0)	28(37.3)	17(22.7)	3(4.0)	27(36.0)	75(7.1)	7.0
Propulsion	1(1.3)	27(35.5)	4(5.3)	8(10.5)	36(47.4)	76(7.2)	13.9
Antenna	0(0)	17(43.6)	3(7.7)	13(33.3)	6(15.4)	39(3.7)	3.6
Spacecraft harness	0(0)	28(58.3)	3(6.2)	5(10.4)	12(25.0)	48(4.5)	11.3 (Combined)
Structural and mechanical	0(0)	6(42.9)	0(0)	3(21.4)	5(35.7)	14(1.3)	
All others	0(0)	2(33.3)	1(16.7)	1(16.7)	2(33.3)	6(0.6)	0.2
Total	23(2.2)	421(39.8)	238(22.5)	137(13.0)	238(22.5)	1057(1000)	100

NOTES: 1) Unit and systems level FRs from all programs
2) ( ) = % of total

The spacecraft harness subsystem was segregated from the structural subsystem because of the high number of FRs generated at the systems level. Of the 110 primary FRs of Program 1 at the systems level, 33 were wire harness (30 percent), and of 38 I&T FRs, 7 (18 percent) were against the harness. The harness group was less significant on Programs 2 and 3, where only two (6.4 percent) primary and two I&T (14 percent) FRs were reported. The total number of FRs per subsystem can be normalized, based upon number of units. As shown in Table 3-2, a weighted calculation indicates that approximately 50 percent of all hardware from which FRs could have been produced were from the communications subsystem. Therefore, 38.2 percent of all programs primary and I&T FRs from the communications subsystem shows a better than average performance. Poor performers include the attitude control, 10.5 versus 3.6 percent; T&C (both) 27.7 versus 11 percent. Other "good" performers include propulsion, 7.2 versus 13.9 percent, and structures, 5.8 versus 11.3 percent.

3.5 ORBIT FR ANALYSIS

All of the reported orbital problems have been coded and classified in a manner similar to that used for the primary FR data base at unit and systems levels. The data has been divided into basic groups: 1) unique problems (those which occurred only once on a program or are the first occurrence of several of similar problems on the same and/or subsequent (time) spacecraft); and 2) duplicate problems (previously reported symptoms with the same determined cause but not necessarily on the same serial numbered spacecraft). A third group called operator error has been added to cover a single problem on Program 1 wherein a ground controller caused a major outage of payload communications through the use of improper procedures. These groupings are summarized on Table 3-21 for the two programs (Programs 2 and 3 are combined). The percentage data shows the similarity between the programs. Approximately 50 percent of all problems are duplicates. This would imply generic defects which cause repeated occurrences of the same problem on several spacecraft within a program. The problems either did not occur or were not identified in time to enable the correction of subsequent spacecraft prior to launch.

Of the unique problems, 52 percent were determined to be design defects. This implies that the test program was relatively ineffective in screening design defects. In fact, the vast majority of the design problems were not tested for during the acceptance test program. These problems involved life, static charge/discharge, orbital alignments, and zero gravity conditions, deficiencies for which the test program was not designed to detect. Therefore, the problem is one of overall test program design. It should also be noted that the data of Table 3-21 reflects almost 44 orbit years of experience on the 12 spacecraft in orbit (see Table 3-1).

The criticality of any residual problem is an important criteria in evaluating the effectiveness of a test program. This was evaluated as a function of this effect and is presented in Table 3-22. Anomalies, periodic/intermittance, and hard failures occurred with the same frequency. All of the catastrophic problems (i.e., loss of spacecraft or payload function of the spacecraft) occurred on Program 1 and were associated with a single type of unit. This was a life/design problem which occurred on a particular lot of units which were all launched prior to detection. Many of the orbit problems were not critical (36 percent), (e.g., loss of a single telemetry channel). Others were operationally overcome by special ground station command sequencing (11 percent). The remaining problems (39 percent) were mission degrading in that some critical function (payload) was lost, impaired, or interrupted.

TABLE 3-21. ORBIT FAILURE REPORTS FOR ALL PROGRAMS

Type	Program		Total
	1	2 & 3	
Unique problems			
Unknown	1(1.5)*	1(4.5)	2(2.2)
Workmanship	2(2.9)	1(4.5)	3(3.3)
Part	13(19.1)	4(18.2)	17(18.9)
Design	18(26.5)	5(22.7)	23(25.6)
Duplicates of above	33(48.5)	11(50.0)	44(48.9)
Operator error	1(1.5)	0	1(1.1)
Total	68	22	90

*Percentage shown in parentheses; the data have not been normalized

TABLE 3-22. ORBIT FAILURE REPORTS SYMPTOM VERSUS SEVERITY FOR ALL PROGRAMS

Severity	Symptom						Total
	Unique Problems			Duplicate Problems			
	Anomaly	Periodic Intermittent	Hard Fail	Anomaly	Periodic Intermittent	Hard Fail	
Not critical	10	5	2	10	4	1	32
Operationally overcome		4			6		10
Mission degrading	5	6	11	5	5	3	35
Catastrophic			2*			10*	12
Total	15	15	15	15	15	14	89**

*All same unit on Program 1

**Operator error on Program 1 not included

4. TEST EFFECTIVENESS

4.1 INTRODUCTION

For this study, test effectiveness has been generally defined as the number of failure reports (FRs) generated in a given test screen to the number of FRs (i.e., defects) which were available to be screened. The analysis was approached from two perspectives. First, "Specific Test Effectiveness" was calculated to assess the effectiveness of a particular test in detecting defects peculiar to that environment (e.g., the effectiveness of unit vibration tests in screening the total FR group which was generated during unit, system, and launch vibration. Second, a "General Test Effectiveness" was calculated to assess the effectiveness of a given test relative to the total FR population without regard to screening eligibility (i.e., all remaining FRs were considered available to the next test environment). Using both of these approaches, comparisons were made between the individual unit and systems level tests. A methodology is outlined to vary the content of the total test program in response to the observed FR rates.

4.2 DATA BASE

The primary FRs for test and in-orbit from Programs 1 through 3 were evaluated and classified in terms of the tests or particular environment where the FRs originated. Test data were taken from the summaries of section 3.3 and in-orbit data were taken from section 3.5. For this analysis, the qualification and acceptance test data were regrouped in the following categories:

Unit Tests	Systems Level Tests
Initial ambient	DCTV/ESTV and initial IST
Vibration	Vibration and SPT
Thermal vacuum/ temperature cycling	Thermal vacuum
Final ambient	Final IST
All unit tests	Launch operations
	All system tests

TABLE 4-1. PROGRAM 1 IN-ORBIT PROBLEM SUMMARY

Problem Classification	Number of Different Problems	Number of Repetitive Problems	Total
Vibration	1	–	1
Thermal vacuum	20	8	28
Not tested for during system test	14	25	39
Totals	35	33	68

The in-orbit data were similarly divided into three categories: 1) vibration related (launch), 2) thermal vacuum, and 3) all problems (totals). The nature of the in-orbit data suggested further categorization. As shown in section 3.5, many of the in-orbit problems were duplications of generic deficiencies. In addition, a significant number of the in-orbit problems could not have been screened by the systems level acceptance test program as defined. A summary of the in-orbit problem data for Program 1 is presented in Table 4-1. The criticality of the in-orbit FRs is also of interest for analysis and, therefore, additional sorts were defined as follows: 1) catastrophic degradation in performance of the communications payload, 2) unique problems, and 3) significant degradation of any spacecraft performance parameter. The total test effectiveness data base for the Specific Test Effectiveness evaluation is presented in Table 4-2. It includes, from all three programs, qualification and acceptance test and in-orbit data on vibration, thermal vacuum/thermal cycling, all tests, and the special categories of in-orbit problems defined above.

4.3 SPECIFIC TEST EFFECTIVENESS

The equations which define each "Specific Test Effectiveness" calculation are summarized in Table 4-3. The results, using the data base of Table 4-2, are presented in Table 4-4.

4.3.1 Qualification Tests

The test effectiveness calculations for qualification tests are limited to the unit level because the total FR population is assumed to be those FRs actually generated in the combined unit and systems level test programs. With reference to Table 4-4, Program 1 unit qualification tests (all tests) were significantly less effective than Programs 2 and 3 (36 versus 72 percent). This is probably due to the increased complexity of Program 1, which resulted in a greater number of relatively subtle design and interface problems. The learning curve relative to test technique is also a likely significant factor. In any event, for a new complex spacecraft, the data suggests that systems level tests will probably be substantially more important than the unit level. From Table 4-2, the Program 1 systems level qualification test program screened more problems than unit level tests by a factor of 1.8 (60 versus 33 FRs). For the combined programs, systems level tests are more important by a factor of 1.2.

TABLE 4-2. SPECIFIC TEST EFFECTIVENESS DATA

Tests	Unit Test FRs						Systems Test FRs						In-Orbit FRs		
	Program						Program						Program		
	1		2 and 3		1, 2 and 3		1		2 and 3		1, 2 and 3		1	2 and 3	1, 2 and 3
	Accept	Qual	Accept	Qual	Accept	Qual	Accept	Qual	Accept	Qual	Accept	Qual			
Vibration*	73	7	10	4	83	11	2	25	3	5	5	30	1	0	1
Thermal vacuum/* temperature cycling	89	5	25	3	114	8	12	8	6	1	18	9	28	2	30
All tests*	472	33	125	23	597	56	85	60	27	9	112	69	68	22	90
• Total communications loss	–	–	–	–	–	–	–	–	–	–	–	–	1**	0	1**
• Unique problems	–	–	–	–	–	–	–	–	–	–	–	–	35	11	46
• Performance degradation (all subsystems)															
All data	–	–	–	–	–	–	–	–	–	–	–	–	40	7	47
Items which could have been screened	–	–	–	–	–	–	–	–	–	–	–	–	25	5	30

*F-6 deleted in Program 1; F-6 and shelf deleted in Programs 2 and 3; indented categories overlap and consequently cannot be summed.

**Only one catastrophic failure of all redundant units

TABLE 4-3. TEST EFFECTIVENESS EQUATIONS

SPECIFIC TEST EFFECTIVENESS*

Test	Equation
Vibration	$TE_1 = \dfrac{\text{Total vibration FRs (unit or systems test)}}{\text{All subsequent vibration FRs (systems test and in-orbit or only in-orbit)} + \text{Total vibration FRs (unit or systems test)}}$
Thermal vacuum/thermal cycling	$TE_2 = \dfrac{\text{Total thermal vacuum/thermal cycling FRs (unit or systems test)}}{\text{All subsequent thermal vacuum/thermal cycling (systems test and in-orbit or only in-orbit)} + \text{Total thermal vacuum/thermal cycling FRs (unit or systems test)}}$
All tests	$TE_3 = \dfrac{\text{All FRs (unit or systems test)}}{\text{All subsequent FRs (systems test and in-orbit or only in-orbit)} + \text{All FRs (unit or systems test)}}$
• Total communications loss	$TE_4 = \dfrac{\text{All FRs (systems test)}}{\text{All subsequent in-orbit FRs} + \text{All FRs (systems test)}}$
• Unique problems	$TE_5 = \dfrac{\text{All FRs (systems test)}}{\text{Number of operational problems revealed} + \text{All FRs (systems test)}}$
• Performance degradation (all subsystems)	$TE_6 = \dfrac{\text{All FRs (systems test)}}{\text{Number of failures which degraded performance or mission} + \text{All FRs (systems test)}}$

GENERAL TEST EFFECTIVENESS*

Test	Equation
Unit ambient	$TE_7 = \dfrac{\text{Total initial ambient unit FRs}}{\text{All subsequent FRs (unit test, systems test, and in-orbit)} + \text{Total initial ambient unit FRs}}$
Unit vibration and SPT	$TE_8 = \dfrac{\text{Total vibration and post vibration performance test FRs}}{\text{All subsequent FRs (unit test, systems test, and in-orbit)} + \text{Total unit vibration and post vibration performance test FRs}}$
Unit thermal vacuum/temperature cycling	$TE_9 = \dfrac{\text{Total unit thermal vacuum/thermal cycling FRs}}{\text{All subsequent FRs (unit test, systems test and in-orbit)} + \text{Total unit thermal vacuum/thermal cycling FRs}}$

*The qualification test effectiveness equations do not include in-orbit failures

Table 4-3 (continued)

Test	Equation
Unit final ambient	$TE_{10} = \dfrac{\text{Total final unit ambient FRs}}{\text{All subsequent FRs (systems test and in-orbit} + \text{Total final unit ambient FRs}}$
Systems test initial ambient DCTV/ESTV and first IST	$TE_{11} = \dfrac{\text{Total initial ambient, DCTV/ESTV and first IST FRs}}{\text{All subsequent FRs (systems test and in-orbit)} + \text{Total initial ambient and DCTV/ESTV and first IST FRs}}$
Systems test vibration and SPT	$TE_{12} = \dfrac{\text{Total systems test vibration and SPT FRs}}{\text{All subsequent FRs (systems test and in-orbit)} + \text{Total systems test vibration and SPT FRs}}$
Systems test thermal vacuum	$TE_{13} = \dfrac{\text{Total systems test thermal vacuum FRs}}{\text{All subsequent FRs (systems test and in-orbit)} + \text{Total systems test thermal vacuum FRs}}$
Systems test final IST and Cape	$TE_{14} = \dfrac{\text{All systems test final IST and Cape FRs}}{\text{All subsequent in-orbit FRs} + \text{All systems test final IST and Cape FRs}}$

DESIGN TEST EFFECTIVENESS

Test	Equation
Unit initial ambient	$TE_{15} = \dfrac{\text{Total initial ambient unit design problems}}{\text{All design problems in unit, systems test, and in-orbit}}$
Unit vibration and SPT	$TE_{16} = \dfrac{\text{Total vibration and post vibration unit design problems}}{\text{All design problems in subsequent tests and in-orbit} + \text{Total unit vibration and post vibration design problems}}$
Unit thermal vacuum/ temperature cycling	$TE_{17} = \dfrac{\text{Total unit thermal vacuum/thermal cycling design problems}}{\text{All design problems in subsequent tests and in-orbit} + \text{Total unit thermal vacuum/thermal cycling design problems}}$
Unit final ambient	$TE_{18} = \dfrac{\text{Total final unit ambient design problems}}{\text{All design problems in subsequent tests and in-orbit} + \text{Total final unit ambient design problems}}$
Systems initial ambient and DCTV/ESTV and IST (first)	$TE_{19} = \dfrac{\text{Total initial ambi[illegible]TV/ESTV and first IST design problems}}{\text{Design problems in subsequent tests and in-orbit} + \text{Total initial ambient, DCTV/ESTV and first IST design problems}}$

Table 4-3 (continued)

Test	Equation
Systems test vibration and SPT	$TE_{20} = \dfrac{\text{Total systems test vibration and SPT design problems}}{\text{All design problems in subsequent tests and in-orbit} + \text{Total systems test vibration and SPT design problems}}$
Systems test thermal vacuum	$TE_{21} = \dfrac{\text{Total systems test thermal vacuum design problems}}{\text{All design problems in subsequent tests and in-orbit} + \text{Total systems test thermal vacuum design problems}}$
Systems test final IST and Cape	$TE_{22} = \dfrac{\text{All systems test final IST and Cape design problems}}{\text{All design problems in-orbit} + \text{All systems test final IST and Cape design problems}}$

TABLE 4-4. SPECIFIC TEST EFFECTIVENESS ESTIMATES (PERCENT)

Tests	Unit Test Effectiveness						System Test Effectiveness			All Tests
	Program						Program			
	1		2 and 3		1, 2 and 3		1	2 and 3	1, 2 and 3	Programs 1, 2, and 3
	Accept	Qual*	Accept	Qual*	Accept	Qual*	Accept	Accept	Accept	Accept
Vibration	96	22	77	44	93	27	67	100	83	99
Thermal vacuum/ temperature cycling	69	39	76	75	70	47	30	75	38	81
All tests	76	36	72	72	76	45	56	55	55	90
• Total communications loss	–		–		–		99	100	99	
• Unique problems	–		–		–		71	71	71	
• Performance degradation (all subsystems)										
All data	–		–		–		68	79	70	
Items tested for	–		–		–		77	84	79	

*Based on total FRs for the test program

For qualification vibration tests, the systems level screen on Program 1 was more important by a factor of 3.6 (25 versus 7 FRs), and 2.7 for the combined programs. For the combined programs, a total of 33 percent of all qualification FRs (unit and systems level) were generated in a vibration test environment. This relatively large fraction (vibration tests are only 12 percent of the total acceptance test FR group) is partly because many structural elements are tested for the first time at systems level.

For qualification thermal tests, the systems level screen on Program 1 was more important by a factor of 1.6, and 1.1 for the combined programs. For the combined programs, a total of 14 percent of all qualification FRs were generated in a thermal test environment (thermal tests are 19 percent of the total acceptance test FR group).

The effectiveness of the qualification test screen relative to lesign problems is discussed in Section 4.5.

4.3.2 Acceptance Tests

The Specific Test Effectiveness calculations for acceptance tests are based on the entire FR history for the flight spacecraft, including in-orbit performance to date. The F-6 spacecraft for both programs was excluded from this data base because no in-orbit data exists (the booster failed for F-6 spacecraft on Program 1; and the F-6 spacecraft of Program 3 has not been launched.) The effectiveness of the acceptance test screens was relatively consistent across the programs with a combined program test effectiveness of 75 percent for unit tests, 56 percent for systems tests, and 89 percent for the entire test program.

The vibration acceptance test screen was highly effective at 93 percent for unit tests, 83 percent for systems tests, and 99 percent for the entire test program. Only one vibration problem occurred for 12 spacecraft launches. A total of only 12 percent of the acceptance test FRs occurred in a vibration test. The data strongly suggests that a three-axis vibration test at spacecraft level is probably not required. For all systems level vibration tests (including F-6 spacecraft), a total of only six FRs were generated, with five occurring in the post-vibration performance test (SPT) and one occurring during the initial vibration test which is along the Z-axis (see Table 3-12). The risk associated with limiting systems level acceptance vibration testing to a single axis is probably negligible.

The thermal acceptance test screens were 70 percent effective at the unit level, but only 38 percent effective at systems level. The combined test programs were 81 percent effective. With reference to the Program 1 in-orbit FR summary in Table 4-1, note that 28 (41 percent) in-orbit problems occurred which could have been detected in either unit or systems level thermal testing. These problems included telemetry instrumentation errors, circuit design deficiencies, noisy earth sensor bolometers, and RF anomalies. However, a total of 39 (57 percent) of the in-orbit problems could not have been screened by the test program because of test technique deficiencies. These included static discharge problems, early wearout of life limited items, thermal deformation of structures, ground station operator errors, and zero-gravity effects. Generally, the deficiencies in test technique resulted from either the omission of tests such as static discharge or TWT life tests, or failure to test units in this entire operating mode. Therefore, the relatively low effectiveness of the systems level thermal vacuum environment is probably a result of test technique deficiencies and not total test time in that environment. The limited effectiveness of increased test time in the thermal vacuum environment is presented in more detail in section 6 of this report.

Relative to problems which are critical to the communications satellite mission, the system acceptance test screens were highly effective at 99 percent for the combined programs (see Table 4-4). The single catastrophic problem occurred on the initial spacecraft of Program 1 and was the result of an early wearout mechanism which could not have been screened by the acceptance test program as designed. Programs 2 and 3 have experienced no failures which could constitute catastrophic degradation.

The effectiveness of the systems test screens relative to FRs with some in-orbit performance implication was 70 percent for the combined programs (see Table 4-4), using all FRs. If the FRs which could not have been screened are excluded, the systems test effectiveness increased to 79 percent for performance related failures.

4.4 GENERAL TEST EFFECTIVENESS

The equations which define each "General Test Effectiveness" calculation are also summarized in Table 4-3. The results, using the data base of Table 4-5 are presented in Tables 4-6 and 4-7. Figure 4-1 shows the relative test effectiveness of unit and systems level tests and the first year in orbit.

TABLE 4-5. GENERAL TEST EFFECTIVENESS DATA BASE

Tests	Number of FRs			
	Qualification Tests		Acceptance Tests	
	Program		Program	
	1	2	1	2 and 3**
Unit				
Ambient (initial)***	19	14	304	102
Vibration and SPT	7	4	89	10
Thermal vacuum/temperature cycling	5	3	99	31
Ambient (final)	1	2	10	2
Systems				
Initial ambient, DCTV/ESTV, and IST (initial)	27	3	65	17
Vibration and SPT	25	5	3	3
Thermal vacuum	7	0	9	2
IST (final) and Cape (acceptance only)	1	1	33	7
In-orbit problems – 10 year projection*	–	–	91	64

*The total failure estimates are based on the predicted 10 year failure rate for individual spacecraft. When failure rate curves were not available, the failure rate was estimated by considering the relative number of in-orbit failures and the predicted failure rates for spacecraft which were launched in the same period.

**The shelf is excluded

***Performance tests not identified as to time of occurrence have been included herein.

NOTE: RCS tests have not been included.

TABLE 4-6. GENERAL TEST EFFECTIVENESS OF QUALIFICATION AND ACCEPTANCE TESTS

Tests	Test Effectiveness, Percent			
	Qualification Tests		Acceptance Tests*	
	Program		Program	
	1	2	1	2 and 3
Unit				
Ambient (initial)	21	44	43	43
Vibration	11	22	22	7
Thermal vacuum/cycling temperature	8	21	32	25
Ambient (final)	2	18	5	2
Combined unit tests	35	72	71	61
Systems				
Ambient, DCTV/ESTV, and IST (initial)	–	–	32	18
Vibration and SPT	–	–	2	4
Thermal vacuum	–	–	7	3
IST (final) and Cape (acceptance only)	–	–	27	10
Combined systems tests	–	–	55	31
Combined unit and systems tests	–	–	87	73

*General unit ambient acceptance test effectiveness was calculated with the following equation: TE_7 = 304/703 - 0.43 (43 percent). All other acceptance test effectivenesses were calculated in the same way.

TABLE 4-7. SUMMARY OF ACCEPTANCE TEST SCREENING EFFECTIVENESS FOR COMBINED PROGRAMS

Tests	Number of FRs	Remaining Defect Population	Test Effectiveness, Percent	Percent of Test FRs	Percent of Systems Test FRs
Unit					
Initial ambient	406	942	43	51.6	–
Vibration	99	536	18.5	12.6	–
Thermal	131	437	29.9	16.6	–
Final ambient	11	306	3.6	1.4	–
Unit Total	647	942	68.7	82.2	–
Systems					
DCTV	11	295	3.7	1.4	7.9
Initial IST	71	284	25	9.0	50.7
Vibration and SPT	6	213	2.8	0.8	4.3
Thermal vacuum	11	207	5.3	1.4	7.9
Final IST and Cape	41	196	20.9	5.2	29.3
Systems total	140	295	47.5	17.8	100
All tests total	787	942	83.5	100	–
Orbit (1 year estimate)	112	155	72.3	–	–
Orbit total (10 years)	155	155	100	–	–
Grand total (14 spacecraft)	942	–	–	–	–

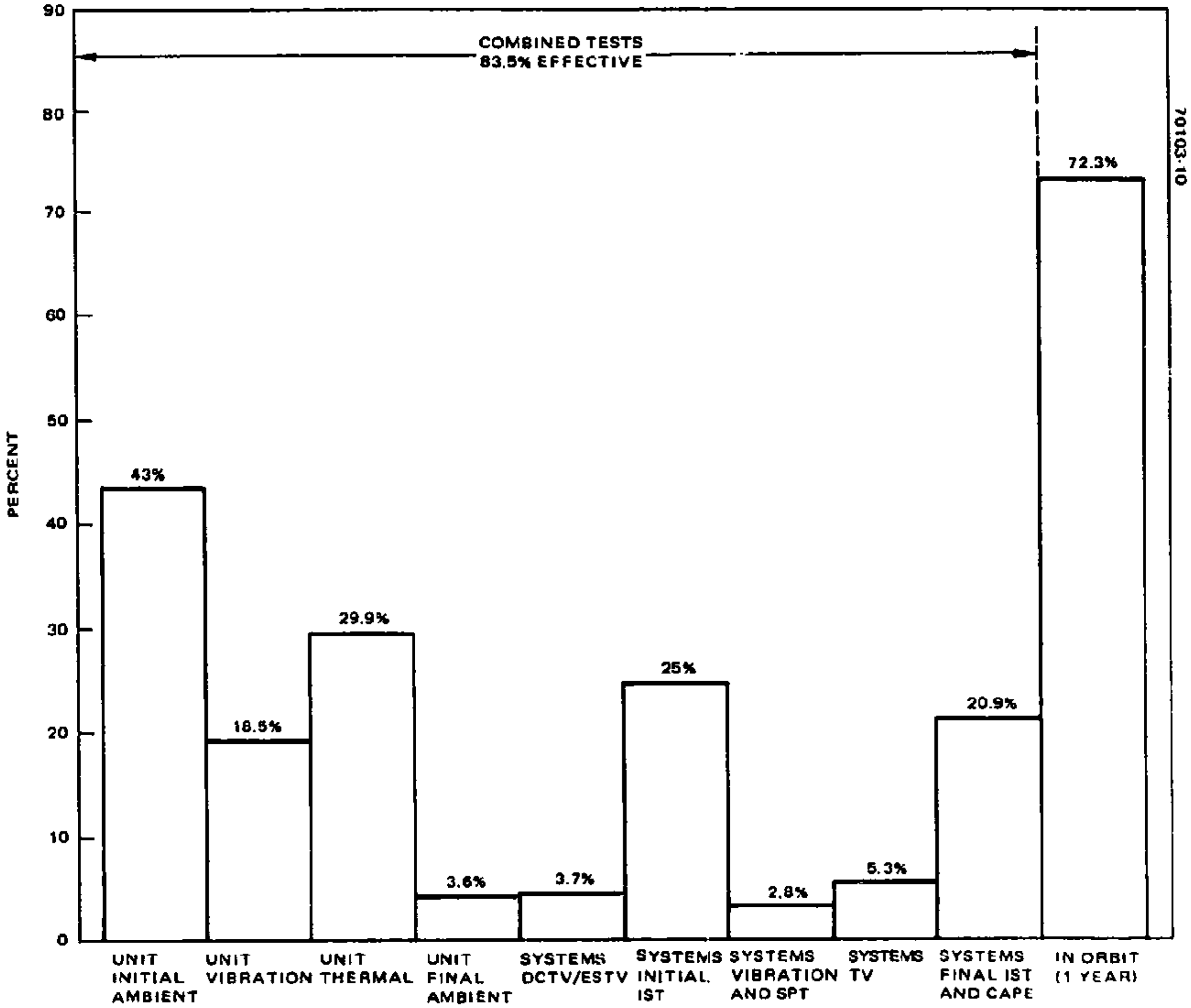


FIGURE 4-1. ACCEPTANCE TEST SCREENING EFFECTIVENESS FOR COMBINED PROGRAMS

4.4.1 Qualification Tests

The test effectiveness calculations for qualification tests are limited to the unit level because the total FR population is assumed to be those FRs actually generated in the combined unit and systems level test program. With reference to Table 4-6, the Program 2 combined unit qualification tests were significantly more effective than Program 1 (72 versus 35 percent). This probably is a result of the higher complexity of Program 1 which resulted in a greater number of subtle design and interface problems combined with the learning curve relative to test technique. The Program 2 systems level prototype spacecraft tests benefitted from the Program 1 test experience. However, for a new complex spacecraft, the data suggests that systems level prototype spacecraft tests will probably be substantially more important than the unit tests by almost a factor of two (60 versus 32 FRs).

The unit ambient test screen was the single most effective unit test in both programs (21 percent for Program 1 and 44 percent for Program 2). The unit qualification vibration test was the second most effective unit test and was 11 percent effective in Program 1 and 22 percent effective in Program 2. The thermal vacuum test was more effective for Program 2 (21 percent) than for Program 1 (8 percent). The final ambient test was only 2 percent effective in Program 1, whereas it was 18 percent effective in Program 2.

4.4.2 Acceptance Tests

The general test effectiveness calculations for acceptance tests are based on the entire FR history for the flight spacecraft, including the predicted in-orbit performance for 10 years. The ground test FRs include the F-6 spacecraft data. A predicted 10 year failure estimate has been included in the in-orbit failure estimate to correct for the loss of the F-6 spacecraft. Table 4-5 describes this calculation procedure. The Programs 2 and 3 in-orbit total failure estimate has been similarly corrected to include a prediction of the yet unlaunched F-6 spacecraft in-orbit failures. The effectiveness of the acceptance test screens was relatively consistent across the programs with a test effectiveness for units of 71 percent in Program 1 and 61 percent in Programs 2 and 3; 55 and 31 percent, respectively, for systems level tests, and 87 percent for Program 1 and 73 percent for Programs 2 and 3 for the entire test program.

The Programs 1 through 3 unit initial ambient tests were the most effective tests (43 percent) performed at the unit level (see Figure 4-1). The unit thermal tests were the second most effective tests with 29.9 percent (32 percent Program 1 and 25 percent Programs 2 and 3). The data indicate that these tests provide the best opportunity for defect detection and correction.

The systems level ambient, DCTV/ESTV, and initial IST provide the next most effective defect detection test. The test was 32 percent effective on Program 1 and 18 percent effective on Programs 2 and 3, with 28 percent overall.

The systems level final IST and cape tests were 20.9 percent effective (27 percent on Program 1 and 10 percent on Programs 2 and 3). The unit vibration test was 22 percent effective in Program 1 and only 7 percent effective in Programs 2 and 3. The combined test effectiveness of the unit final ambient, systems level vibration, and systems thermal vacuum tests ranged between 2.8 to 5.3 percent. The data indicate that these tests are not very effective in screening defects and failures. They suggest that reduced vibration and thermal vacuum testing at the spacecraft acceptance level could be implemented without a significant effect on the fraction defective in a delivered spacecraft.

4.5 TEST EFFECTIVENESS FOR DESIGN PROBLEMS

The equations which define each "Design Test Effectiveness" calculation are also summarized in Table 4-3. The results, using the data base of Table 4-8, are shown in Table 4-9.

4.5.1 Qualification Tests

The test effectiveness calculations for the ability of the qualification tests to identify design deficiencies are based on the FR history for all spacecraft (prototype and flight), and includes in-orbit performance to date. Only design problems have been included. The design test effectiveness for the Program 1 qualification test program was 32 percent and for Program 2 it was only 23 percent. The data indicate that the tests were only partially successful in detecting design deficiencies.

The Program 1 unit qualification tests were only 6 percent effective in detecting design problems. Program 2 was 14 percent effective. The Program 1 combined systems level tests were 27 percent effective as contrasted with Program 2, which was only 11 percent effective. The Program 1 systems level vibration test had the highest test effectiveness, with 16 percent, and was the single most important qualification test for the identification of design deficiencies.

The above data show that unit and systems level qualification tests were marginally effective in detecting design problems. Overall, only 47 out of 148 design problems were detected in the unit and systems level test qualification program of Program 1. On Program 2, 15 out of 66 were detected. This suggests that insufficient qualification testing was performed on both programs. In the future, more extensive testing, in terms of environmental and operational extremes, should be performed. Possibly, the duration of the qualification tests should be extended in an attempt to screen out additional design deficiencies and avoid delays to the acceptance test programs due to design problems.

TABLE 4-8. PROGRAMS 1 THROUGH 3 DESIGN PROBLEMS SUMMARY BY TEST PHASE

Spacecraft	Unit Test: Initial Ambient**	Unit Test: Vibration	Unit Test: TV/TC	Unit Test: Final Ambient	Systems Test: Initial Ambient IST and DCTV/ESTV	Systems Test: Vibration and SPT	Systems Test: TV	Systems Test: Final IST	Launch Operations	Orbit Unique Problems
Program 1										
Qual	4	0	4	1	14	20	4	0	–	–
F-1	6	3	7	0	3	0	0	1	0	0
F-2	9	0	8	1	3	0	1	0	0	12
F-3	7	0	2	0	0	1	1	0	0	3
F-4	4	0	2	0	0	0	0	0	0	1
F-5	5	0	1	0	0	0	0	0	0	2
F-6	2	2	2	0	1	0	0	0	0	0
F-7	4	0	0	0	0	0	0	0	0	0
F-8	2	0	1	0	0	0	0	0	1	0
Unassigned	3	0	0	0	0	0	0	0	0	0
Subtotal	46	5	27	2	21	21	6	1	1	18
Programs 2 and 3										
Qual	5	1	2	1	3	3	0	0	–	–
F-1	7	1	1	0	0	0	1	1	0	2
F-2	2	1	1	0	1	0	0	0	0	3
F-3	3	0	0	0	2	0	0	0	0	0
F-4	4	1	4	1	0	0	0	0	0	0
F-5	6	1	1	0	0	0	0	0	0	0
F-6	2	0	1	0	0	0	0	0	-	--
Shelf	1*	0	0	0	–	–	–	–	–	
Unassigned	3	0	0	0	0	0	0	0	0	0
Subtotals	33	5	10	2	6	3	1	1	0	5
Totals	79	10	37	4	27	24	7	2	1	23

*The 9 FRs written against a single capacitor have been considered a single problem.
**The unit FRs where the test phase unknown were included in the initial ambient unit test.

4.5.2 Acceptance Tests

Because of the relatively high numbers of design FRs on the early spacecraft, the acceptance design test effectiveness calculations were based on the ability of the first three flight spacecraft of Programs 1 and 2 to identify those design deficiencies which were undetected by the qualification programs. The combined unit and systems level test were 49 percent effective. Therefore, approximately 75 percent of all design deficiencies were found after the testing of the first four spacecraft (prototype and three flight).

TABLE 4-9. GENERAL TEST EFFECTIVENESS FOR DESIGN PROBLEMS

Tests	Design Test Effectiveness, Percent: Qualification, Program 1	Qualification, Program 2	Acceptance, Programs 1 and 2: F-1, F-2, and F-3 Spacecraft
Unit			
Ambient	3	8	22
Vibration	0	2	4
Thermal vacuum/temperature cycling	3	3	17
Ambient	1	2	1
Combined unit test	6	14	39
Systems			
Ambient, DCTV/ESTV, and IST (first)	10	5	10
Vibration and SPT	16	6	1
Thermal vacuum	4	0	4
IST (final)	0	0	3
Combined systems test	27	11	16
Combined unit and systems test	32	23	49

The unit tests were more effective in detecting design deficiencies than the systems level tests (39 versus 16 percent). The single most effective test was the unit ambient (22 percent). Unit thermal tests were 17 percent effective and the systems level initial ambient, DCTV/ESTV, and initial IST were 10 percent effective. The remaining tests were very ineffective (1 to 4 percent).

The first three flight spacecraft unit and systems level acceptance tests were surprisingly more effective in detecting design problems than were the individual unit and systems level qualification programs. The data tend to suggest that an important factor in detecting design problems is the amount of time spent in testing. What is important in detecting design deficiencies appears to be the full exercising of units and the spacecraft in a variety of environmental and operational conditions. The proper conclusion may be that an effective qualification test program is one in which a spacecraft is fully exercised in changing environments and operational modes (novel conditions) until such time as the test data indicate that the time between the detection of design problems (mean time to failure) is such that it is not cost effective to continue the testing.

4.6 RELIABILITY PREDICTIONS USING A TEST EFFECTIVENESS MODEL

The concept of Test Effectiveness (TE)* assumes the existence of a fixed initial defect population which is steadily reduced by successive test screens of consistent effectiveness (i.e., the test screen effectiveness is independent of the defect population size; and all members of the population are eligible for the screen). The test screens are not required to have equal effectiveness. The TE model (Equation 1) has the property that the total defects screened by "n" environments are independent of the sequence of environments.

Let:

T = total number of defects in the initial population

f_n = number of defects screened in environment "n"

F_N = total defects screened by the combination of "n" environments

E_n = test effectiveness of environment "n"

then:

$$E_n \triangleq \frac{f_n}{T - \sum_{n=1}^{n-1} f_n}$$

or

$$f_n = E_n \left(T - \sum_{n=1}^{n-1} f_n \right) \qquad (1)$$

The TE model also has the property that the number of defects screened in any particular environment (or group of environments) will be linearly related to the number of previously screened defects as follows:

Let:

E_1 = effectiveness of test group 1

E_2 = effectiveness of test group 2

T = total initial defect population

f_1 = defects screened in test group 1

f_2 = defects screened in test group 2

---

*A.F. Timmins, A Study of Relationship between Performance in Systems Tests and Space," Proceedings of the Institute of Environmental Sciences, 1975, 1.172

then

$$f_1 = E_1 T$$

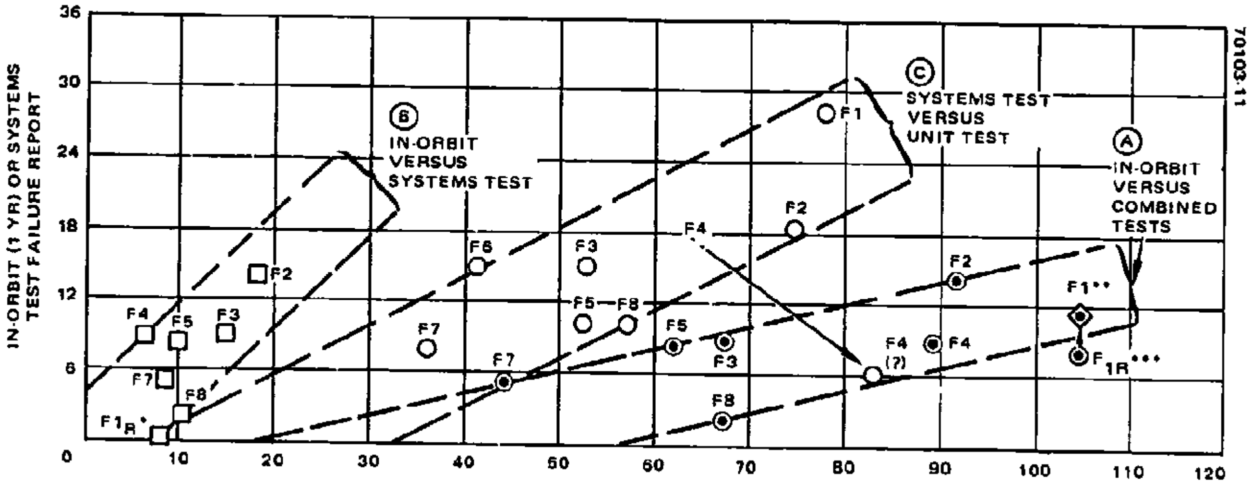
$$f_2 = E_2 (T - f_1) = E_2\left(\frac{f_1}{E_1} - f_1\right) = \left(\frac{E_2}{E_1} - E_2\right) f_1$$

Therefore, a convenient check of the applicability of the TE model for prediction purposes is to evaluate the relationship between the in-orbit failure group for a fixed period (1 year was chosen) and the systems level test group or the combined systems, and unit test groups. A linear relationship would be expected. Similarly, a linear relationship should exist between the systems level test failure group and the unit test failure group.

These data are presented in Figure 4-2. Data Set "A" illustrates the relationship between in-orbit FRs and the combined tests FR group. Similarly, "B" and "C" compare systems level test FRs with in-orbit and unit level, respectively. The FR data is summarized in Table 4-10. The linear relationships predicted by the TE model are apparent. The unit data indicates a bias of approximately 20 FRs which apparently have no reliability significance. The F4 and F8 spacecraft appear unusual but may only be indicative of the statistical variance. The defect analysis data presented earlier in this report (see Figure 3-4) also indicated F4 and F8 were atypical. A more detailed analysis of the data base is required.

As mentioned earlier in this report, the F1 spacecraft was retrofitted to correct all known design problems and was exposed to a complete retest at systems level prior to launch. Given the test history of F-1 (105 FRs; see Table 4-10) and the estimated effectiveness of the Program 1 combined test environments (0.87, see Table 4-6), a total of 8.6 failures would have been predicted for the retest program. In fact, a total of 8 did occur. This data is shown as $F1_R$ in Figure 4-2. The in-orbit (1 year) screening effectiveness is greater than systems level test (0.72 versus 0.55) and, therefore, 11 failures would have been expected if F1 had been launched without the retest program. This is in good agreement with Figure 4-2.

In summary, the data indicates that a TE model could be used as the basis for reliability predictions and, therefore, could be used to adjust the content of a test program as required to achieve an acceptable residual defect group. If the number of test failures were too high for a particular spacecraft, the test program would be continued until acceptable FR rates are achieved. The F-1 retest program is a case in point. Consistent failure reporting criteria would be required and statistical limits for the TE model would have to be developed. The results of this study indicate that such an approach is feasible.



*F1R – THERE WERE NO FAILURES IN THE FIRST YEAR OF ORBITAL OPERATIONS ON THE F-1 RETESTED SPACECRAFT

**THE F1 WAS PREDICTED TO HAVE 11 IN-ORBIT FAILURES DURING THE FIRST YEAR IN ORBIT WITHOUT RETEST

***THE F_{1R} ACTUALLY HAD 8 FAILURES DURING RETEST

FIGURE 4-2. PROGRAM 1 FAILURE RATE GROUPS COMPARISON

TABLE 4-10. SUMMARY OF PRIMARY FAILURE REPORTS FOR PROGRAM 1
(Unassigned FRs Excluded)

Program 1 Spacecraft	Total Unit FRs	Total Systems FRs	Combined Total Test FRs	Systems Level Retest FRs	In-orbit* FRs (1 year)
F-1	77	28	105	8	0
F-2	74	18	92	–	14
F-3	52	15	67	–	8.8
F-4	83	6	89	–	8.8
F-5	52	10	62	–	8.2
F-6	41	15	56	–	–
F-7	36	8	44	–	5.3
F-8	57	10	67	–	2
Totals	472	110	582	8	47.1
Average per spacecraft	59	13.75	72.8	1	6.7**

*The number of in-orbit FRs for the first year in orbit was estimated from the orbital Duane plots in Figures 5-28 through 5-98 or 5-16 for the F-8 spacecraft.

**The average was based on seven in-orbit spacecraft since the F-6 booster failed.

5. FAILURE RATES AS A FUNCTION OF TIME

The system level and in-orbit failure data was analyzed to assess time dependent trends and relationships. In particular, the applicability of the Duane reliability growth model was evaluated.

Each of the systems level and in-orbit primary failure reports (FRs) have been associated with their respective time of occurrence to establish the required data base. The test time-to-failure was used for the systems level FRs and was estimated from either the individual spacecraft test logs or test reports. The time recorded was only an indication of the time-to-failure, since spacecraft "off" time during test reconfiguration, third shift, weekends, etc., was not extracted. The in-orbit times were those reported by the customers at the time the event occurred. By definition, in-orbit time is current to the date of this report (March 1, 1977). Because systems test times varied from a low of 4752 test hours to a high of 12,288; and accumulated in-orbit time varies between 15,593 and 53,173 hours on the different spacecraft of the three programs, it is difficult to combine all data. Therefore, analysis was basically made by individual spacecraft. A summary of accumulated time for systems level test (launch), in-orbit to date (3-1-77), and combined is provided in Table 5-1. The total data base represents over 12 years of systems level test experience and almost 44 in-orbit years.

Where sufficient data were available (e.g., three or more FRs), three different graphs were prepared for each spacecraft as applicable. The first, graph A, was a Duane plot of the systems level test data. Graph B, the second, was a Duane plot of the in-orbit report data, using each launch date as the initialization point. The third, graph C, is a combined systems test and in-orbit plot, using the start of systems test as the initialization point and assuming that the in-orbit data was contiguous to the end of systems test (launch). It was not analyzed rigorously with the Duane model, but was presented in the same format.

Scatter diagrams of the $\log_{10}$ cumulative failure rate versus $\log_{10}$ of the cumulative time of each event were plotted. A weighted linear regression was used to fit least squares estimates for the appropriate Duane equations. The weighting scheme used for the linear regressions is to give the first failure of each spacecraft a weight of 1, the second failure a weight of 2, the

TABLE 5-1. SYSTEMS TEST, IN-ORBIT, AND COMBINED ACCUMULATED TIMES (HOURS)

Spacecraft	Systems Level Test (Launch)	In-Orbit to 3-1-77	Combined (If Applicable)
Program 1			
Prototype	7,248 (EOT)*	N/A	N/A
F-1	12,288	15,593	27,881
F-2	5,184	53,173	58,357
F-3	5,592	45,552	51,144
F-4	4,752	44,764	49,516
F-5	6,552	41,347	47,899
F-6	6,700	N/A	N/A
F-7	5,832	30,835	36,667
F-8	7,848	19,973	27,821
Program 2			
Prototype	7,536 (EOT)	N/A	N/A
F-1	6,384	37,756	44,140
F-2	6,864	33,813	40,677
F-3	8,808	15,943	24,751
Program 3			
F-4	5,832	25,228	31,060
F-5	5,328	20,936	26,264
F-6	5,088 (EOT)	N/A	N/A
Totals	107,836 (12.3 years)	384,913 (43.9 years)	466,177 (53.2 years)

*All times in hours; EOT = End of test

nth failure a weight of n, etc.; thus, giving the last event the most weight. All data is assumed to fit the Duane analysis model of the form:

$$\lambda = \frac{\Sigma F}{T} = KT^{-\alpha}$$

where:

λ = cumulative failure rate

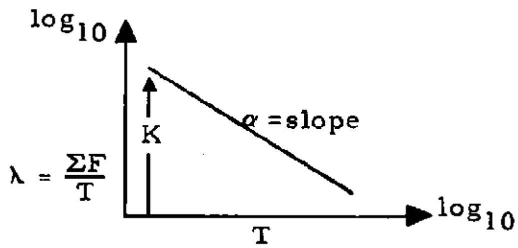
T = total time in hours

F = failure events during time T

K = constant determined by circumstances

α = growth rate (slope)

When plotted on on $\log_{10}$ - $\log_{10}$ paper, the model appears as a straight line of the form:



Therefore, equation (1) reduces to:

$$\log_{10}\lambda = \log_{10}K - \alpha\log_{10}T \tag{2}$$

and regression fitting will determine the constant K and slope α.

The failure rate at time t_i is λ_i; this equates to instantaneous failure rate for any event as:

$$\lambda_i = \int_o^T \frac{F_i\,dt}{T} = KT^{-\alpha} \tag{3}$$

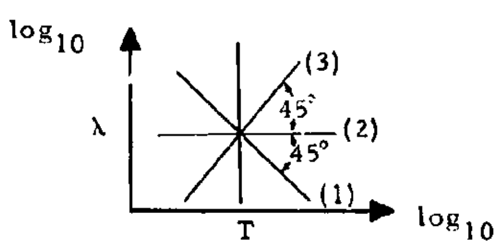
where:

F_i = instantaneous fails per unit time at time t_i

Therefore:

$$F_i = K\ (1-\alpha)\ t_i^{-\alpha} \tag{4}$$

Using the above expressions, the Duane model can be interpreted for three different conditions:



For condition 1, $\alpha = +1$; therefore, $F_i = 0$ and $\lambda_i = K/T$. This is a limiting condition and $\alpha \leq +1$.

For condition 2, $\alpha = 0$; therefore, $F_i = K$ and $\lambda_i = K$ (i.e., the failure rate is a constant, not influenced by time.).

For condition 3, $\alpha = -1$; therefore, $F_i = 2KT$ and $\lambda_i = KT$. For this case, the failure rate linearly increases with time. This implies either an increasing test screen effectiveness or that the test article is degenerating as a result of test (i.e., the defect population is increasing with time).

An understanding of these boundary conditions will assist in the evaluation of the Duane plots.

Figures 5-1 through 5-15 contain the scatter diagrams for each spacecraft on Programs 1 through 3 where sufficient data existed. Figures A and B contain the weighted least squares Duane estimate curves. Figure C contains only visual estimations of the curve (or curves) to fit the data, using straight lines. For this combined data no meaningful weighted least squares Duane estimate could be found for most spacecraft because of the marked contrasts in slope between early systems test and in-orbit data. A time reference from Table 5-1 has been provided on each graph. The operational time-to-date for B figures has been indicated but was not used as a data point for purposes of determining Duane curves; only data points corresponding to the occurrence of failure events were utilized in the Duane analysis.

Table 5-2 contains the parameter estimates derived for each Duane analysis shown in the figures. This table also contains the number of data points utilized to derive each Duane curve estimate. It should be noted that where two or more failure events occurred at the same point in time, the total cumulative failure rate was calculated as one data point.

A review of the spacecraft Duane curves and Table 5-2 reveals that systems level failure rates vary widely as a function of time. This is best seen in Figures 5-18a and 5-19a, wherein all systems level Duane curves have been redrawn on one figure for Program 1 and Programs 2 and 3, respectively. The slopes shown appear to represent all three Duane curve bounding conditions described above. Early spacecraft trend toward negative slopes ($-\alpha$) and later spacecraft approach the $\alpha = +1$ limit. The final failure rate of systems level test (λ_{fst}) does tend to converge for all flight spacecraft (especially on Program 1) and gets lower with each succeeding spacecraft (fewer FRs per spacecraft, therefore, lower failure rate).

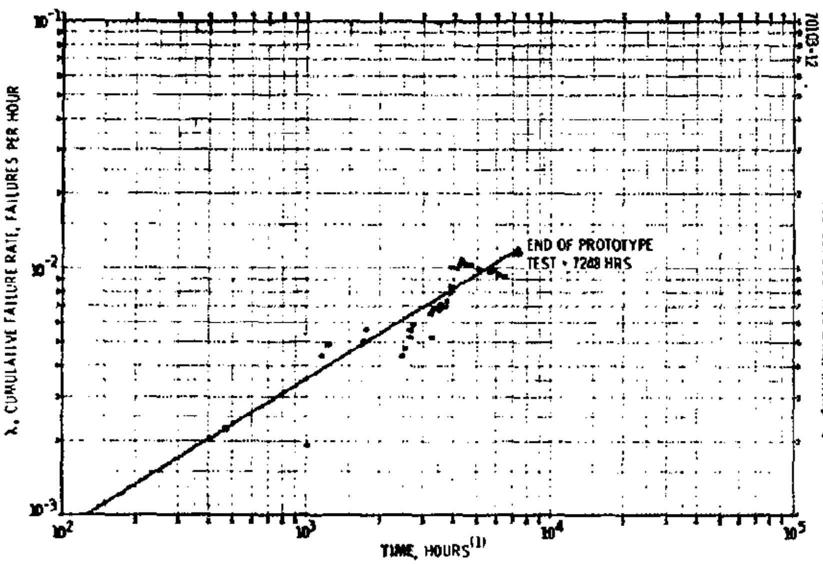
When all of the in-orbit Duane curves are compared and plotted on Figures 5-18b and 5-19b, it is shown that all slopes are similar (positive α) and form a family of curves, again with the property that successive spacecraft generally have lower failure rates at any given time in orbit. There are three special cases for the in-orbit data. Both Program 1 F-1 spacecraft and Program 2 F-3 spacecraft have had no orbital anomaly reports in over 15,500 hours of operation, and, therefore, have not been shown. The F-5 spacecraft of Program 3 has not had an orbit problem in its last 11,000 hours of operation. The F-5 spacecraft also has the Duane curve with a negative slope ($-\alpha$). If one assumes a failure occurring on 1 March 1977 for this spacecraft, a worst case estimation line could be drawn through the calculated failure rate at 20,936 hours. As a failure has not in reality occurred, the actual slope should be even more positive than that drawn on Figure 5-15b. This estimator would appear to depict the F-5 spacecraft in-orbit performance more realistically than the Duane curve and was used in Figure 5-19b.

TABLE 5-2. SYSTEMS TEST AND IN-ORBIT DUANE CURVE PARAMETER ESTIMATES

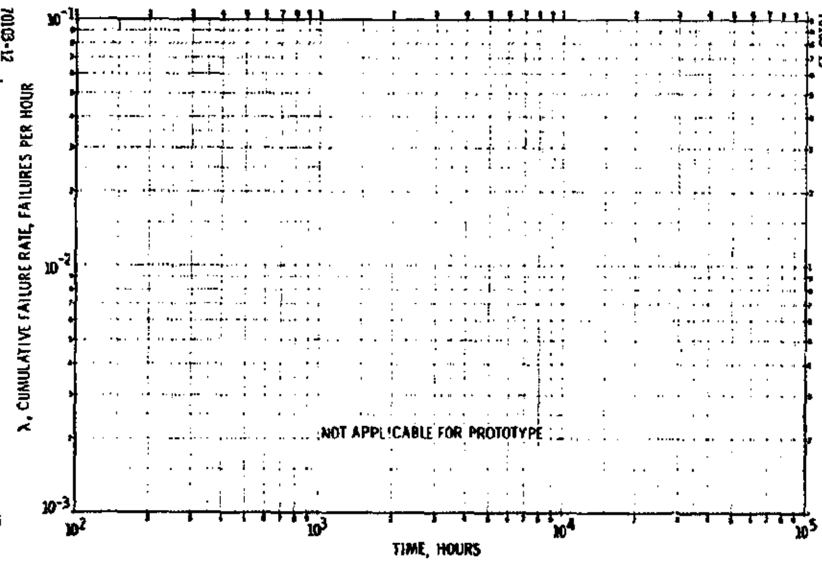
Spacecraft	Systems Test			In-Orbit		
	No. of Data Points	Estimate for K-Constant	Estimate for α-slope	No. of Data Points	Estimate for K-Constant	Estimate for α-slope
Program 1						
Prototype	33	$5X10^{-5}$	-0.61	N/A	N/A	N/A
F-1	23	0.512	0.58	None	Insufficient data	
F-2	11	$1X10^{-10}$	-2.03	12	3.09	0.83
F-3	15	0.0275	0.23	9	1.74	0.82
F-4	6	0.00195	-0.06	8	2.29	0.85
F-5	9	0.00041	-0.16	12	0.955	0.76
F-6	13	0.457	0.61	N/A	N/A	N/A
F-7	7	0.155	0.55	8	0.309	0.69
F-8	10	0.145	0.54	3	Insufficient data	
1-Composite	113	0.126	0.19	45	6.16	0.78
Program 2						
Prototype	8	0.00141	0.02	N/A	N/A	N/A
F-1	11	0.0562	0.37	5	0.00093	0.16
F-2	2	Insufficient data		6	0.0933	0.58
F-3	6	0.089	0.53	None	Insufficient data	
Program 3						
F-4	3	$9.5X10^{-45}$	-11.36	3	Insufficient data	
F-5	3	0.209	0.67	6	0.000323	-0.06
F-6	2	Insufficient data		N/A	N/A	N/A
2-3 Composite	32	0.0537	0.26	18	0.0115	0.26

Note: 1) All equations are assumed to be of the form: $\lambda = KT^{-\alpha}$.

2) Number of data points are not equal to the number of events as some events occurred on the same day.

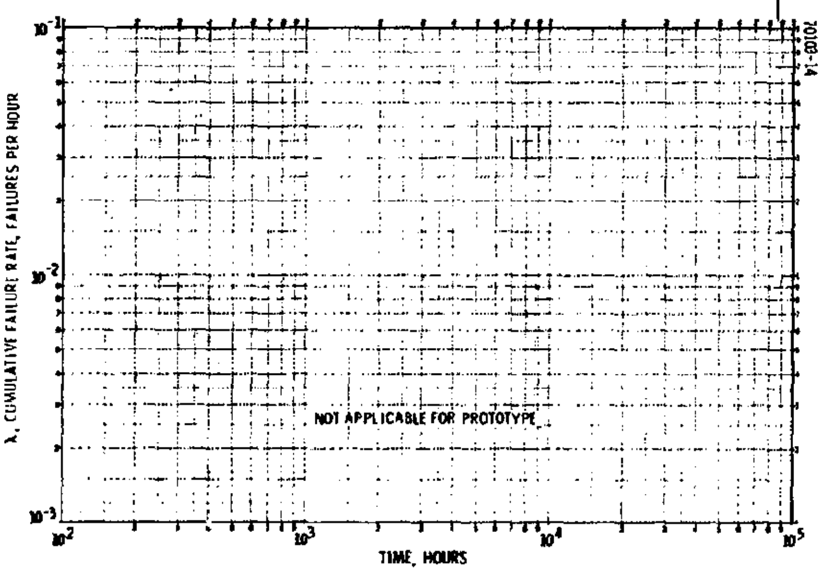


a) SYSTEMS TEST



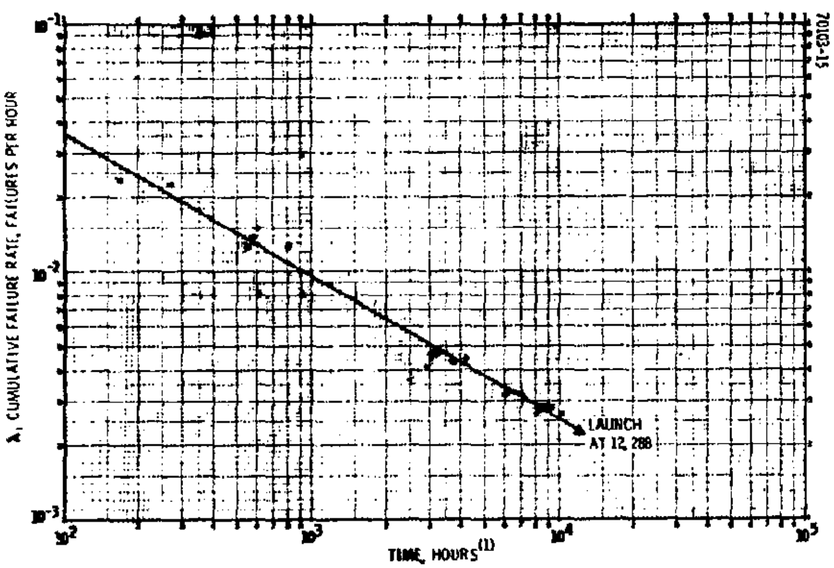
b) IN-ORBIT

(1) REPRESENTS SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME

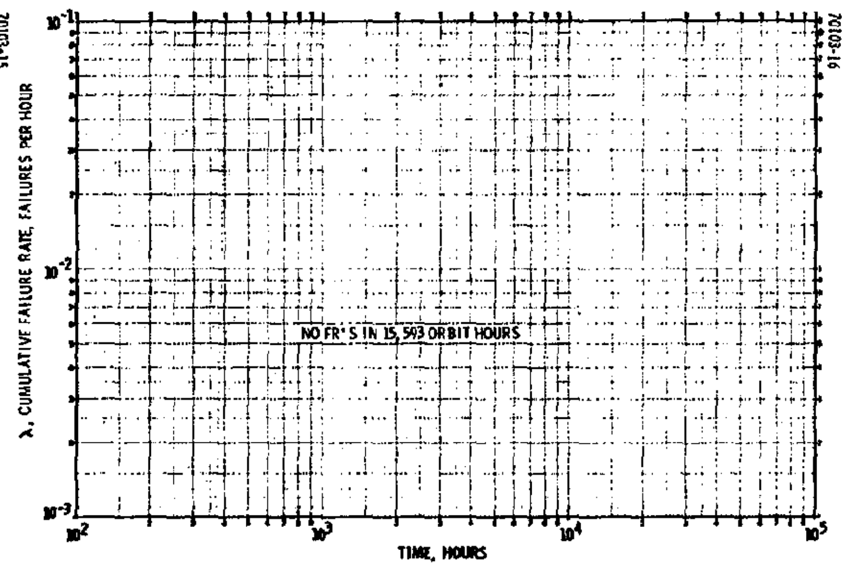


c) SYSTEMS TEST AND IN-ORBIT COMBINED

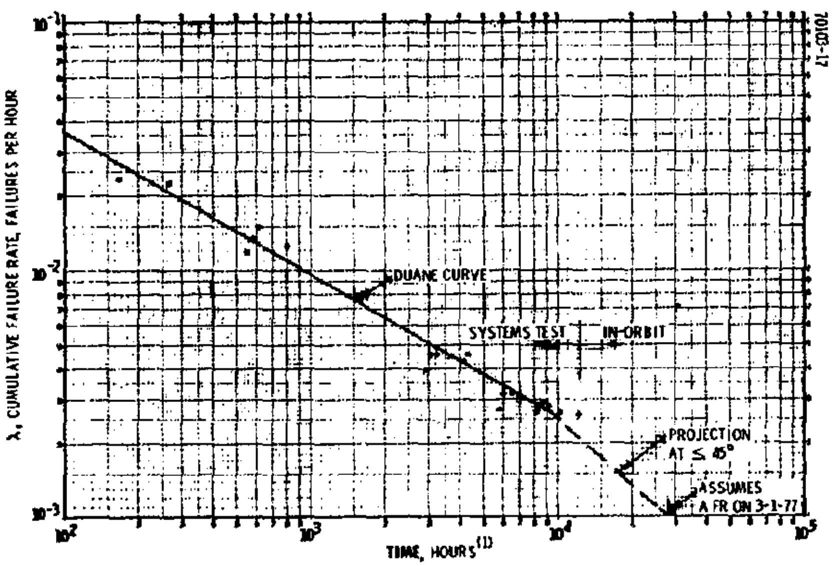
FIGURE 5-1. PROGRAM 1 PROTOTYPE DATA



a) SYSTEMS TEST



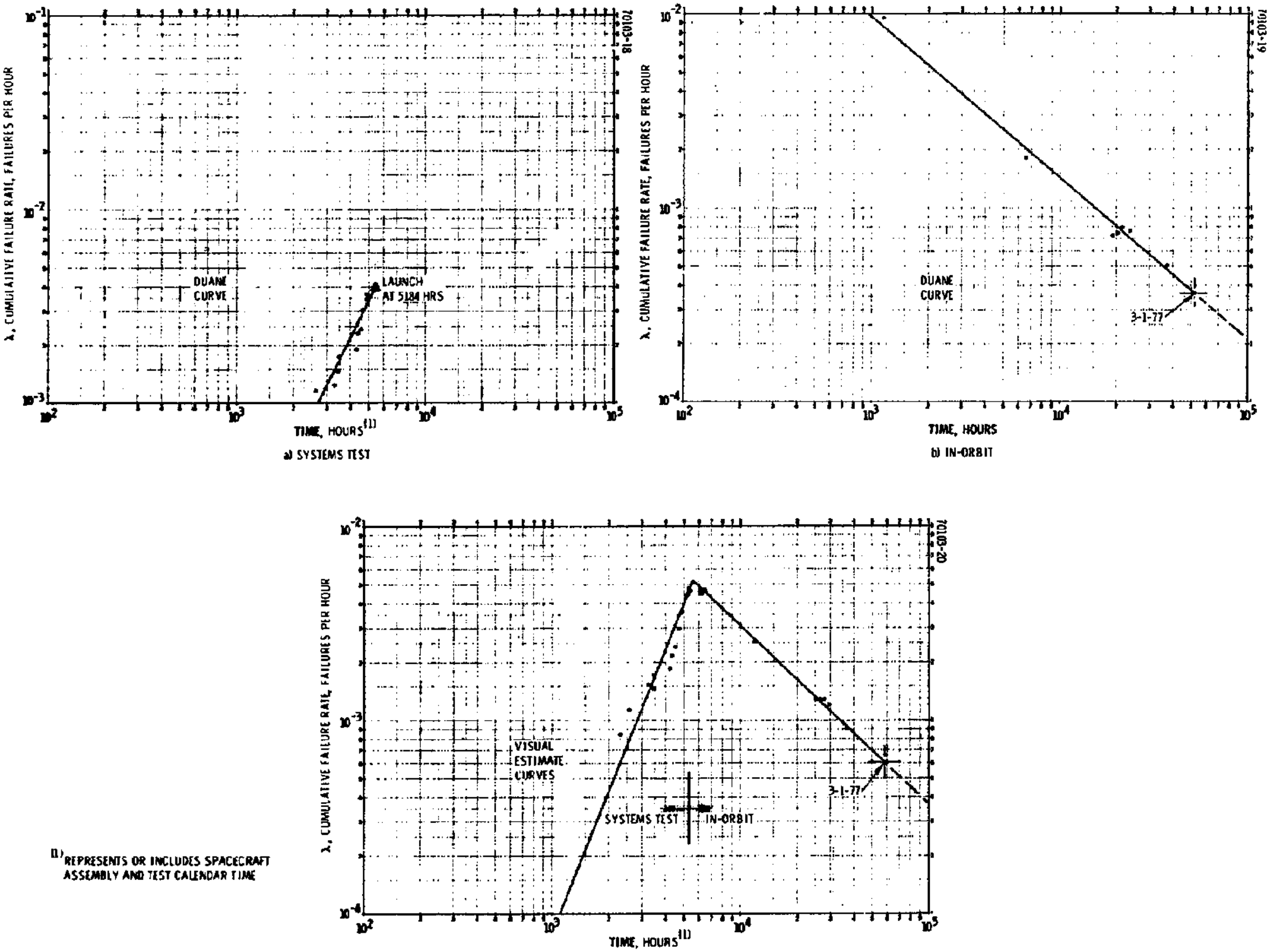
b) IN-ORBIT



c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-2. PROGRAM 1 F1 DATA

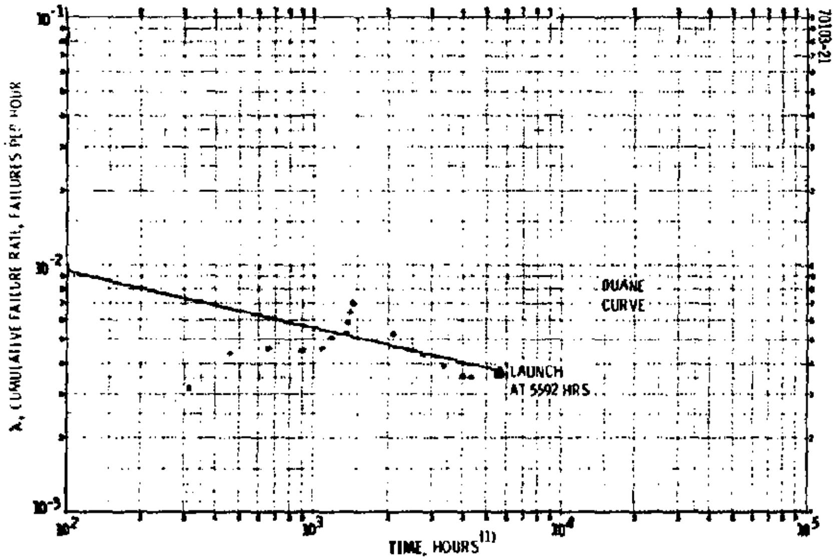
(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



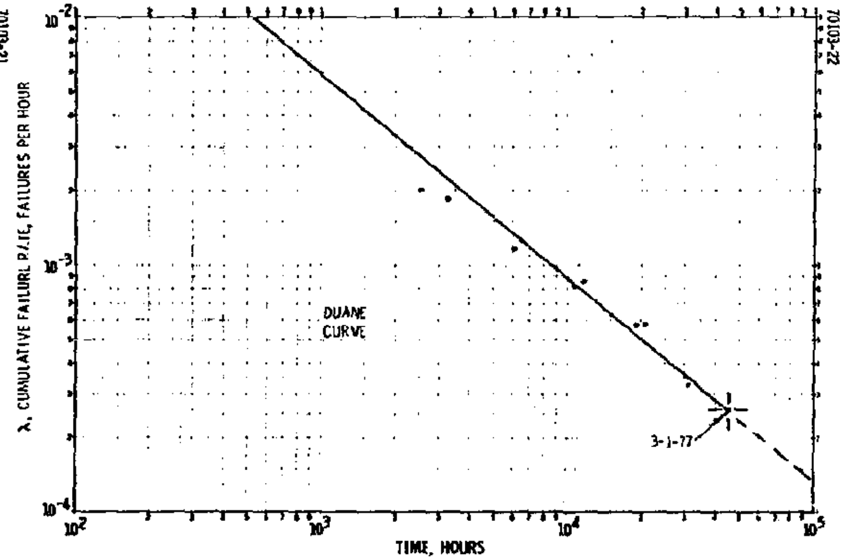
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-3. PROGRAM 1 F2 DATA

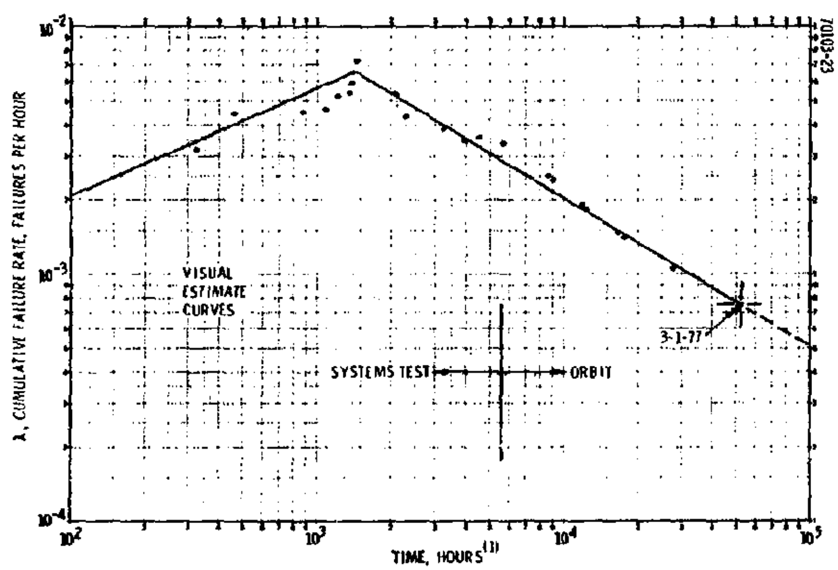
(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



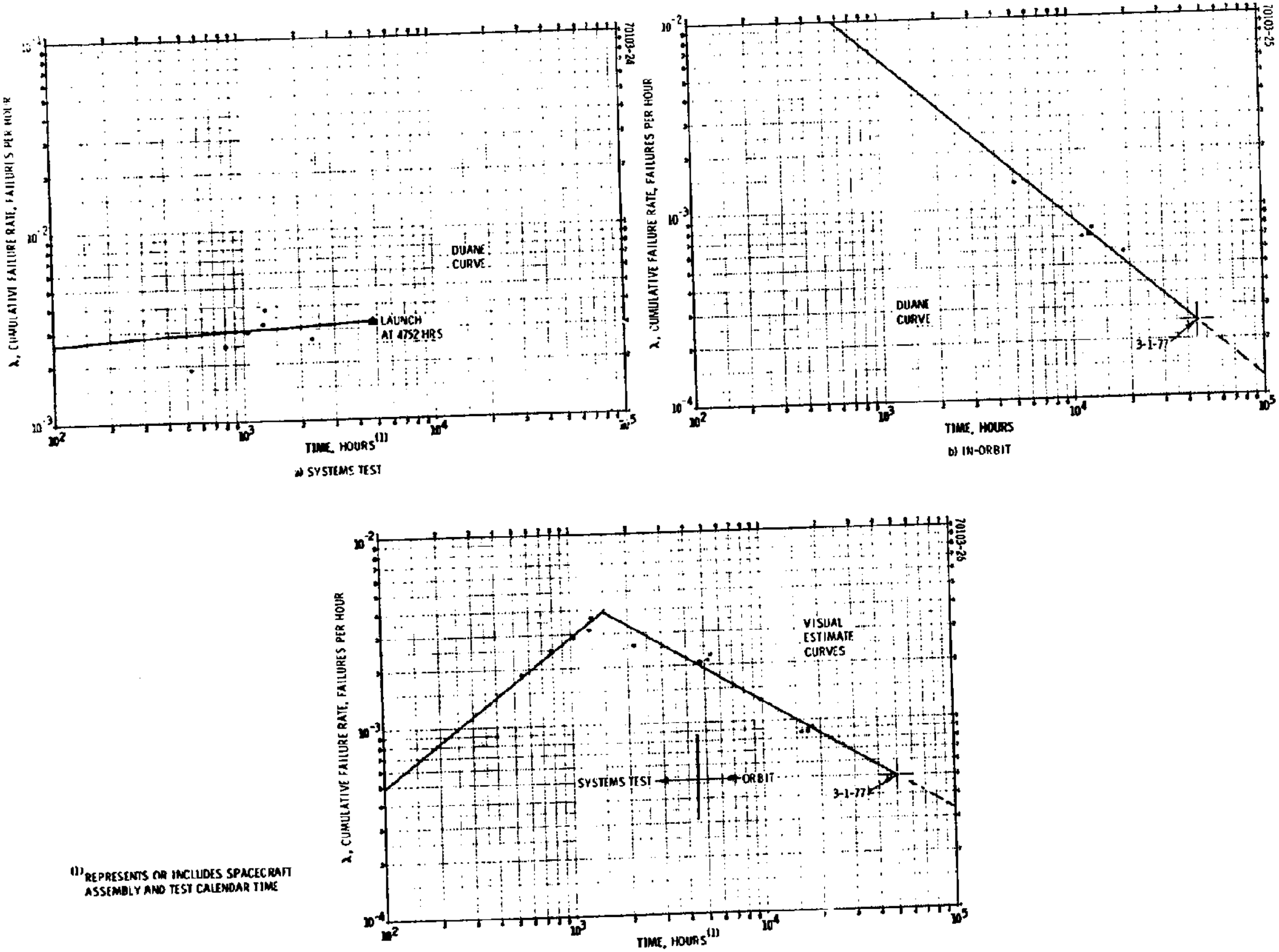
b) IN-ORBIT



c) SYSTEMS TEST AND IN-ORBIT COMBINED

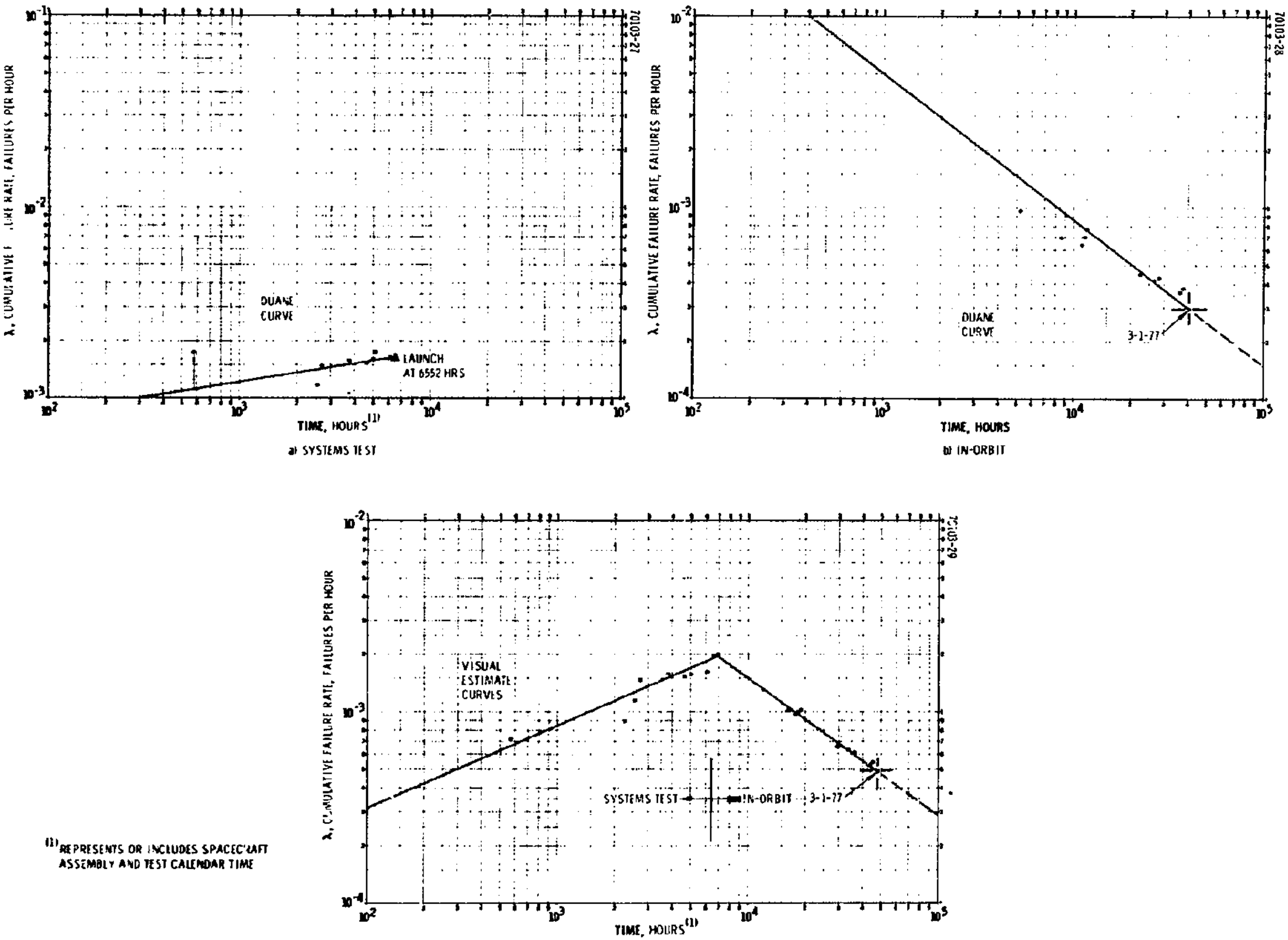
FIGURE 5-4. PROGRAM 1 F3 DATA

(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME

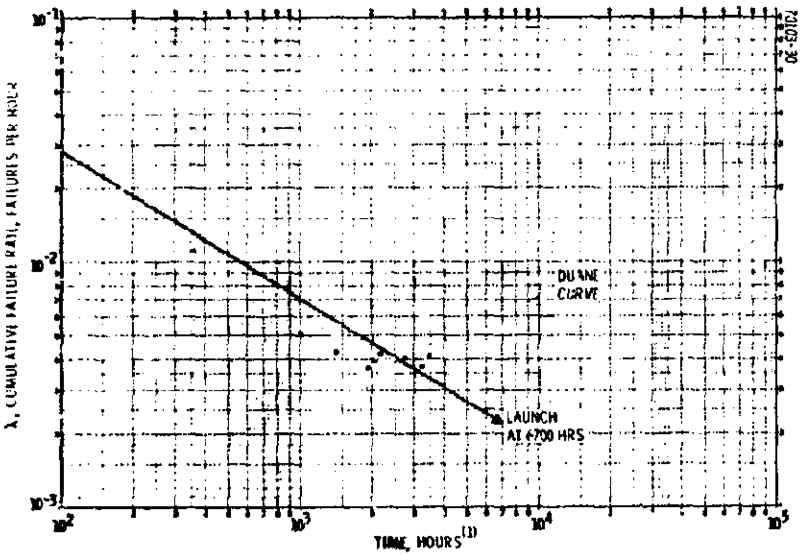
FIGURE 5-5. PROGRAM 1 F4 DATA



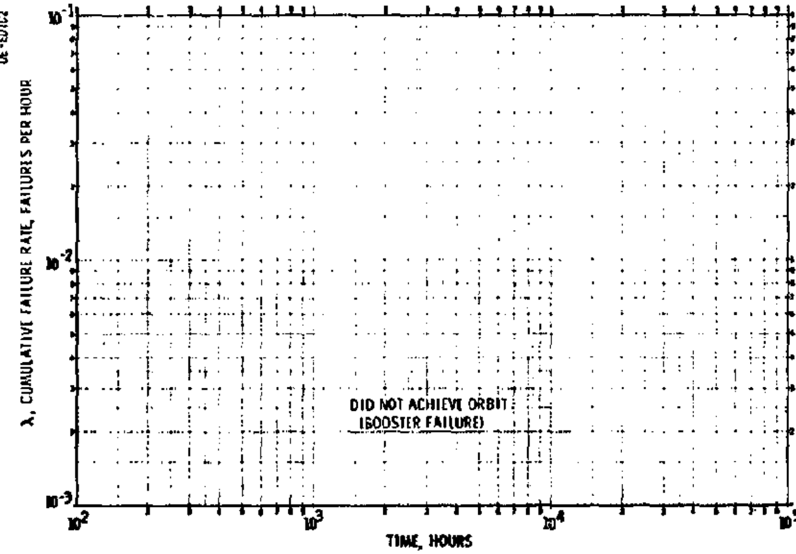
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-6. PROGRAM 1 F5 DATA

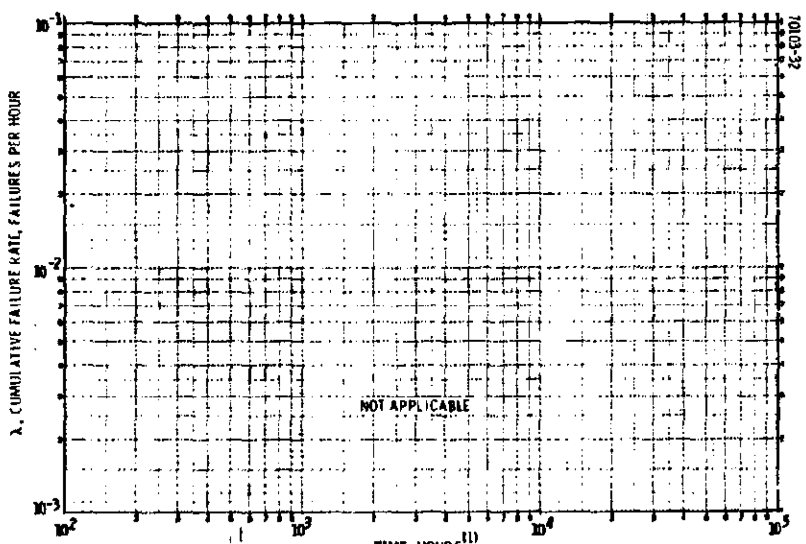
(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



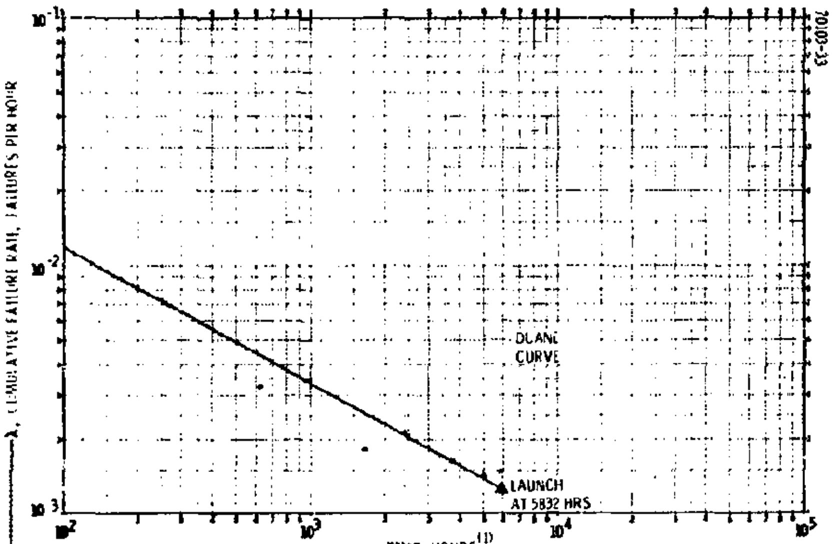
b) IN-ORBIT



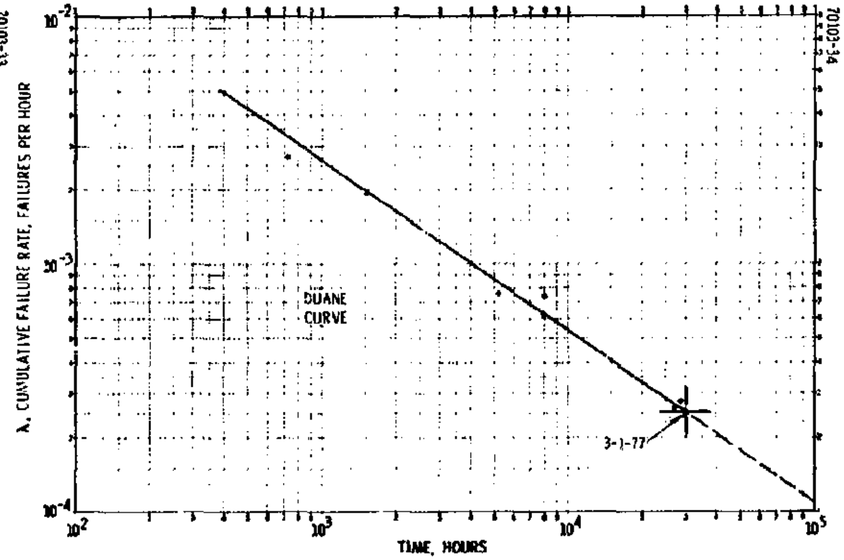
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-7. PROGRAM 1 F6 DATA

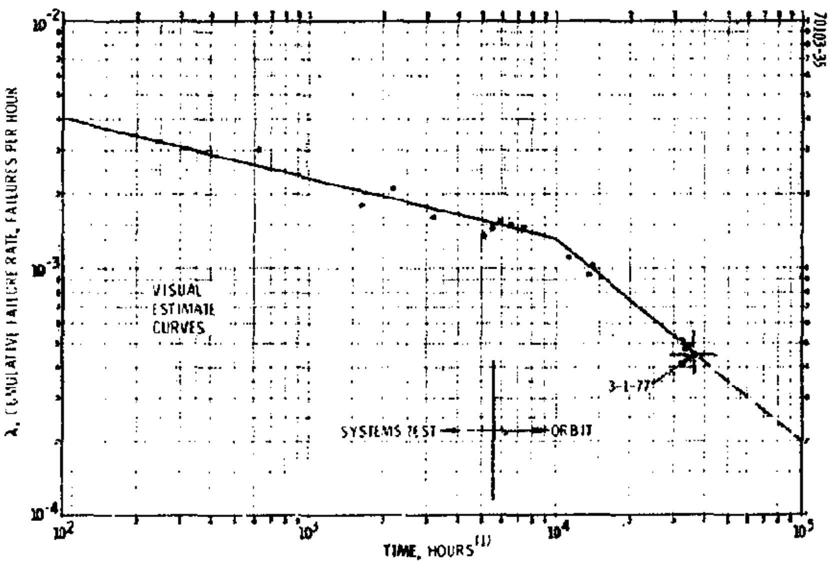
(1) REPRESENTS SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



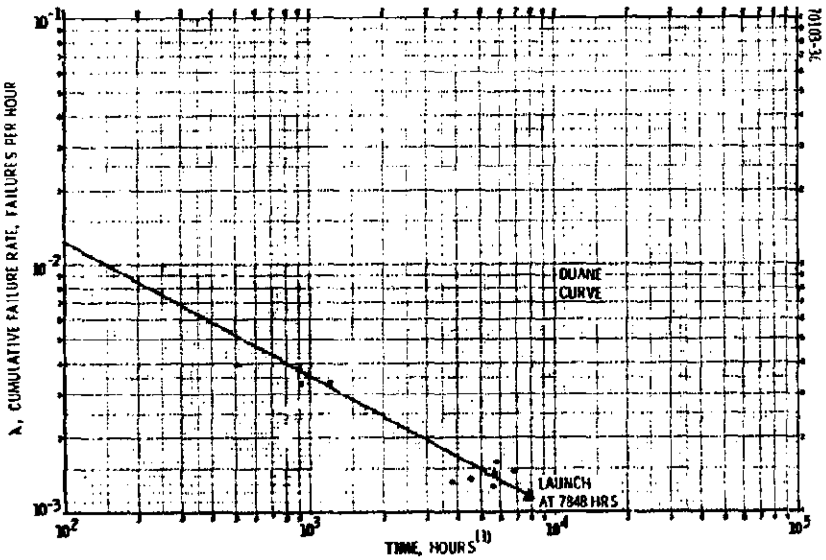
b) IN-ORBIT



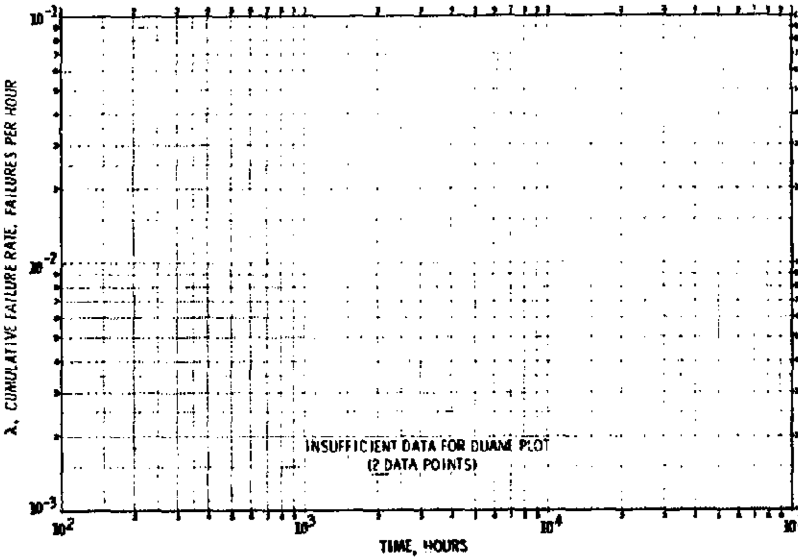
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-8. PROGRAM 1 F7 DATA

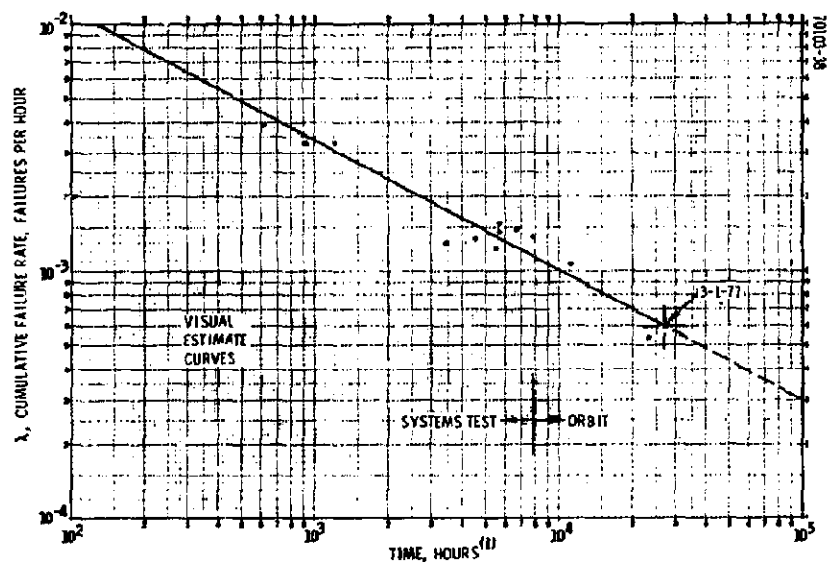
(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



b) IN-ORBIT



c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-9. PROGRAM 1 F8 DATA

(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME

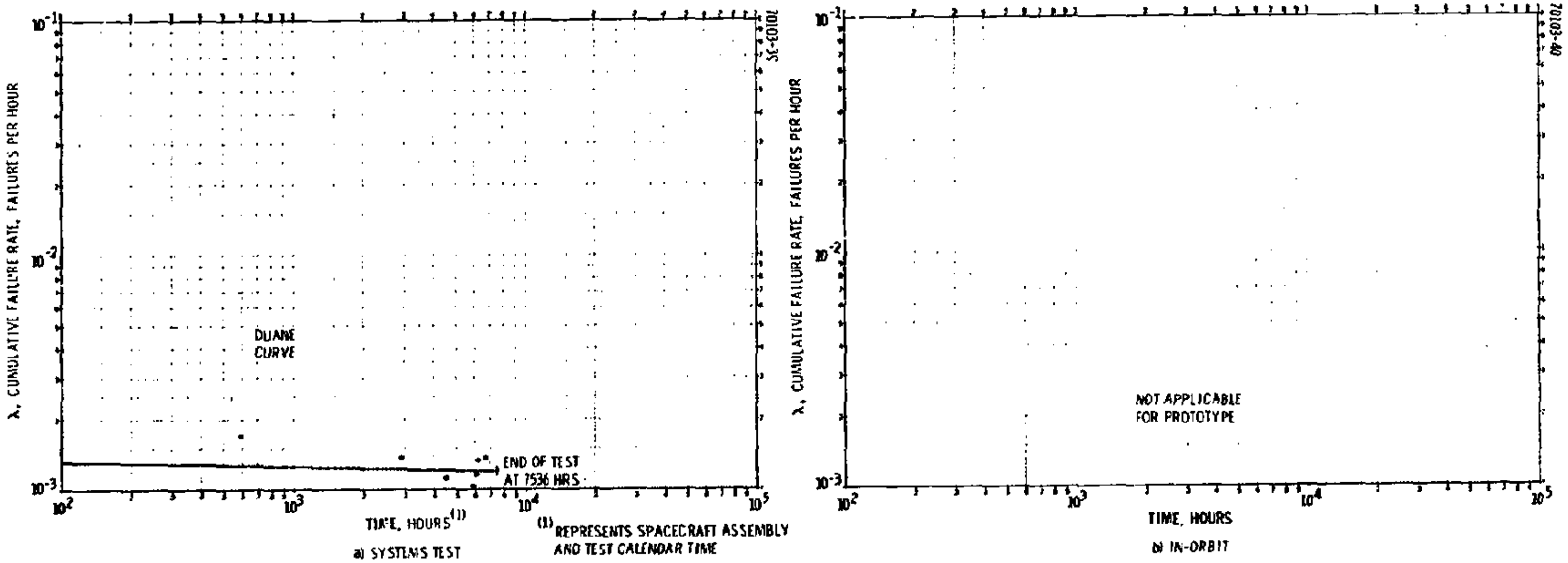
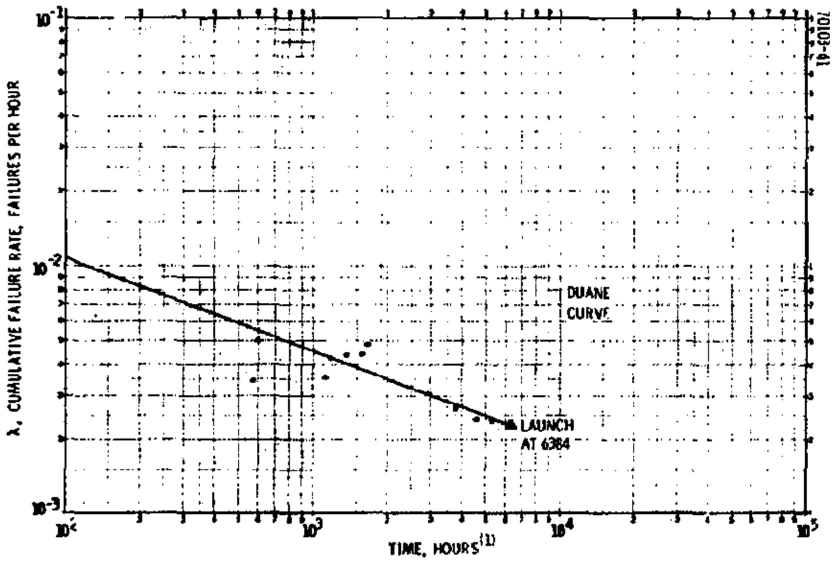
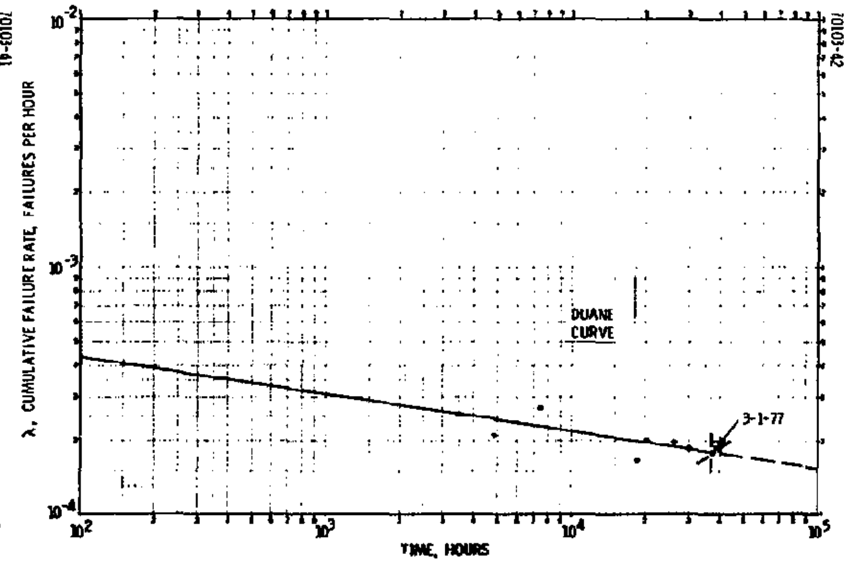


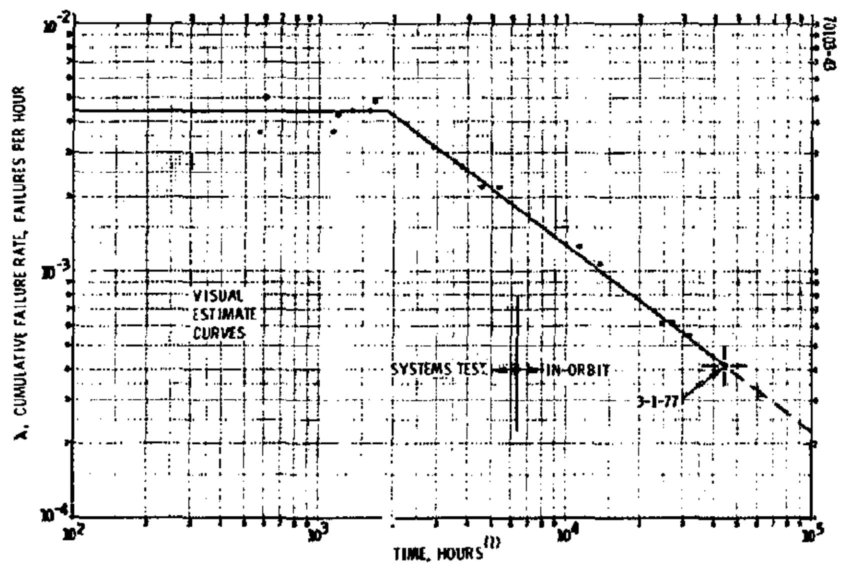
FIGURE 5-10. PROGRAM 2 PROTOTYPE DATA



a) SYSTEMS TEST



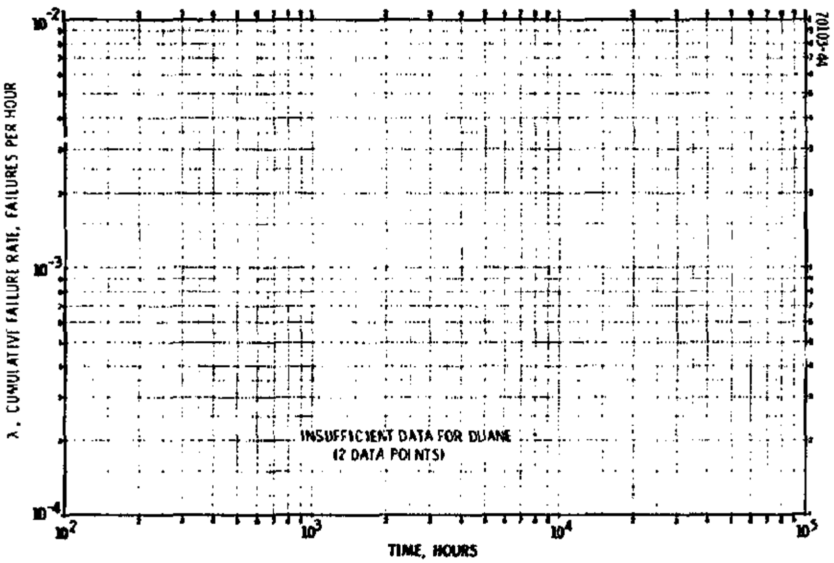
b) IN-ORBIT



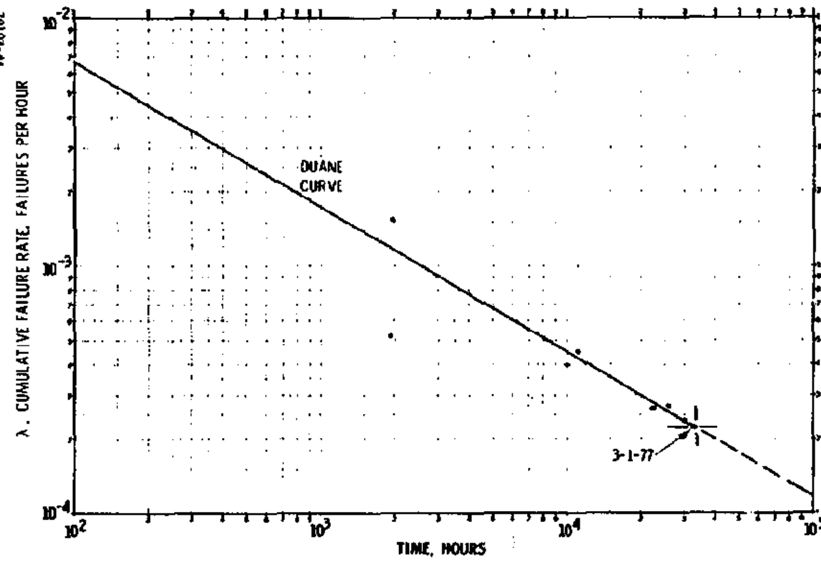
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-11. PROGRAM 2 F1 DATA

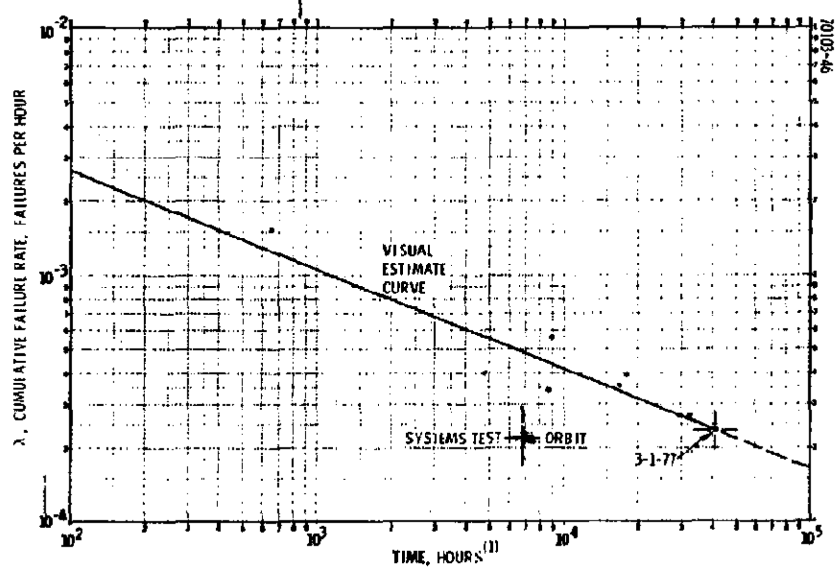
(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



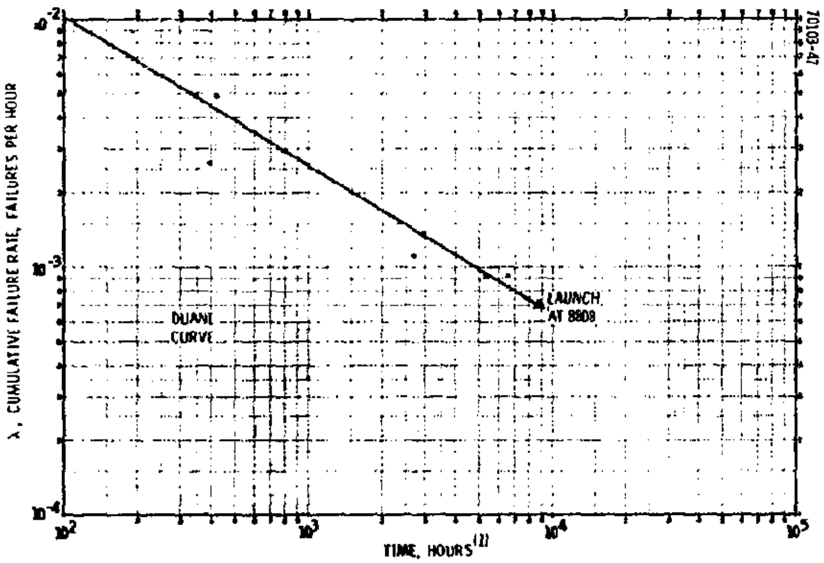
b) IN-ORBIT



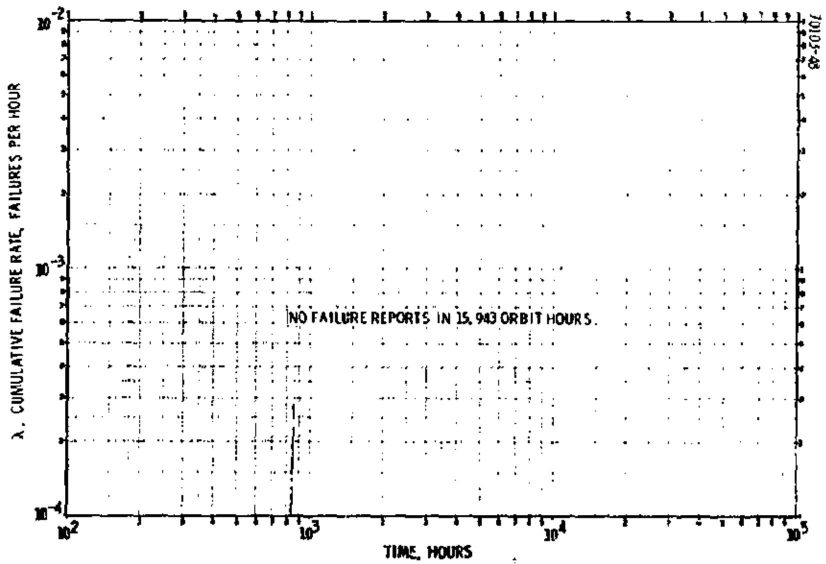
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-12. PROGRAM 2 F2 DATA

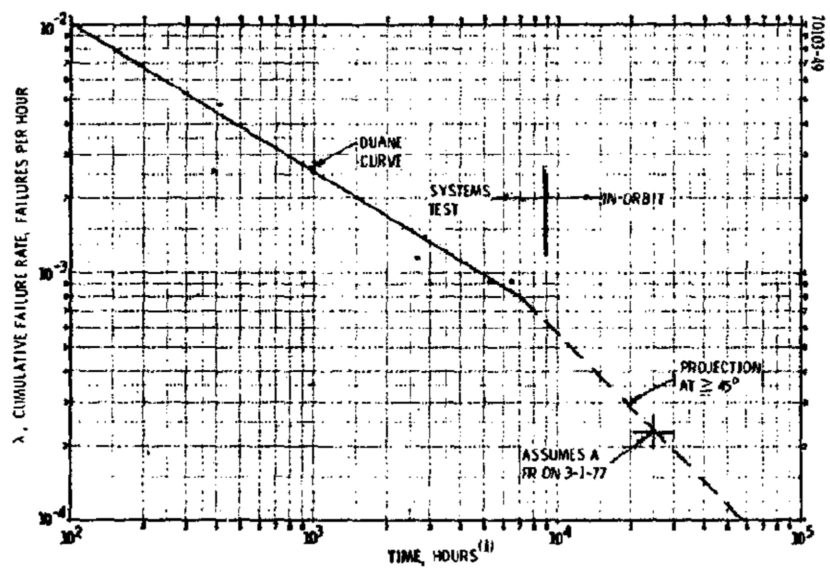
[1] INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



a) SYSTEMS TEST



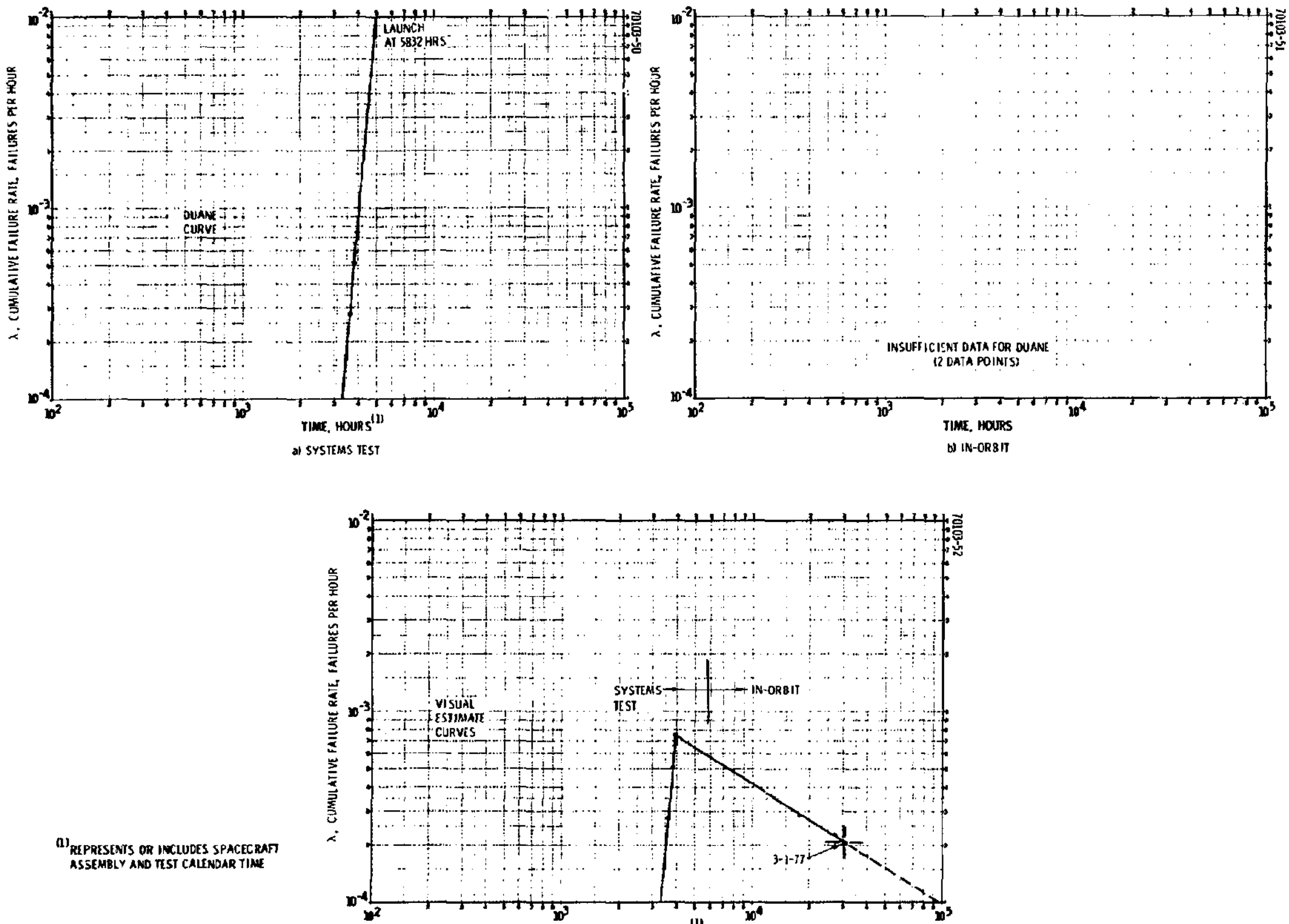
b) IN-ORBIT



c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-13. PROGRAM 2 F3 DATA

(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



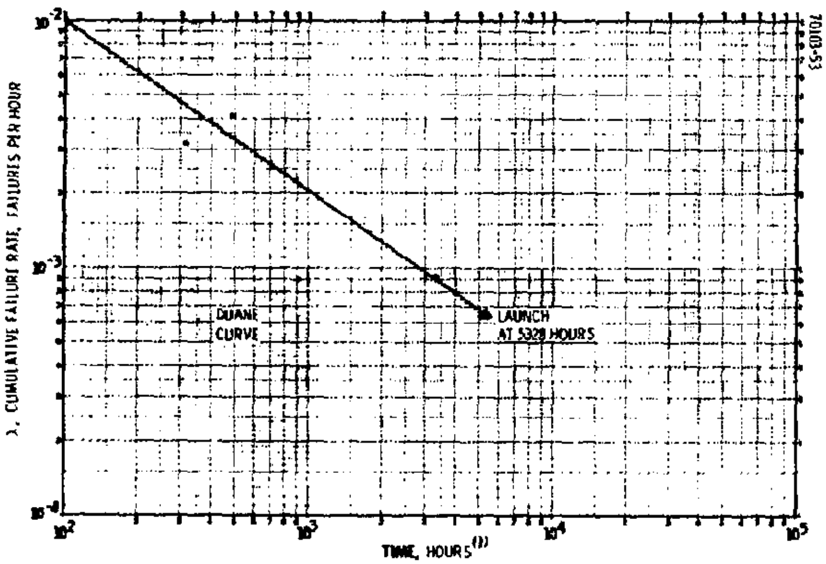
a) SYSTEMS TEST

b) IN-ORBIT

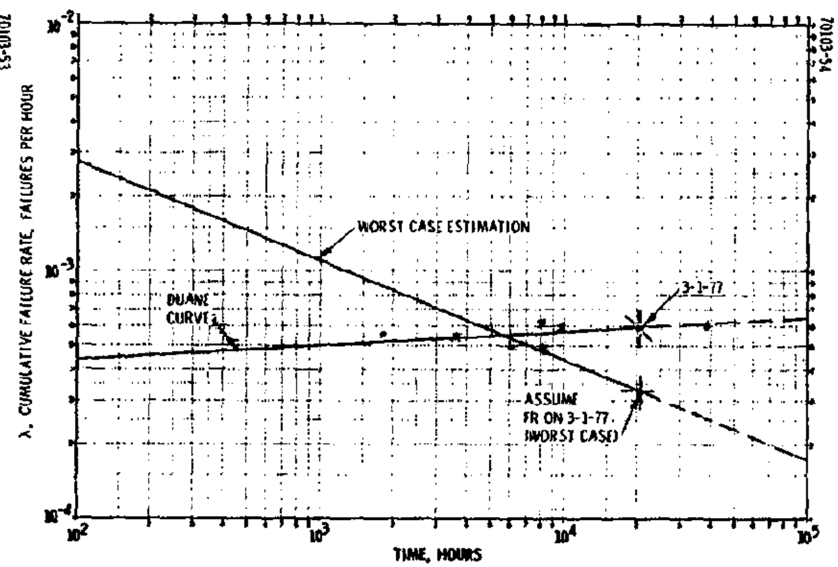
c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-14. PROGRAM 3 F4 DATA

(1) REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME

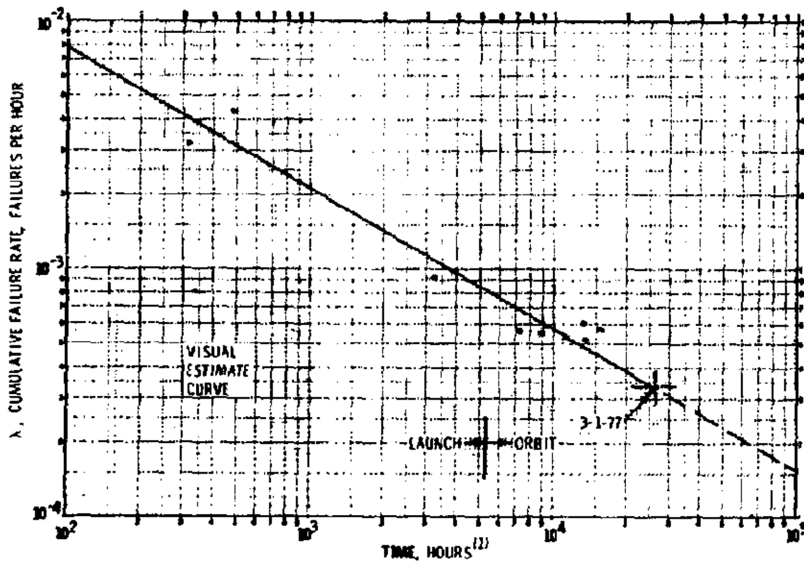


a) SYSTEMS TEST



b) IN-ORBIT

(1)REPRESENTS OR INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



c) SYSTEMS TEST AND IN-ORBIT COMBINED

FIGURE 5-15. PROGRAM 3 F5 DATA

The Duane curve provides a reasonable fit of the systems level test failure rates. When this is combined with the visually estimated in-orbit continuous curve after launch, a reasonable approximation of each spacecraft total profile is obtained. These profiles have been drawn for every spacecraft on Figures 5-16 and 5-17 for Program 1 and Programs 2 and 3, respectively. A convergence of the failure rates at the end of systems test (and a high correlation) with the in-orbit initial failure rate can be seen. The two spacecraft which have had no failures in orbit were estimated on the individual C figures by the method of assuming a failure on 1 March 1977, with the additional requirement that $\alpha \leq +1$. The use of λ_{fst} as a predictor of initial in-orbit performance and the fact that the in-orbit slopes within each program are similar (0.78 for Program 1 and 0.26 for Programs 2 and 3) form a basis for future prediction of in-orbit performance.

A composite of figures C was prepared as shown in Figures 5-20 and 5-21 for Program 1 and Programs 2 and 3, respectively. These differ from Figures 5-16 in that only the visually estimated curves are shown.

The varying slopes during the systems test phase are primarily caused by changes in the screening effectiveness of the various test phases and by the total number of defects presented to the systems level screen. This is an inherent weakness of the Duane model. However, as the spacecraft design matures, the unit level and initial systems level tests become more effective and the cumulative failure rates assume a more classic Duane profile.

The definition of a physical model which could provide a basis for the observed Duane trends is considered in Appendix B. A relatively simple model is shown to fit most of the empirical data evaluated in this study.

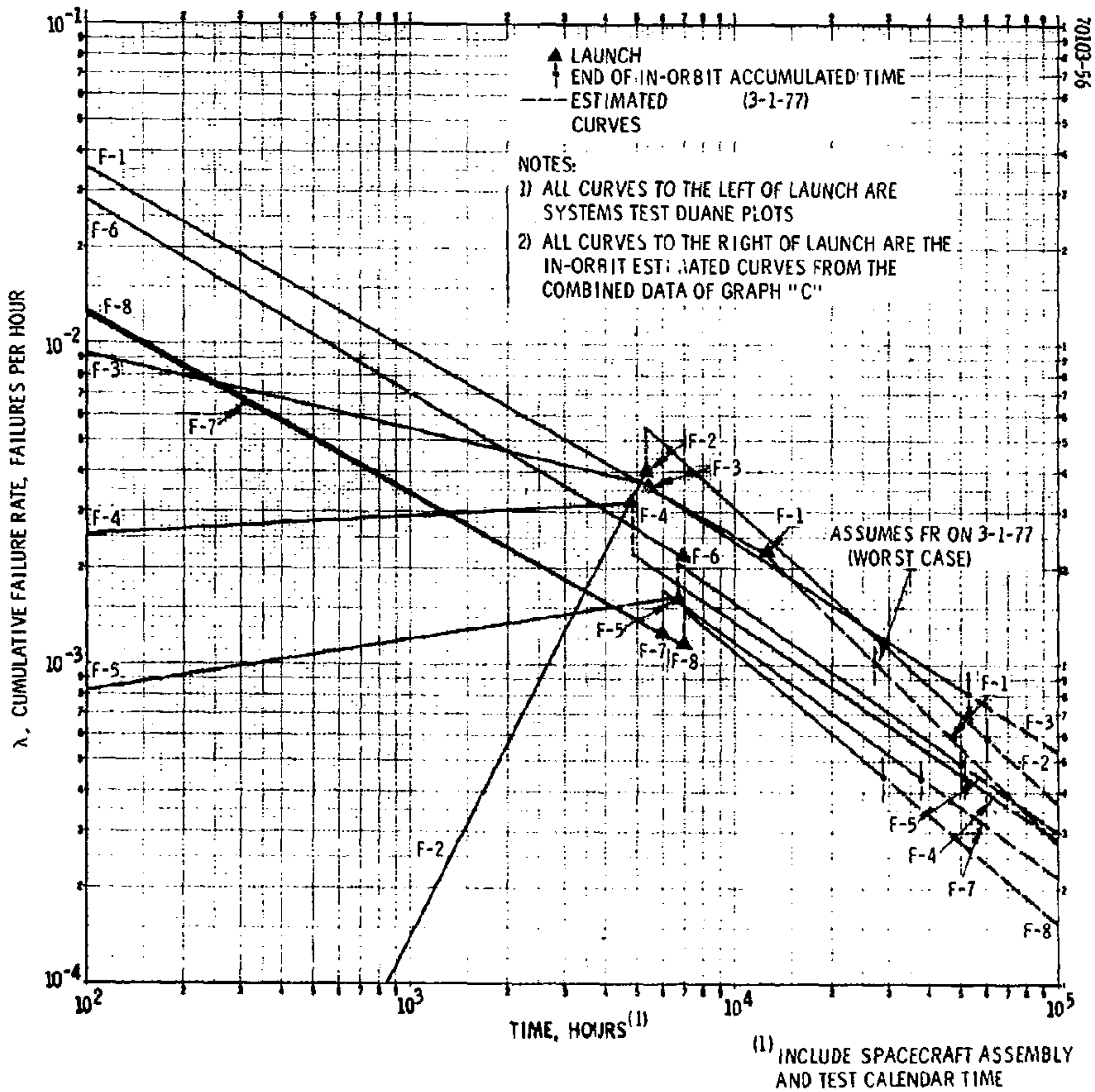


FIGURE 5-16. PROGRAM 1 – COMBINED SYSTEMS TEST DUANE PLOTS AND IN-ORBIT ESTIMATES

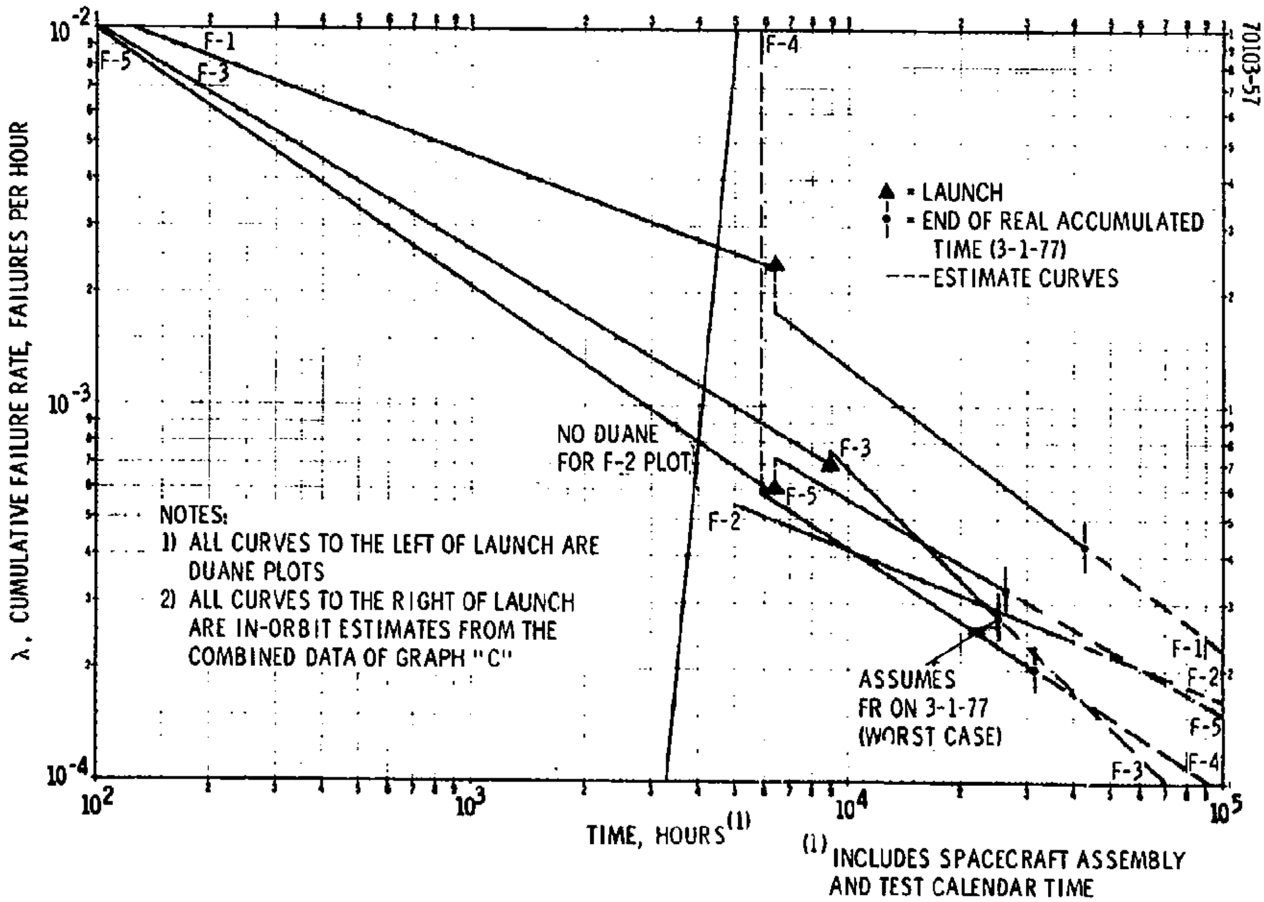
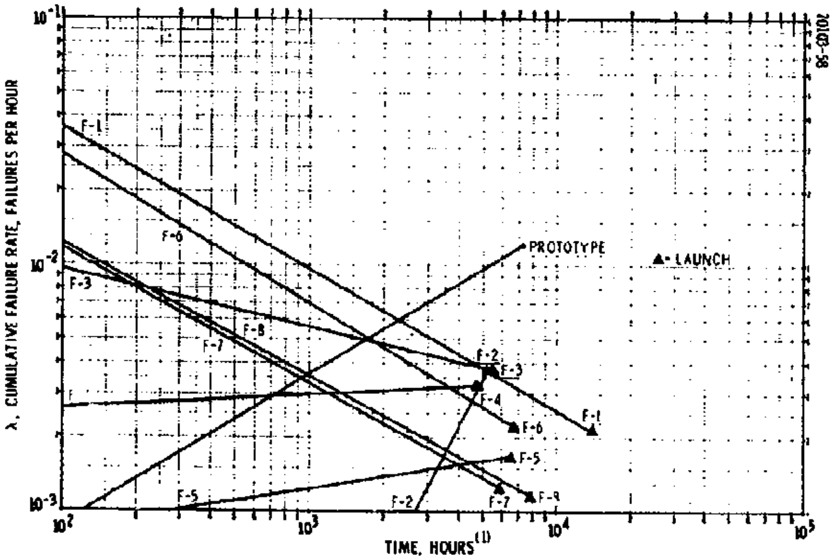
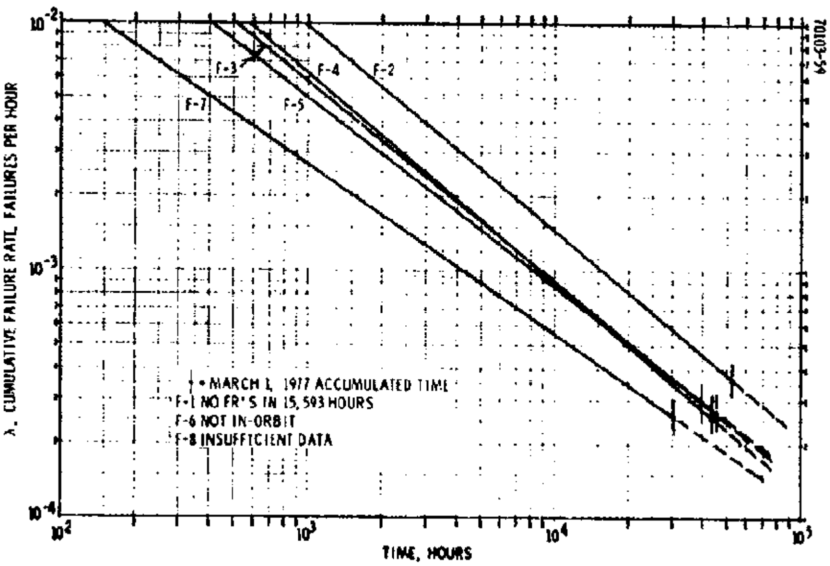


FIGURE 5-17. PROGRAMS 2 AND 3 – ALL SYSTEMS TEST DUANE PLOTS AND IN-ORBIT ESTIMATES



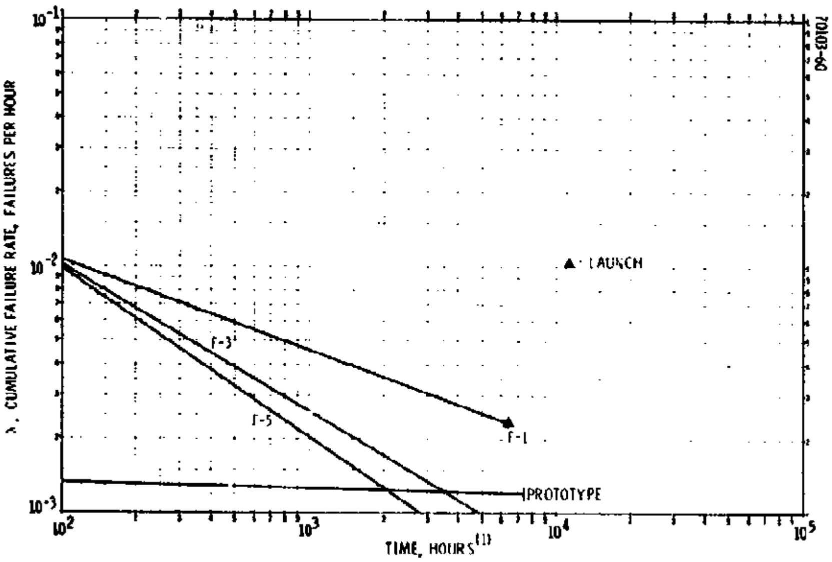
a) SYSTEMS TEST PROTOTYPE & F-1 THROUGH F-8

(1) INCLUDES SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



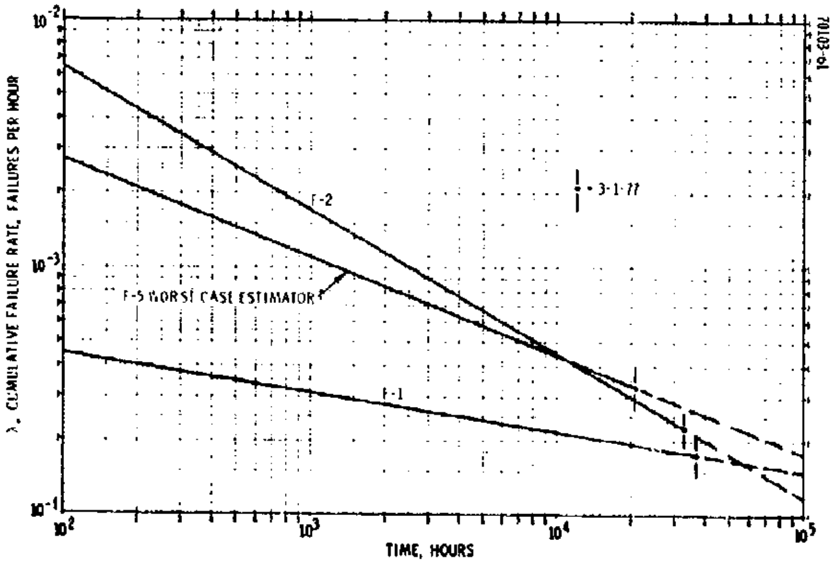
b) IN-ORBIT SPACECRAFT

FIGURE 5 18. PROGRAM 1 - ALL SPACECRAFT DATA



a) SYSTEMS TEST PROTOTYPE, F-1, F-3 AND F-5

(1) REPRESENTS SPACECRAFT ASSEMBLY AND TEST CALENDAR TIME



b) IN-ORBIT SPACECRAFT

FIGURE 5-19. PROGRAMS 2 AND 3 - ALL SPACECRAFT DATA

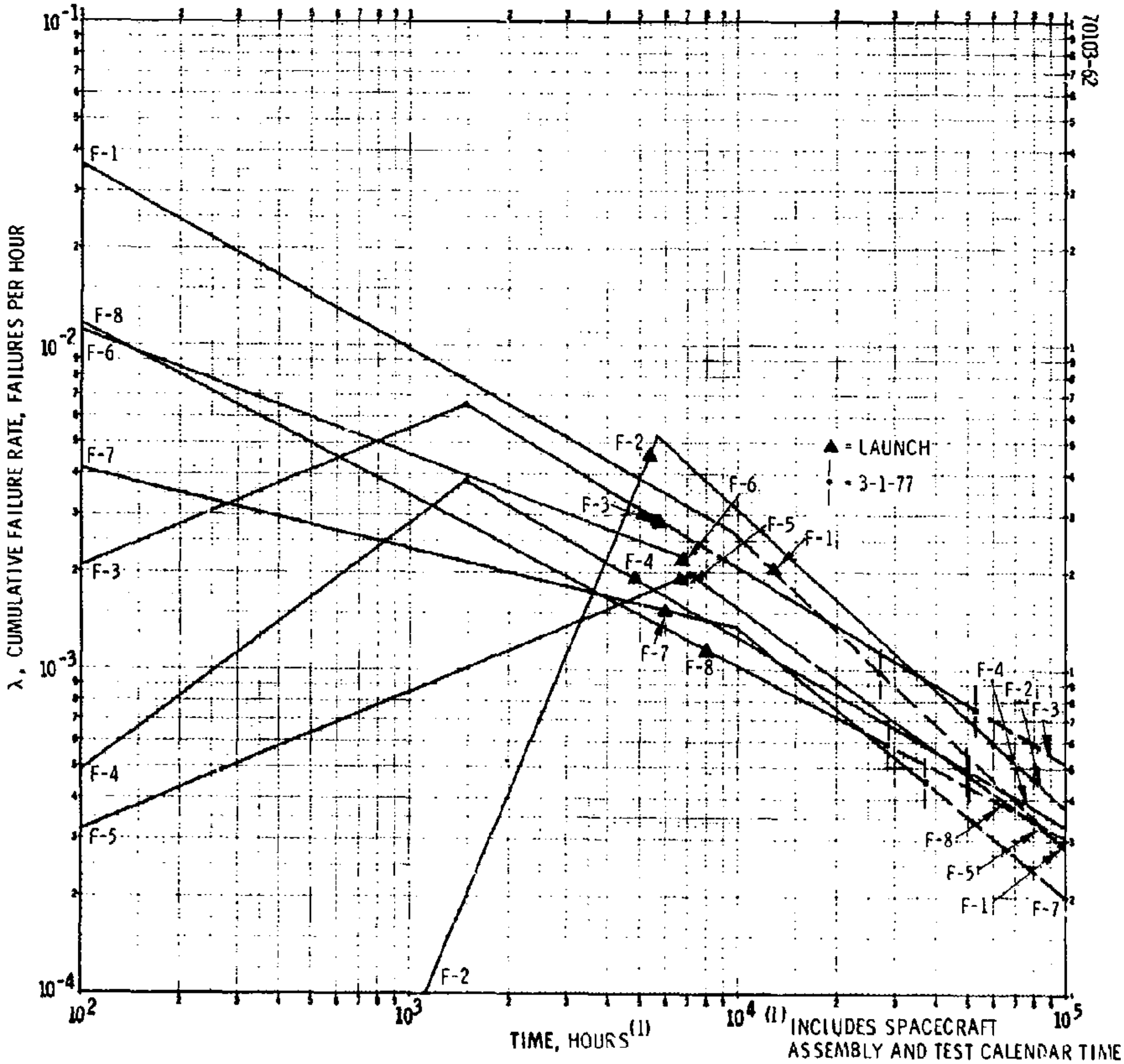


FIGURE 5-20. VISUAL ESTIMATE CURVES OF ALL PROGRAM 1 SPACECRAFT–SYSTEMS TEST AND IN-ORBIT COMBINED

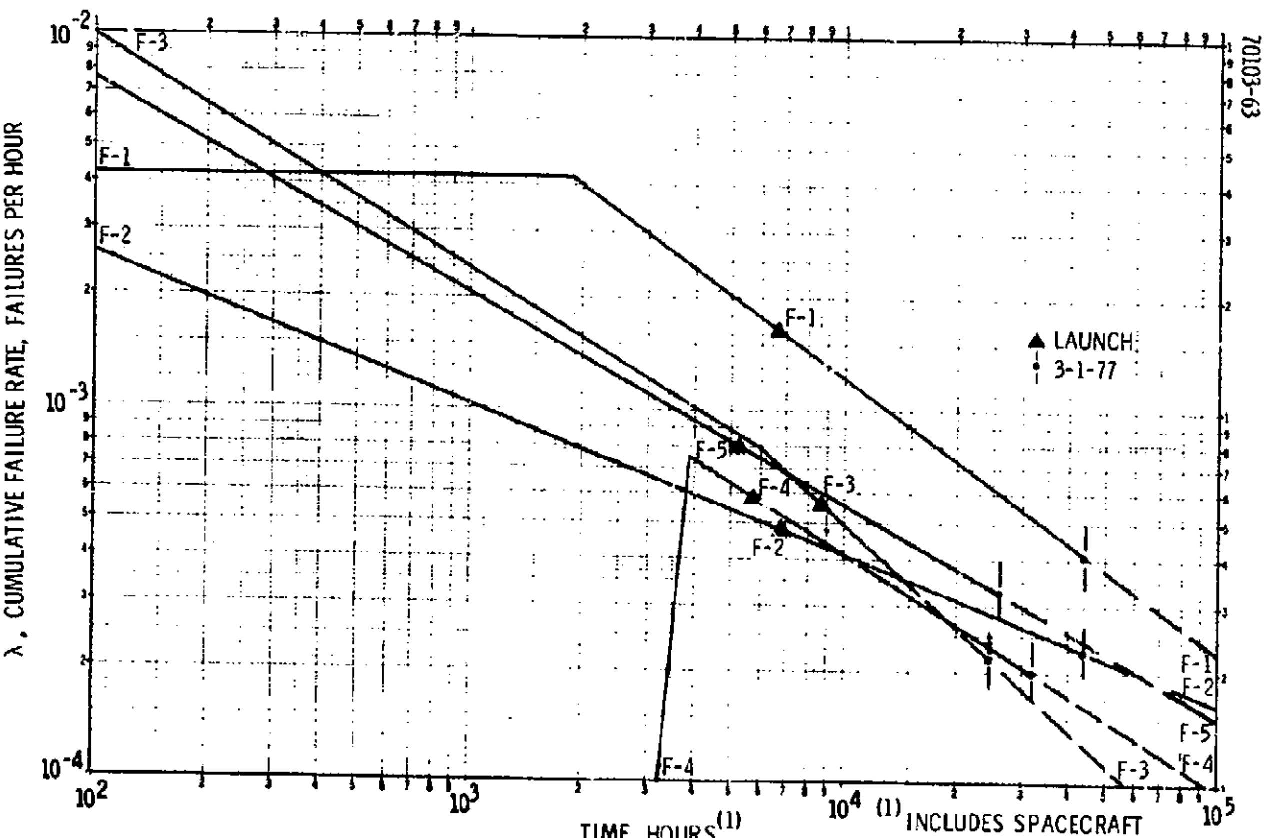


FIGURE 5-21. VISUAL ESTIMATE CURVES OF ALL PROGRAMS 2 AND 3 SPACECRAFT—SYSTEMS TEST AND IN-ORBIT COMBINED

6. SCREENING EFFECTIVENESS OF SYSTEMS THERMAL-VACUUM TESTS

The specific test effectiveness and general test effectiveness of thermal environments were discucsed in Sections 4.3 and 4.4. The test effectiveness of systems thermal vacuum acceptance tests was very low at 7 percent for Program 1 and 3 percent for Programs 2 and 3. A summary of systems level FRs is presented in Table 6-1. The test time-to-failure was determined for each of these systems level thermal vacuum FRs, and a failure histogram is presented in Figure 6-1. The screening effectiveness of the environment decreases rapidly with time with most of the failures occurring in the first 60 hours of test (87 percent).

The relative importance of this acceptance test screen is illustrated in Figure 6-2 for Program 1 systems thermal vacuum with DCTV excluded. The test effectiveness of the thermal vacuum environment (7 percent) is a factor of four less important than the following final ambient tests (27 percent) and a factor of ten less important than the first year in-orbit. The relatively low test effectiveness of the thermal vacuum environment may be explained in terms of the test not inducing sufficient stresses to cause failure, or there being enough novelty, e.g., changes in environment and operating modes to cause marginal hardware to fail. A cumulative failure rate curve was prepared, Figure 6-3, for the thermal vacuum FR data. From Figure 6-3 it is shown that additional time in thermal vacuum environment is of little value because the instantaneous failure rate (λ_i) is nearly zero ($\alpha = +1$) after 70 hours. This is equivalent to a test effectiveness asymptotically approaching 0.075. As shown in Figure 6-2, the cumulative test effectiveness for Program 1 reaches 0.06 after only 50 hours in test. The data strongly suggests that the test technique should be improved to add novelty and significantly increase the test effectiveness relative to the first year in-orbit. Time in the environment is not necessarily the most important factor.

TABLE 6-1. THERMAL VACUUM FR SUMMARY
(See Tables 3-11 and 3-12)

Thermal Vacuum Test	Program 1		Programs 2 and 3		Totals
	Qual	Accept	Qual	Accept	
DCTV/ESTV	N/A	5	1	6	12
Systems TV	7	9	0	2	18
Totals	7	14	1	8	30

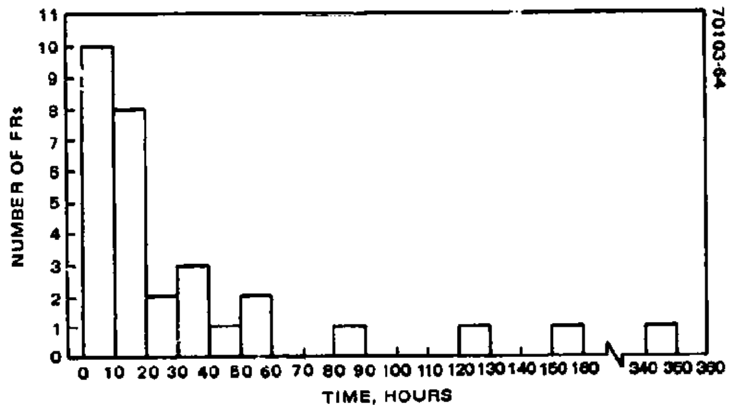


FIGURE 6-1. PROGRAMS 1 THROUGH 3 – DCTV/ESTV AND SYSTEMS LEVEL THERMAL VACUUM FAILURE REPORTS VERSUS TIME

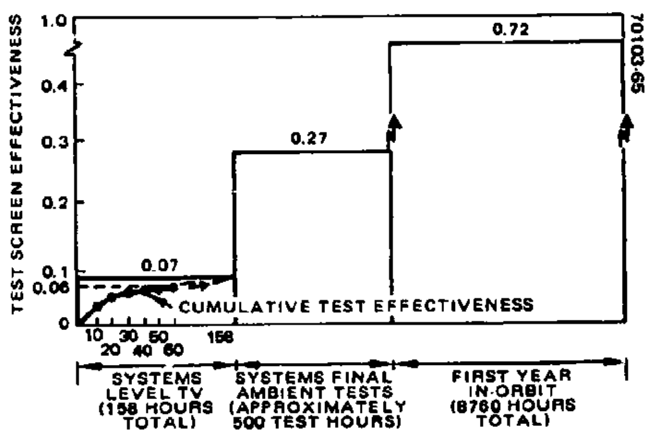


FIGURE 6-2. PROGRAM 1 TEST EFFECTIVENESS COMPARISONS

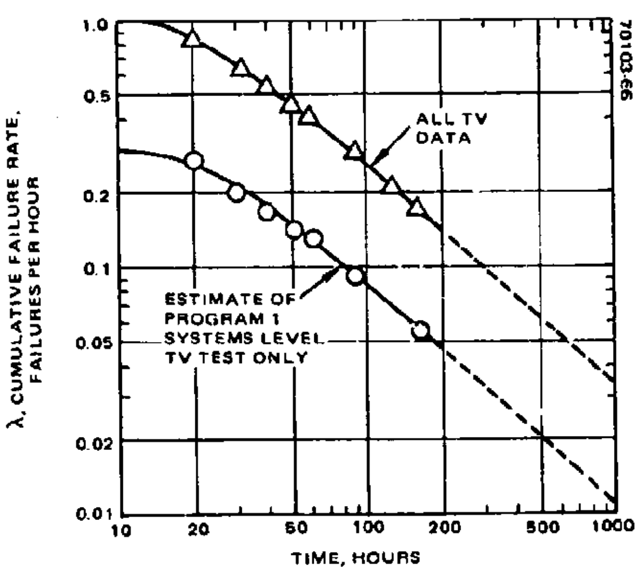


FIGURE 6-3. CUMULATIVE FAILURE RATE FOR THERMAL VACUUM TESTS AT SYSTEMS LEVEL (DCTV/ESTV AND SYSTEMS TEST THERMAL VACUUM)

7. COST EFFECTIVE TESTING

The issue of test cost effectiveness is a question of value received balanced against test costs. For a complex communications spacecraft, the total investment in an in-orbit spacecraft can easily exceed $40 million (including $20 million for a booster) with an operating revenue on the order of $40 million per year. Total recurring unit and systems test costs would typically be less than 7 percent of the spacecraft costs, or approximately $1.5 million. Therefore, from a return on investment viewpoint, it is apparent that all reasonable test screens should be utilized to ensure in-orbit reliability. It also should be noted that the relatively high operating revenues create a very significant launch schedule incentive, particularly for the initial spacecraft.

Typical recurring cost data for an acceptance test program are shown in Tables 7-1 and 7-2. The data are for programs of moderate complexity with 1 flight spacecraft and for 8 flight spacecraft, respectively. The data are strictly approximations and do not represent actual data for any specific program. However, it will be apparent that large variances are possible without significantly affecting the conclusions. With reference to Table 7-1, note that the fixed components of unit and systems level acceptance tests are 67 and 76 percent, respectively. Therefore, the basic test plan is the predominate test cost driver. Hardware problems and failures create only 20 to 30 percent of the total test cost. However, secondary program stretch-out costs resulting from systems level problems override all test cost considerations by a large factor. The schedule implications of the basic test plan are also obviously important because of the secondary program costs (typically on the order of $25K/day in spacecraft cost and as much as $110K/day of in-orbit operating revenue). Therefore, based on the cost drivers, it seems clear that all practical steps to achieve a near zero-failure condition at systems level are warranted. The optimum approach would be zero-failures at both unit and systems level through an effective emphasis on defect prevention at the subunit level. The test effectiveness of the unit level screens should also be substantially greater than either the systems test environment or the in-orbit environment. In addition, the duration of the acceptance test program should be varied in accordance with the actual failure experience on the particular set of spacecraft hardware as proposed in section 4.6 of this report.

For a program with eight flight spacecraft, the relative cost drivers shift significantly as shown in Table 7-2. The program stretch-out costs

TABLE 7-1. TYPICAL RECURRING ACCEPTANCE TEST COSTS
(1 Flight Spacecraft)

Item	Dollars in Thousands		
	Fixed	Variable	Total
Basic unit tests	200		
Unit rework and retest due to unit level problems		100	
Unit total	200 (67%)	100 (33%)	300 (100%)
Basic systems test	750		
Systems retest due to problems		200	
Unit work and retest due to systems problems		35	
Systems total	750 (76%)	235 (24%)	985 (100%)
Program stretch-out costs due to systems problems		1700	
Program total		1700	1700
Grand total	950	2035	2985
		$\Delta$$	Ratio
Potential savings			
If all problems screened at unit level		1900	
If zero-failures achieved in test program		2035	
Maximum permitted increase in unit costs to achieve zero-failure at systems level			9:1

are not affected, but all other cost elements increase proportionately with quantity of spacecraft. The overall effect is to increase the importance of test costs. Corrective actions at the subunit level, which result in lower failure rates, become more highly levered.

Based on an engineering review of the FR data base included in this study, it is probable that only a small fraction of the failures are random in the sense that corrective action is either impossible or not cost effective. In fact, there are many corrective action paths which do not involve increased manpower. For example, a recent Hughes experiment in an electronics production area indicated that the manufacturing defect rate could be dramatically reduced by simply assuring rapid feedback of problems to the person responsible. It is probable that an aggressive corrective action program to achieve near-zero failures in both the qualification and acceptance test programs would be cost effective.

TABLE 7-2. TYPICAL RECURRING ACCEPTANCE TEST COSTS
(8 Flight Spacecraft)

Item	Dollars in Thousands		
	Fixed	Variable	Total
Basic unit tests	1600		
Unit rework and retest due-to unit level problems		800	
Unit total	1600 (67%)	800 (33%)	2400 (100%)
Basic systems test	6000		
Systems retest due to problems		1600	
Unit rework and retest due to systems problems		280	
Systems total	6000 (76%)	1880 (24%)	7880 (100%)
Program stretch-out costs due to systems problems		1700	
Program total		1700	1700
Grand total	1600	4380	11,980
		Δ$	Ratio
Potential savings			
If all problems screened at unit level		3300	
If zero-failures achieved in test program		4380	
Maximum permitted increase in unit costs to achieve zero–failures at systems level			2:1

Relative to qualification tests, the results discussed in section 4.5 of this report indicate that none of the unit or systems environments are very effective as a screen for design problems. The data suggests that an increased emphasis on worst-case analysis, development tests, and other nonrecurring activities would be cost effective, including an aggressive corrective action program, intended to achieve near-zero failures in the unit and systems qualification test programs. With such an approach, the need for a qualification spacecraft becomes dubious. It is probable that destructive tests are not required to achieve an effective screen and there is no apparent schedule advantage to a sequenced development. Therefore, a protoflight spacecraft is probably the most cost effective path wherein the first flight spacecraft is exposed to a more comprehensive test program.

In summary, a cost effective test program would most likely have the following attributes:

1) An aggressive corrective action program would be pursued to achieve near-zero failures in both qualification and acceptance test programs. Such a program would include standards regarding acceptable defect rates in various categories (e.g., design, workmanship, etc.).

2) The effectiveness of the various test screens would be measured and an aggressive policy would be pursued to maximize the effectiveness of the overall test program. The key is to ensure a high level of novelty in each test period. Time alone is probably not an effective screen in any given stable environment.

3) In particular, the effectiveness of the unit test environments would be increased significantly to ensure a near-zero failure rate at systems level.

4) Observed failure rates would be used to predict both systems level test results and in-orbit performance. The content of the test program would be varied in accordance with these predictions.

8. CONCLUSIONS

The test programs considered by this study were extremely effective in ensuring a high level of in-orbit reliability. Only a single catastrophic problem occurred during almost 44 years of in-orbit operation on 12 flight spacecraft. This problem was the result of an early wearout phenomenon which could not have been screened by the qualification and acceptance programs as designed.

The results of this study indicate that in-orbit failure rates are highly correlated with failure rates observed in the unit and systems test programs. Further, the data suggest that a test effectiveness model is more generally applicable than a Duane extrapolation because of reduced sensitivity to variations in the test screen effectiveness. With suitably standardized failure criteria, it is probable that predictions based on observed FRs and a test effectiveness model can be used to guide the content of a test program to ensure that the stated in-orbit reliability goals are achieved.

Cost considerations suggest that an aggressive corrective action program to achieve a near-zero failure rate for all systems level testing would be cost effective for both qualification and acceptance. Very few failures are considered to be random in the sense that preventive actions are impractical. In general, test activities at the black box level and above should be considered insurance functions, not screens routinely used to improve product reliability.

8.1 GENERAL

1) Normalization of data bases from similar programs is feasible; and when based on electronic parts count, number of units, and number of different unit types, provides comparable program to program data.

2) Approximately 1 FR of any type was generated against each 127 electronic parts produced on each of the three programs.

3) Only 68 percent of the FRs were primary, 20 percent were integration and/or test (I&T) induced, and 11 percent were secondary. Of the I&T FRs, 41 percent were electrical test errors and 24 percent were electrical overstress. I&T FRs remained essentially constant as programs matured, and eventually almost

equalled the primary FRs generated on later spacecraft at the systems level. Increased emphasis on test procedures and safety practices would have been cost effective in reducing the I&T FRs. Of secondary FRs, 64 percent were for acceptable out-of-specification test results. Again, better engineered procedures would have reduced the FR rates.

4) Z and X were more effective than Y in vibration testing, and the eclipse simulation was the most significant phase of the systems level thermal vacuum test (48 percent of all thermal vacuum FRs).

5) The ambient initial checkout and performance tests are very effective with at least 42 percent of all unit and 48 percent of all systems level FRs being detected at these test phases.

6) As the programs matured, the FRs generated at both unit and systems level declined. The first spacecraft has three times as many FRs as the average of all the subsequent spacecraft. Additionally, the last spacecraft in any program set has an increased number of FRs compared with the preceding spacecraft because of delays which have resulted from unit problems.

7) Workmanship problems dominate the primary FRs (40 percent of total and 51 percent of acceptance). Sixty percent of all workmanship FRs are installation and/or assembly errors. Little evidence of learning in this area was evident. An emphasis placed on reduction of installation/assembly problems should result in substantially reduced FR rates.

8) Part problems remained essentially constant throughout the test programs at 22.5 percent of all FRs. Forty four percent of the qualification and thirty one percent of acceptance part problems were caused by manufacturing deficiencies.

9) Subsystems with high concentrations of electronics (T&C and ACS) had a disproportionately high number of FRs. Power and propulsion subsystems had a disproportionately high nunber of I&T FRs.

10) Half of all in-orbit problems were duplications of previous occurrences. Fifty-two percent of the unique problems were design, the majority of which could not have been detected by the acceptance or qualification test programs as designed.

8.2 QUALIFICATION TESTING

1) The unit and systems level qualification tests were found to be marginally effective in detecting design deficiencies. Approximately 8.4 percent were screened at unit level and 22.3 percent at systems level, for an overall effectiveness of 28.8 percent.

These results suggest not only that the test techniques should be improved, but also that greater emphasis should be placed on subunit development tests, worst-case analysis, design review techniques, etc.

2) Systems level vibration testing is an effective qualification test.

3) In general, systems level tests are more effective for qualification purposes than for acceptance purposes.

8.3 ACCEPTANCE TESTING

1) Most vibration related problems are detected during qualification tests or unit level acceptance tests. Three-axis, systems level vibration testing is probably not cost effective.

2) All systems level acceptance thermal tests were only 8 percent effective. This is probably the result of deficiencies in test technique and not strongly related to total test time in the environment. Virtually all of the failures were detected during the first 60 hours of thermal vacuum testing. The effectiveness of the system thermal environments may be severely constrained by the inherent limits imposed on temperature ranges and rates of change.

3) Overall, the acceptance test programs were highly effective in screening critical defects (99 percent). For all problems, the combined unit and systems test effectiveness was 87 percent for Program 1 and 73 percent for Programs 2 and 3.

8.4 RELIABILITY GROWTH

1) Test effectiveness models can be used at both unit and systems level to estimate the population of defects and to predict performance in systems test and in-orbit. In-orbit failure rates are strongly correlated with test program results. These predictions could be used to vary the content of the test program in accordance with reliability requirements.

2) The in-orbit cumulative failure rate data exhibits a consistent slope which implies a constant equivalent test effectiveness.

3) The in-orbit failure rates generally decrease with successive members of a spacecraft series, indicating significant reliability growth. The cumulative failure rate at the end of systems test is a good predictor of in-orbit performance.

4) The systems test and in-orbit cumulative failure rate data cannot be accurately combined into a single Duane plot because the slopes are generally different. The disparity between the slopes generally decreases with successive members of a spacecraft series. For a mature design (e.g., quantity greater than 10), it is probable that a single Duane curve could be utilized for extrapolation purposes.

8.5 COST EFFECTIVENESS

1) The consequences of an in-orbit critical failure are generally too great to permit significantly reduced testing. However, the test program should be considered a matter of insurance and not a technique to achieve reliability growth.

2) The most cost effective test program will probably be achieved through aggressive corrective action to achieve a near-zero failure rate at both unit and systems level. The cost inpact of schedule delays at the systems level is particularly significant,

3) A continuing effort is warranted to measure and improve the screening effectiveness of each environment in the test program.

4) Approximately 70 percent of all direct test costs are fixed by the definition of the test program.

APPENDIX A. SUMMARY OF INFORMATION CODES

(A DESCRIPTION OF INFORMATION
OR CODES ENTERED IN
EACH COLUMN)

Column	Code	Description
1 through 5		FR or IFR number entered here. Columns 1 and 2 contain a non-U.S. subcontractor code of IFRs only.
7		Level — Indicates reporting level where failure occurred. A code is used to indicate these levels:
	N	Unit (foreign or domestic)
	F	Spacecraft (systems test)
	P	Post-launch (orbit)
8		Subsystem — Spacecraft subsystem from which part failed, coded:
	0	Other (thermal, etc.)
	1	Communications
	2	Antenna
	3	Power
	4	Despin control subsystem
	5	T&C digital
	6	T&C RF
	7	RCS
	8	Structure/mechanical
	9	Harness assembly
9		Activity Type — During which failure occurred, coded:
	1	Acceptance tests (any)
	2	Qualification tests (any)
	6	Alignment
	7	Launch operations
	8	Launch/boost
	9	Orbit

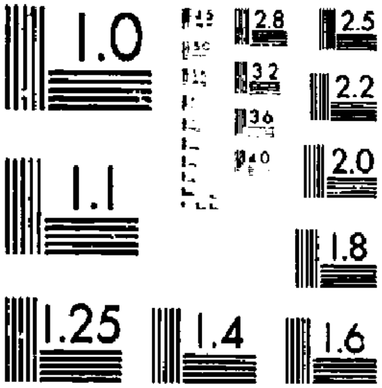
Column	Code	Description
10		Test Phase — In process when failure occurred, coded:
	0	Inspection
	1	Electrical manufacturing
	2	Mechanical manufacturing
	3	Electrical performance test (first, not-defined, or last)
	4	IST (first or last)
	5	ESTV/DCTV (electrical shelf or despun compartment thermal vacuum)
	6	Vibration (sine or random)
	7	TV/STV (spacecraft (or solar) thermal vacuum)
	8	Thermal cycle
	9	RCS tests
12 through 14		Unit Part Number (P/N) — Of failed item.
16 through 18		Part P/N — Of failed component, if applicable.
19		Importance Code:
	-	Primary type of failure/anomaly
	T	Test or integration induced
	S	Secondary problem — no action taken except NCMR or TFR close out
	A	Anomaly not verified w/retest
	D	Duplicate of in-orbit problem
20		Type code for test (& integration) induced:
	1	Damaged wires, heaters, and other leads
	2	Electrical test overstress error
	3	Electrical test error
	4	Hardware handling problems
	5	Wrong environment imposed
	6	Bad part selection in test
	7	RCS test error
	8	Test fixture induced
	9	Spacecraft operator error; in-orbit
20		Type code for secondary problem:
	0	Out-of-spec/OK
	1	Measurement technique bad/no retest
	2	Test equipment error/no retest

Column	Code	Description
20 (continued)	3	Manufacturing problem or assembly error multiple duplications
	4	Test error/no rework
	5	Miscellaneous/no rework
	6	Anomalies/no retest
21 and 22		<u>Month</u> during which failure occurred.
23 and 24		<u>Day</u> of month failure occurred.
25		Last digit of <u>Year</u> in which failure occurred.
27 through 51		<u>Verbal</u> complement to failure code. Attempt to achieve a complete description of failure cause between this verbiage and failure code of Column 62 and 63.
53		<u>Symptoms</u> – Of failure, coded:
	1	Anomaly or specification variance
	2	Random occurrence
	3	Periodic/intermittent
	4	Hard failure
	5	Normal life or wearout
	6	Not hardware
	7	Visual observation
54		<u>Severity</u> – Of failure, coded:
	0	Not critical
	1	Rework to existing paper
	2	Retest to existing paper
	3	Degradation of performance or mission
	4	Catastrophic failure
	5	Redesign hardware or paper
	6	Replace failed part
	7	Not hardware
	8	Not verified
	9	Operationally overcome
56		<u>Generic Type</u> – Of failure, coded:
	0	Unknown/unverified
	1	Workmanship
	2	Part failure
	3	Design or performance
	4	No fail

Column	Code	Description
57		Type code for Workmanship failure:
	0	Installation/assembly
	1	Test error
	2	Paper error
	3	Bad electrical connection
	4	Accidental handling
	5	Misuse
	6	Wrong part used
	7	Damaged wires/coax
	8	Contamination/corrosion
	9	Bad alignment or machining
57		Type code for Part failure:
	0	Unknown
	1	Workmanship
	2	Contamination
	3	Manufacturing process
	4	Design
57		Type code for Design failure:
	0	No change
	1	Wrong part
	2	EMI
	3	Physical properties
	4	Specification error
	5	Process error
	6	Circuit design
	7	Mechanical installation
	8	Analytical error
59		Environmental Condition — Attributed to or existing at time of failure.
	0	Other/orbit
	1	Ambient
	2	Random
	3	Sine
	4	Hot
	5	Cold
	6	Eclipse (simulation)
	7	Thermal/transition/pumpdown
	8	EMI/EMC and static change
	9	Radiation

2 OF 2

7 25200 UNC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

Column	Code	Description
60		Test Modifier – For activity, coded:
	I	Initial
	F	Final
	X, Y, Z	Axis (X, Y, Z)
	B	Before
	D	During
	A	After
62 and 63		Numerical portion of Failure Code (cause of failure).
	01	Unknown (failure verified, but cause unknown)
	02	In process
	03	Contamination (e.g., weld splatter, etc., in electronic part or RCS)
	04	Failure not verified (failure of part not verified by test/analysis)
	05	Incorrect specification
	06	Suboxide arc-over
	07	Handling
	08	Test error (operator error)
	09	Manufacturing error (a mistake)
	10	Design inadequate
	11	Electrical overstress
	12	Micro cracks
	13	Cracked chip
	14	Thermal overstress
	15	Component (internal) short
	16	Drawings or procedure unclear or incorrect
	17	Wrong part used
	18	Open, missing, or defective weld or solder joint
	19	Normal and wearout
	20	Intermetallic formation
	21	Excessive and unnecessary test conditions
	22	High window impedance
	23	Mechanical tolerance
	24	Component (internal) open, missing, or defective connection
	25	Test equipment malfunction/design
	26	Manufacturing process
	27	Minor RF anomaly
	28	Paperwork lags spacecraft hardware
	29	Minor out of specification condition-analysis unprofitable
	30	Hermetic seal not sufficient
	31	Performance specification too tight
	32	No previous failure history – nonrecurring, isolated occurrence

Column	Code	Description
62 and 63 (continued)	33	Miswiring
	34	Design subject to environmental parameter drift
	35	Test equipment or structure unsafe or subject to human error
	36	Part lost — no analysis
	37	Inspection error (a mistake)
	38	Checking/inspection insufficient or non-existent
	39	Design/procedure subject to human error
	40	Premature wearout
64		Letter Portion of Failure Code (corrective action)
	P	Primary part failure (part caused), Part = Hi Rel in FPS01000. Basic corrective action involved a part supplier.
	S	Secondary failure, induced by external error such as poor workmanship or design.
66 through 68		Unit Serial Number — Of failed item.
70		S/N of Spacecraft upon which unit "flew"
72 through 74		Time Duration represented as XX.X (where available)
75		Time Base (as applicable):
	0	Seconds
	1	Hours
	2	Months
	3	Years
	4	Days
77 through 78		Test Cycle, Retest Number (R-), or multiple of similar failure.
80		Program Data Set
	0	Program 1
	1	Program 2
	2	Program 3

APPENDIX B. ANALYSIS OF SOME SIMPLE FAILURE RATE MODELS

SUMMARY AND CONCLUSIONS

The fact that empirical failure rate data appears to fit so well to a Duane curve (i.e., the plot of failures versus test hours appears to fit a simple straight line on log-log paper) suggests that the test/failure process ought to be explainable in terms of a simple analytical model based on at least an intuitive understanding of the process. This appendix investigates the applicability of several such models.

The first model considered here assumes an initial population of equally vulnerable (equal probability of detection) defects, where the instantaneous failure rate is proportional to the number of remaining defects. This results in a failure rate which is an exponential function of time, being nearly constant initially and approaching zero after several test "time-constants". The log-log plot of the cumulative failure rate versus test hours for this model clearly does not fit a Duane straight line and, therefore, this simplistic model is rejected.

The second model considered involves a simple sophistication of the first model wherein the defect population is assumed to be distributed in terms of probability of detection. That is, in the initial population of defects, some are easy to detect (these are culled out of the population early in the test cycle), and some are very difficult to detect (these become the iminishing source of failures later on in the test cycle). This model, even in the elementary form considered here, produces a failure rate (failures versus hours) curve which is vulnerable to a straight line fit on log-log paper. The slope of the cumulative failure rate curve derived from the model is in the same range observed in actual ground test and in-orbit failure rate data.

Further sophistication of the model (e.g., in terms of defect detection probability distributions, varying test effectivity, etc.) can provide further "shaping" of the predicted failure rate curve to more closely match the trends seen in the empirical data.

It is concluded from this preliminary analysis that the Duane property of empirical failure rate data can be explained in terms of a simple analytical model based on an intuitive understanding of the test/failure process. The key element of this model is the distribution of the defect population in terms of probability of (or vulnerability to) detection. Further development of the model can provide more complete correspondence with trends seen in the empirical failure rate data.

EQUAL PROBABILITY OF DETECTION FAILURE RATE MODEL

The first model considered is similar to the one investigated by Timmins*. Here we assume the existence of a fixed initial population of defects which is steadily reduced by successive test screens. The defects all have equal probability of detection, and the test screens all have equal effectiveness. We further assume that the failures which occur in a test screen (interval) are a constant fraction of the remaining defects. (These constraints on the model can be removed later, but are imposed to keep this first illustrative example as simple as possible). Let N_o = the number of defects in the initial population

N_k = the number of remaining defects after k test screens (or intervals)

F_k = total failures (defects detected) after k test intervals.

$= N_o - N_k$

T = time interval of each test screen

a = the fraction of the remaining defects which are detected in each test interval.

λ_k = cumulative failure rate

$$= \frac{F_k}{kT}$$

In accordance with the above assumptions, the number of defects remaining after (k + 1) test intervals is given by,

$$N_{k+1} = N_k - aN_k = (1 - a) N_k$$

*Timmins, A. F., "A Study of Relationship Between Performance in Systems Tests and Space", Proceedings of the Institute of Environmental Sciences, 1975, 1. 172

The solution of this difference equation is a simple geometric progression decrease of the defect population,

$$N_k = N_o \, (1 - a)^k$$

which produces the following relations for F_k and λ_k.

$$F_k = N_o \left[1 - (1 - a)^k\right]$$

$$\lambda_k = \frac{aN_o}{T} \left[\frac{1 - (1 - a)^k}{ak}\right] \tag{1}$$

Figure B-1 shows a plot of Equation 1 on log-log paper for two different values of the parameter a (which corresponds physically to test effectivity). From this plot it is clear that the model does not fit a Duane straight line. As indicated in the figure, the implied failure rate is nearly constant initially, and then approaches zero after many test intervals. The deficiency stems from the assumption of equal probability of defect detection, which results in a relatively rapid initial depletion of the defect population, followed by a diminishing trickle of failures from the depleted population.

A continuous-time version of the same model can be developed as follows:

N_o = the initial number of defects

$N(t)$ = the number of remaining defects after t hours

$F(t) = N_o - N(t)$ = total failures over the interval $(0, t)$.

Then, in accordance with the assumption that the failure rate is proportional to the number of remaining defects,

$$\dot{N}(t) = -\frac{N(t)}{\tau}$$

where, τ = test effectivity time constant (a parameter characterizing test effectivity)

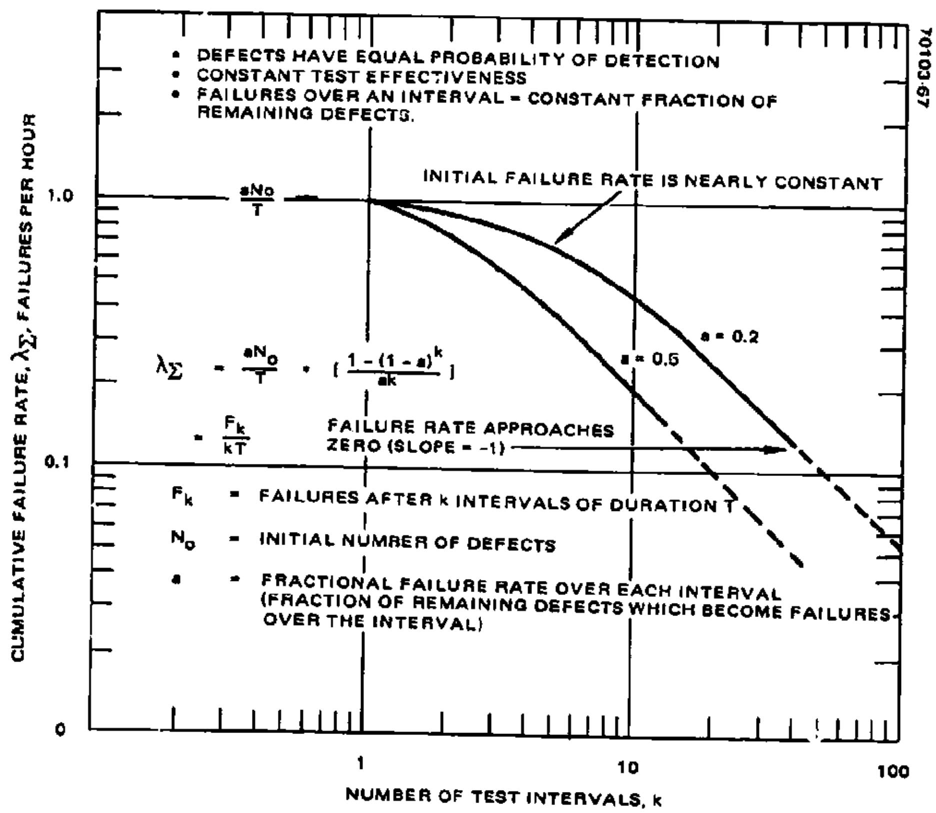


FIGURE B-1. EQUAL PROBABILITY OF DETECTION FAILURE MODEL – DISCRETE VERSION

Then,

$$N(t) = N_o e^{-t/\tau}$$

$$F(t) = N_o \left(1 - e^{-t/\tau}\right)$$

and the cumulative failure rate is expressed as,

$$\lambda_\Sigma(t) = \frac{F(t)}{t} = \frac{No}{\tau} \left[\frac{1 - e^{-t/\tau}}{(t/\tau)}\right] \qquad (2)$$

Figure B-2 shows a plot of Equation 2 on log-log paper, where the normalized quantity (t/τ) is used for the time scale. Here we note the same deficiencies (relative to a Duane fit) as seen in the discrete version of the model displayed in Figure B-1. That is, the initial failure rate is nearly constant as the defect population is rapidly depleted, followed by a failure rate which rapidly approaches zero (slope on log-log paper = -1). This follows from Equation 1 since,

$$\lambda_\Sigma(t) \approx \left(\frac{N_o}{\tau}\right) \text{ for } \left(\frac{t}{\tau}\right) << 1$$

$$\approx \left(\frac{No}{\tau}\right) \cdot \left(\frac{t}{\tau}\right)^{-1} \text{ for } \left(\frac{t}{\tau}\right) >> 1$$

We can also note that the same exponential character would be displayed even if time is "started" after some initial defect depletion period, since this affects only the initial defect population number, and not the exponential character of the ensuing failure rate curve.

DISTRIBUTED PROBABILITY OF DETECTION FAILURE MODEL

The deficiencies (relative to a Duane fit) in the previous model are largely corrected by simply introducing the concept that the defects within a given population have unequal probability of detection. That is, some are easy to detect while others are very difficult, requiring many test data samples to detect them. This concept agrees both with the empirical facts and our intuitive understanding of the test/failure process.

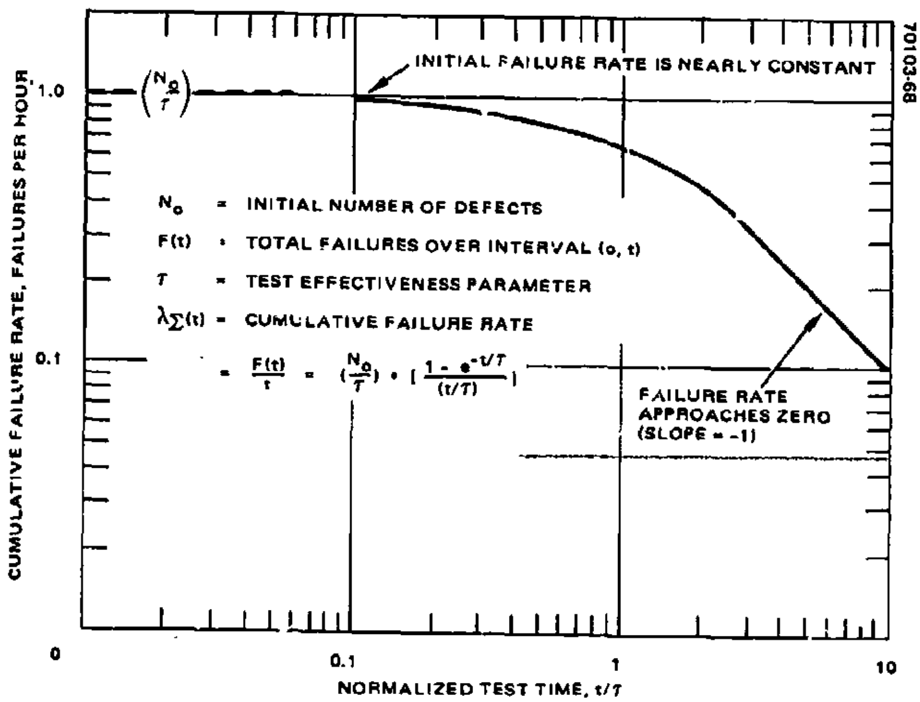


FIGURE B-2. EQUAL PROBABILITY OF DETECTION FAILURE MODEL – CONTINUOUS TIME VERSION

Decrease of Defect Population with Test Time

This approach is pursued mathematically by defining a defect density function, $D(v, t)$, which is both a function of time (since the defect population decreases with test time) and a function of the defect vulnerability to detection, v. In this illustrative case we consider the simple density function shown in Figure B-3. Here the defects within the population are distributed in terms of vulnerability to detection between the values $v = 0$ (impossible to detect) and $v = 1$ (easy to detect). In this example we also assume that the initial density is uniform; that is, the initial defect population has equal density at all values of detection vulnerability. (Further shaping of the resulting failure rate curve is accomplished by introducing a different shape for the initial density function, e.g., a Gaussian type curve).

We can therefore make the following definitions:

v = the relative vulnerability of a defect to being detected in test (resulting in a failure)

$D(v, t)$ = the defect density function, which, for any instant t, gives the number of defects having test vulnerability between v and $v + \Delta v$

$D(v, 0) = N_o$ = the initial defect density (assumed constant in this example) at $t = 0$

$N(t)$ = the total number of defects remaining in the population

At any instant the total number of defects remaining, $N(t)$, is found by integrating the density function $D(v, t)$ over the range $v = 0$ to $v = 1$, i.e.,

$$N(t) = \int_o^1 D(v, t)\, dv \qquad (3)$$

where, for $t = 0$,

$$N(0) = \int_o^1 N_o\, dv = N_o$$

is the total number of defects initially.

The decrease in the defect population (as a result of test action) is obtained from the differential equation,

$$\dot{D}(v, t) = -\frac{v}{T} D(v, t) \qquad (4)$$

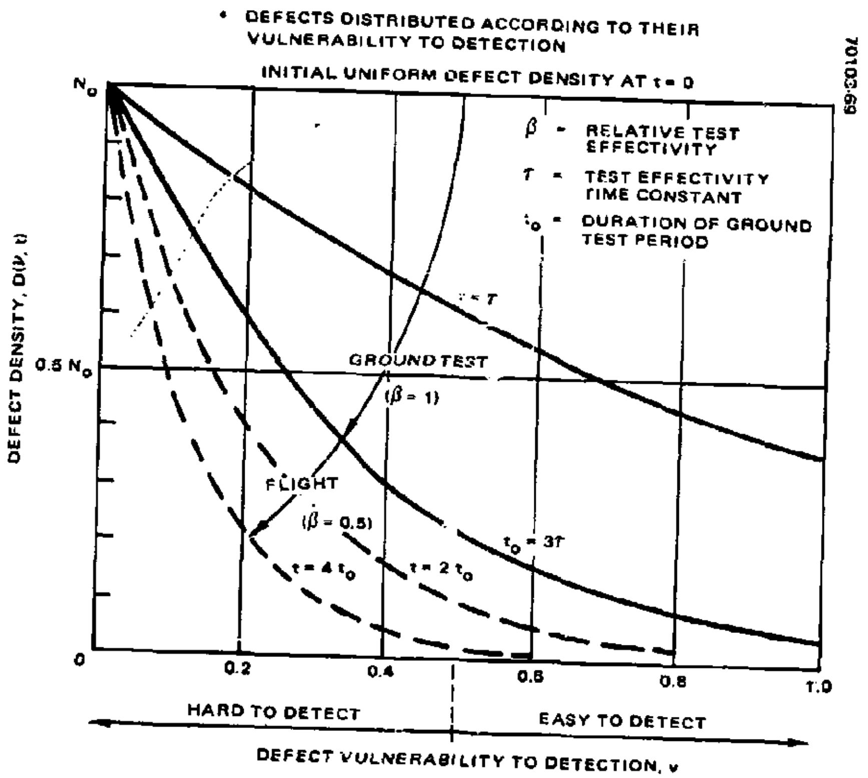


FIGURE B-3. DEFECT POPULATION DENSITY VERSUS TIME where,

τ = test effectiveness time-constant (a parameter of the test process).

Equation 4 states that the rate of decrease of the defect population at a particular vulnerability level, v (i.e., the population contained in the vulnerability range of v to v + Δv) is proportional to the test vulnerability of the defects as well as the number of defects remaining. The solution to Equation 4, giving the density as a function of time is then,

$$D(v, t) = D(v, 0) \cdot e^{-v \frac{t}{\tau}} \tag{5}$$

and, since for the particular case being considered, $D(v, 0) = N_o$

$$D(v, t) = N_o \cdot e^{-v \frac{t}{\tau}} \tag{6}$$

Figure B-3 shows a plot of Equation 6 for several values of t, illustrating the resulting decrease of the defect population with test time. As indicated, the population is depleted rapidly in the region of highest vulnerability, and more slowly elsewhere, according to the exponential relation (Equation 6).

The effect of a change in test effectivity can be accounted for by introducing one more parameter, β, defined as follows:

β = relative test effectivity

Thus, in the present case we assume a ground test period of duration $t_o = 3\tau$, followed by an in-orbit operation period. The relative test effectiveness of the in-orbit operation is taken to be β times the effectiveness of ground test. The differential equation describing defect population decrease in orbit is then,

$$\dot{D}(v, t) = -\beta \cdot \frac{v}{\tau} \cdot D(v, t) \text{ (for } t > t_o) \tag{7}$$

which has the solution,

$$D(v, t) = N_o \, e^{-v \frac{t_o}{\tau}} \cdot e^{-v \frac{t_o}{\tau} \beta \left(\frac{t}{t_o} - 1 \right)} \text{ (for } t > t_o) \tag{8}$$

Figure B-3 shows the plot of Equation 8, illustrating the further decrease in defect population for in-orbit operation relative test effectivity of $\beta = 0.5$.

Resulting Failure Rate

The failure rate implied by the defect population model described above is formed as follows:

$N(t)$ = total number of remaining defects

where,

$$N(t) = \int_0^1 D(v, t)\, dv \tag{9}$$

$F(t) = N_o - N(t)$ = total failures over the time interval $(0, t)$

$\lambda_\Sigma(t) = \frac{F(t)}{t}$ = cumulative failure rate

Then, after carrying out the inte_:ation in Equation 9, the cumulative failure rate (which combines both ground test and in-orbit failures) is determined to be,

$$\lambda_\Sigma(t) = \frac{N_o}{2\tau} \cdot \frac{2}{3(t/t_o)} \cdot \left[1 - \frac{1 - e^{-x}}{x}\right] \tag{10}$$

where

$$x = \begin{cases} \left(\frac{t_o}{\tau}\right) \cdot \left(\frac{t}{t_o}\right) & \text{for ground test } (t < t_o) \\ \left(\frac{t_o}{\tau}\right) \cdot \left[1 + \beta\left(\frac{t}{t_o} - 1\right)\right] & \text{in orbit } (t > t_o) \end{cases}$$

Figure B-4 shows a plot of Equation 10 for the specific parameters chosen for this example ($t_o = 3\tau$, $\beta = 0.5$). The following observations are relative to the character of the failure rate curve and the degree of its Duane fit.

1) Ground Test Period. Over the ground test period the average log-log slope of the curve is roughly -0.4, which ~~is in the range~~ of values typically used with the Duane model. Thus, the ground test curve can be fit reasonably well with a Duane curve (straight line on log-log paper) having a slope of -0.4. Further, "straightening" of the curve can be achieved by starting with a different initial defect distribution function — one that is skewed toward lower vulnerability rather than the simple uniform initial distribution used in this example.

2) In-Orbit Period. The portion of the curve corresponding to in-orbit operation has a slope which varies from -0.77 to -0.95 over the first decade. The average slope is -0.86 over the decade, and the curve can be fit very closely by a straight line of that slope. This agrees reasonably well with actual in-orbit failure rate data (e.g., Figure 5-20) for which the average slope for the F-1 through F-8 data is approximately -0.75.

Moreover, the transition from ground test to in-orbit operation, which shows up clearly in Figure 3-4 as a change in slope, can also be seen in the actual data. For example, the F-3 data shown previously in Figure 5-4 shows a similar slope change at launch.

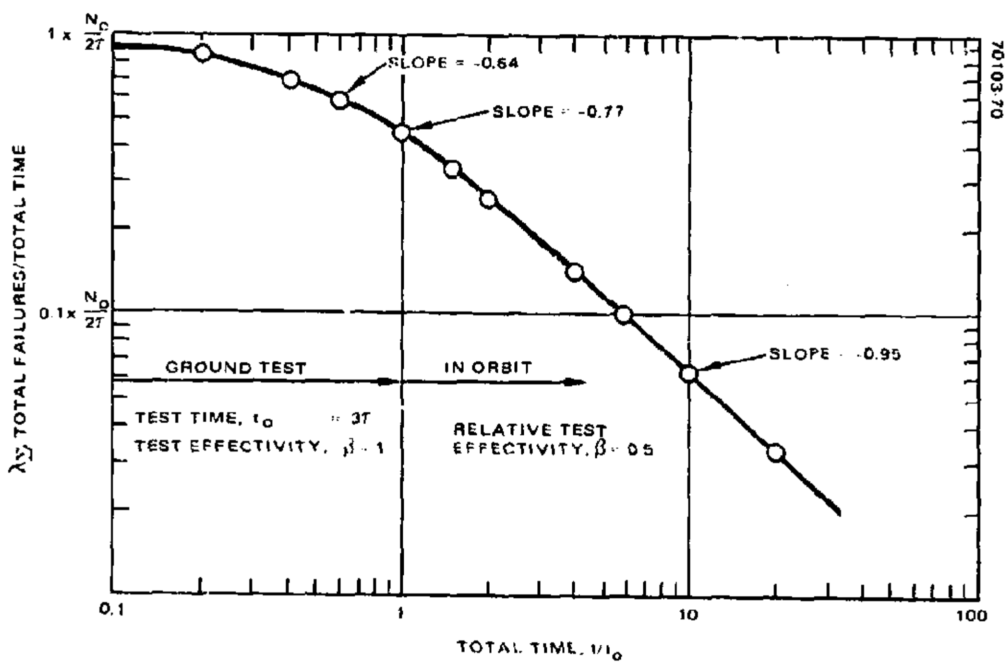


FIGURE B-4. DISTRIBUTED PROBABILITY OF DETECTION FAILURE MODEL COMBINED TEST AND IN-ORBIT FAILURES