A Career in Test and Evaluation

Reflections and Observations

From an Oral History Interview of
Charles E. "Pete" Adolph

Conducted by
Dr. Richard P. Hallion

Air Force History and Museums Program
in association with
Air University Press
1998

Disclaimer

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
1998	na/	

4. TITLE AND SUBTITLE	
A career in test and evaluation : reflections and observations from an oral history interview of Charles E. "Pete" Adolph	5a. CONTRACT NUMBER n/a
	5b. GRANT NUMBER n/a
	5c. PROGRAM ELEMENT NUMBER n/a
6. AUTHOR(S) Hallion, Richard	5d. PROJECT NUMBER n/a
	5e. TASK NUMBER n/a
	5f. WORK UNIT NUMBER n/a
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force History Support Office 3 Brookley Avenue Box 94 Bolling AFB DC 20032-5000	8. PERFORMING ORGANIZATION REPORT NUMBER n/a
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) n/a	10. SPONSOR/MONITOR'S ACRONYM(S) n/a
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) n/a

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

n/a

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Richard I. Wolf
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	221	202-404-2186

Dedicated to the Memory of Jack Strier (1921–1995)

Jack Strier spent his flight-test career—from 1952 until he retired in 1980—at Edwards Air Force Base, California. His legacy was not the aircraft he helped to perfect, but rather the people he mentored, taught, and developed. He was a United States Marine Corps aviator during World War II and flew F–4U Corsairs and F–6F Hellcats. After his release from active duty in 1946, he returned to the University of California at Los Angeles and earned a bachelor of science degree in mechanical engineering in 1949. Following a two-year stint with Hughes Aircraft, he accepted a position as the performance and flying qualities engineer on the H–23B helicopter, followed by assignments as project engineer on the B–36F bomber heavyweight performance tests, the H–19B helicopter, the Cessna XL–19C turboprop, and the YF–84J and F–104A jet fighters at Edwards in the Performance and Flying Qualities Engineering Branch. In 1956 he was promoted to engineering supervisor, and in 1960 he became assistant branch chief. As a supervisor, Jack participated in the test and evaluation of nearly every aircraft to enter the Air Force inventory over the next twenty years. Mr. Strier also contributed to a wide range of Air Force source selections and participated in record-setting attempts with the F–104 and F–106 as project engineer, advisor, and National Aeronautics Association observer. While involved in these projects, Jack trained a generation of flight-test engineers in the basics of performance and flying qualities testing. His skills as a teacher of the fundamentals of test planning, test conduct, data analysis, and reporting were without parallel. He personally mentored more Air Force personnel in the rudiments of flight testing than any other individual. His former students can be found today throughout the flight-test community with aircraft manufacturers, the Federal Aviation Administration, the National Aeronautics and Space Administration, and the Department of Defense.

Contents

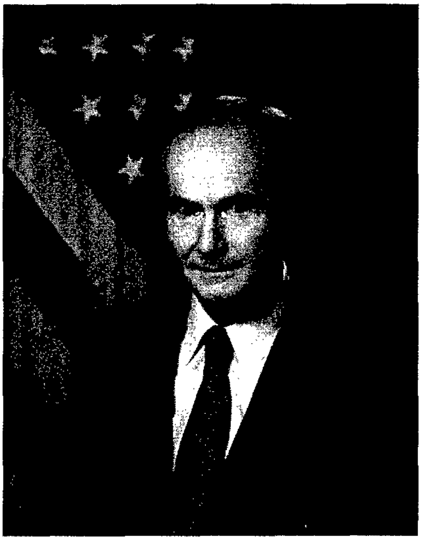
Charts

Foreword

Charles E. "Pete" Adolph retired as Director of Test and Evaluation in the Office of the Under Secretary of Defense (Acquisition and Technology) on 31 January 1994. This completed more than 30 years of federal service—almost all of it within the challenging field of test and evaluation (T&E). Pete—as he was widely known throughout the Department of Defense testing community—enjoyed a remarkable career. It began in the late 1950s, as the heroic era of flight test in the first decade of the jet age was drawing to a close. Pete then played an increasingly prominent role in the transformation of flight testing into a systematic discipline using the latest in information technology to evaluate sophisticated weapon systems. His government career culminated as a senior director at the Office of the Secretary of Defense (OSD). There, in the Pentagon, he brought his many years of engineering and management experience in the field to bear upon the formulation of policies for the acquisition and testing of weapon systems in the post-cold-war era. The text that follows began as a series of five oral history interviews conducted in the Office of the Air Force Historian between 29 July 1993 and 15 April 1994. Ms. Pauline Tubbs of the United States Air Force Historical Research Agency at Maxwell Air Force Base (AFB), Alabama, expertly transcribed these interviews from approximately eight hours of audio tape. Mr. Lawrence R. Benson, the Air Force Historian's Assistant for Field Programs (and previously the Director of Research Services at the Air Force Operational Test and Evaluation Center), organized, revised, and edited the transcript—adding explanatory material in brackets or footnotes as appropriate. Mr. Adolph was accompanied at most of the interviews by Mr. Douglas Nation of the 46th Test Wing at Eglin AFB, Florida, who was on a special assignment to the OSD T&E Directorate. Dr. James O. Young, Historian of the Air Force Flight Test Center (AFFTC), and his staff at Edwards AFB, California, helped with details on flight test and provided most of the photographs. Although Mr. Adolph's responsibilities within the OSD encompassed testing of all types of systems throughout the four armed services, our interview focuses most sharply on Air Force flight testing at Edwards. This is where Pete spent the majority of his career, and where I first met him in 1980 after becoming the AFFTC Historian.

RICHARD P. HALLION
Air Force Historian

About the Authors



Charles E. "Pete" Adolph

Charles E. "Pete" Adolph was born in Akron, Ohio, on 3 December 1935. He graduated from Saint Mary High School in Akron, received a bachelor of science degree in aeronautical engineering from Saint Louis University in 1956 and was commissioned in the Air Force. His later education included a master of science (M.S.) in aeronautical and astronautical engineering from the University of Michigan in 1964, and another M.S. in systems management from the University of Southern California in 1970. During 1973–1974, he was a Sloan Fellow at the Stanford University Graduate School of Business. Mr. Adolph began his engineering career in 1956 in San Diego with the Convair Division of General Dynamics Corporation. He worked for Convair at Edwards Air Force Base, California, on the F-102 and F-106 flight test programs. In 1958 he was assigned as a second lieutenant in test engineering at the Air Force Flight Test Center (AFFTC) at Edwards. In 1960 Mr. Adolph accepted a civil service appointment at Edwards, where, during the next quarter century, he held a wide variety of engineering and systems acquisition management positions in AFFTC. He was a Project Engineer and Senior Project Engineer in the Flight Test Engineering Division from 1960 to 1964, the Chief of Fighter Projects in the Performance and Flying Qualities Branch from 1964 to 1974, the Assistant Chief of the Systems Engineering Branch from 1974 to 1977, the Project Manager of the Flight Test Mission Control Complex from 1977 to 1979, the Chief of the Development Division from 1979 to 1982, and the Technical Director of the 6520th Test Group from 1982 to 1985. He became the Technical Director of the AFFTC—its senior civilian position—in 1985. While at Edwards, he was involved in the planning and execution of numerous combined tests (developmental/operational), including the F-15 and F-16 fighters and the air launched cruise missile, as well as the testing of prototype programs, including the lightweight fighter (YF-16 and YF-17), close air support attack aircraft (A-9 and A-10), and advanced medium short takeoff and landing transport (YC-14 and YC-15). Mr. Adolph also was the project manager for the development and integration of the Flight Test Mission Control Center and was responsible for managing the acquisition of new test support equipment for Edwards and its ranges. Special assignments while at Edwards included serving as the Air Force project officer for the preparation of MIL-S-86391A (USAF), "Stall/Post-stall/Spin Flight Test Demonstration Requirements for Airplanes," 1970 to 1973, AFFTC Executive Committee Member, Range Commander's Council, 1982 to 1985, Chairman of the Flight Test Techniques Group of the NATO Advisory Group on Aerospace Research and Development, 1985 to 1987, and engineering supervisor for classified project activities, 1967 to 1973. In August 1987 Mr. Adolph was selected as the Deputy Director, Defense Research and Engineering, Test and Evaluation (T&E) in the Office of the Secretary of Defense (OSD), where he was responsible for Department of Defense-wide oversight and policy for developmental T&E. The Secretary of Defense designated Mr. Adolph to perform the duties of Director of Operational Test and Evaluation from July to November 1989 and the duties of Director of Defense Research and Engineering from May to December 1991. He also chaired a congressionally mandated Federal Advisory Commission on the consolidation and conversion of Defense Research and Defense Laboratories and was architect of a plan to reduce duplication of T&E facilities throughout the Department of Defense (DOD). As part of a 1992 reorganization in OSD, Mr. Adolph was made Director, Test and Evaluation for the Under Secretary of Defense, Acquisition and Technology. His duties included the assessment of weapon system development and live fire test results for the Defense Acquisition Board, the oversight of test range improvements, and management of DOD's joint test programs. Mr. Adolph is the author of more than fifty test reports, technical papers, management studies, and articles. His awards include the Distinguished and Meritorious Executive Presidential Rank Award (1990, 1992), the DOD Medal for Distinguished Civilian Service (1994), the International Test and Evaluation Association Award for Distinguished Achievement in Test and Evaluation (1993), the Secretary of Defense Meritorious Service Award (1989, 1991), the Air Force Meritorious Civilian Service Award (1987), the Kelly Johnson Award for Outstanding Achievement in Flight Test Engineering (1986), the Presidential Management Intern Alumni Group Outstanding Public Service Award (1986), and the City of Los Angeles Certificate of Tribute for Working to Improve the Status of Women (1984).

Dr. Hallion, an internationally recognized aerospace historian, has been the Air Force Historian, Washington, D.C., since December 1991. He is responsible for directing the worldwide Air Force historical and museum programs. He earned the Bachelor of Arts and the Doctorate of Philosophy degrees from the University of Maryland, and is a graduate of the National Security Studies Program, John F. Kennedy School of Government, Harvard University. Dr. Hallion has served as Curator of Science and Technology at the Smithsonian Institution, where during the Persian Gulf War he was the Charles A. Lindbergh Professor of Aerospace History at the National Air and Space Museum. A former professor at the Army War College, he is the author of fifteen books relating to aerospace history. His publications include *Air Power Confronts an Unstable World* (1977), *Storm over Iraq: Air Power and the Gulf War* (1992), *Strike from the Sky: The History of Battlefield Air Attack, 1911–1945* (1989), *Test Pilots: The Frontiersmen of Flight* (1988), *The Wright Brothers: Heirs of Prometheus* (1978), and *Supersonic Flight: Breaking the Sound Barrier and Beyond* (1972). Dr. Hallion has flown a wide range of military and civilian fixed- and rotary-wing aircraft, has broad experience in museum development, historical research, management analysis, and has been a consultant to a variety of professional organizations.

Dr. Richard P. Hallion

Oral History Interview of Charles E. "Pete" Adolph

By Dr. Richard P. "Dick" Hallion

Classic Flight Testing at Edwards

H*: Pete, let's begin by talking about how you as a second lieutenant were involved in testing at Edwards Air Force Base (AFB), California, during its golden age. One project you worked on in particular was the Republic F–105 Thunderchief fighter-bomber.

A: That's correct, Dick. It was the F–105B stability and control flight-test program, and I was the assistant to the project engineer. As I recall we had a number of F–105 project pilots at the time. Maj. Bob Titus, who later on became General Titus, was the stability and control project pilot, and Donald K. ["Deke"] Slayton, who later became an astronaut, was the overall project pilot for all F–105 testing at the time. Bob White, who later was the X–15 project pilot and later became an Air Force general [and commander of the Air Force Flight Test Center (AFFTC) at Edwards], also flew in the program—he was the performance pilot. Howard Lane, who later became General Lane, was the chief of fighter test operations. My boss was Capt. Phil Conley [then a flight-test engineer]. He wrote my first OER [officer evaluation report]. He also became a general officer as well as commander of the AFFTC. In retrospect it was a very notable group of test pilots and engineers. They all became general officers.

It was a very challenging and interesting time. We were doing what at the time was a classic Air Force stability and control test. We didn't have any real-time telemetry. Basically, the project engineer and I would prepare the cards before the flight and review them with the pilot. We then went over to the radio room and recorded the start and stop times for the various maneuvers as the pilot relayed them to us. We would also check on center-of-gravity (COG) conditions and the various flight conditions. There was a radio room in the test operations building which we would use to

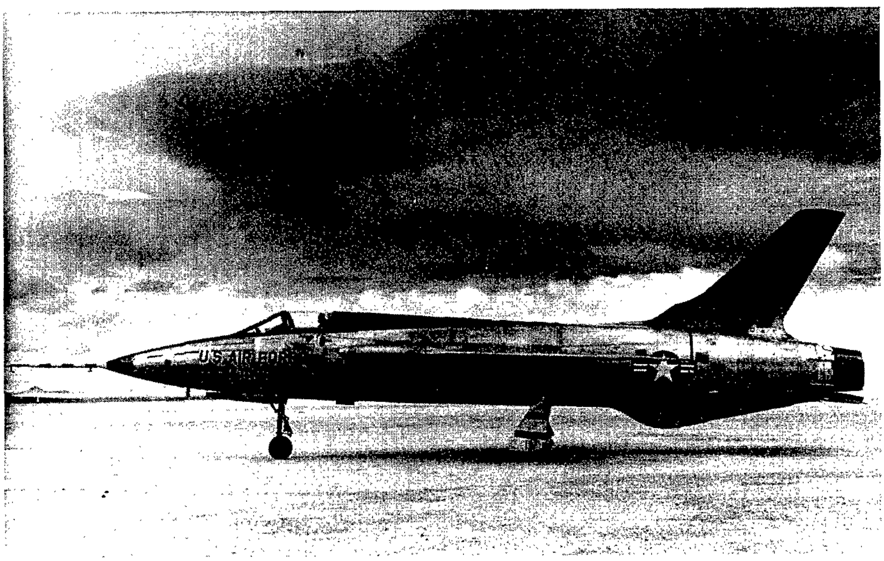
*"H" indicates Dr. Hallion; "A" = Mr. Adolph; "B" = Mr. Benson; and "N" = Mr. Nation.

monitor missions. Prior to going in the Air Force, I had worked for a year as a test engineer for Convair, San Diego, on the F-102 [Delta Dagger] and initial F-106 [Delta Dart] flight-test programs. I later was the government project engineer on a limited test of the [North American] F-100F [Super Sabre] to look at some changes between the initial F two-place version and the so-called dash twenty. In that test we operated out of the contractor facility at Palmdale, California, where the F-100 was being produced. We did some of the data reduction and analysis at the North American facility in El Segundo. The F-105 program was more of an autonomous Air Force activity. For the F-100 program, we relied to a large extent on the contractor for data-processing support. So it was my first very limited exposure to contractor support.

H: The flight testing here seems not all that different in many respects from what was undertaken in the thirties and forties. The card would be prepared to govern the various test conditions that you would examine, and then you would come back and a record would be pulled from the aircraft and the data would be reduced. So the real-time telemetry you mentioned is really not yet a factor. Out of curiosity, how many test conditions would you evaluate on one of your F-105 stability and control investigations?

A: I don't recall exactly, but typically in those days, you would look at conditions at one COG position. We would try to get, to the extent feasible, three COG conditions to evaluate longitudinal stability characteristics, and you would do that by ballasting the airplane and fuel sequencing. You might do a series of stability maneuvers at two or three Mach numbers at a given altitude. The maneuvers would consist of dynamic longitudinal stability tests, lateral directional stability tests, and roll performance. The final test at the condition was a windup turn where the pilot would pull into the maximum allowable G, whether that be dictated by a lift limit or some structural limit. You would do this "stability block" at several Mach numbers at one or two altitudes, depending on the fuel available. Typically, after you defined the envelope at one center of gravity, you would then ballast the airplane and do a series of flights at another center of gravity. The testing we were doing at the time consisted of a detailed definition of characteristics within an already cleared flight envelope. The Air Force, at that time, was not involved in what would be considered flight envelope expansion. As testing evolved to a combined test environment, which we will discuss later, the Air Force and the government became more heavily involved in actual airworthiness, envelope expansion testing. Over time, we have gone to a more integrated test activity.

Back in the late fifties and early sixties, the contractor was totally responsible for expanding the flight envelope. The government would then accomplish a detailed definition of aerodynamic characteristics, airplane performance, and stability and control characteristics within an already cleared flight envelope. There was a significant amount of redundancy between the government and contractor test programs during this era. In addition we weren't keeping up with the data reduction because there were two engineers and only two or three engineering aides helping us. In those days everything was recorded on a photo panel or oscillograph. About all we did between flights was take a look at the data, along with our instrumentation people, to make sure that the instrumentation was functioning properly. We didn't have the time for a complete data analysis until a break would occur in the program as a result of an aircraft layup for airplane or instrumentation repairs. We would try to get off two or three flights a week, and on the weekends we began the detailed analysis. Saturday was a normal workday. Typically, the analysis and the reporting wasn't complete until several months or as much as a year after the flight testing was complete. With very few exceptions, we didn't have to go back and look at data to define what conditions would be flown on the next test.

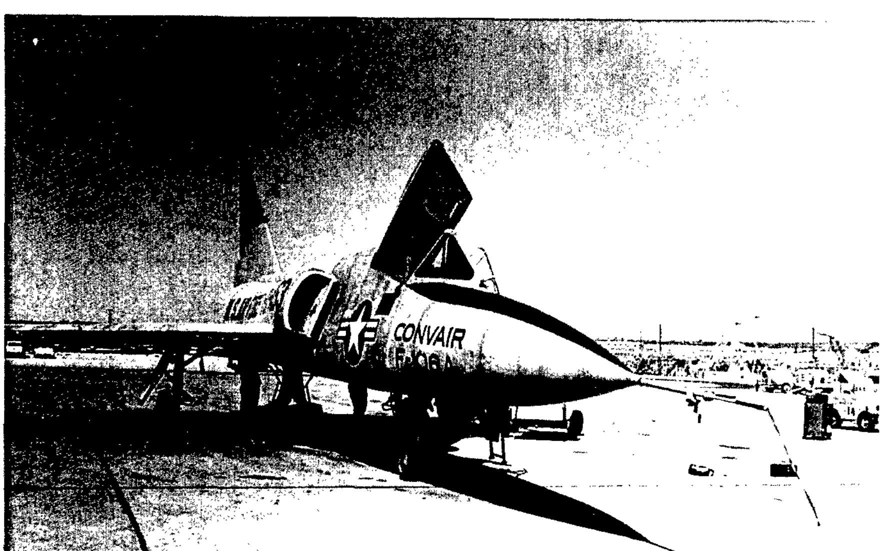


Republic F–105 Thunderchief parked on Rogers Dry Lake at Edwards AFB, California, in the mid-1950s.

H: Were you looking at macro trends? For example, if you were looking at a stability derivative of some sort to see if there was an overall decline in aircraft stability, did you concentrate on the next test point, which might be on a following flight, where for example there might be a safety concern? Or did you look at what was happening at every stage of the flight?

A: Yes, but that sort of thing, even then, was by exception. Typically, the limits had to do with aft COGs, load factor, or roll rates. The limit had already been defined by the contractor, so if we were approaching the test at a limiting COG condition, we would look at the data at a forward and mid [position] and make an extrapolation. If we were evaluating either roll response or roll rate at near peak conditions, we would look at the data from preceding tests. We weren't in a mode of having to review most of the data in order to progress with testing. It was by exception only.

H: In the 1950s one of the problems that had come up very quickly was inertial coupling, which was exemplified on a service airplane, the F–100, and on experimental aircraft such as the X–2 and the X–3. Did the testing you were doing on the F–105 reflect the concern people had at the time with things like inertial coupling and the development of damping and stability augmentation technologies? Were you doing a lot of work in those areas?



Convair Delta Dart on ramp at Edwards in 1958, with a Coke-bottle shape of fuselage clearly visible.

A: There was a lot of work in those areas, but, again, in contrast with the way things are done today, we were using very conservative flight envelope limits which had been imposed as the result of earlier contractor testing. In other words in the inertial coupling area, we were limited to ensuring that we stayed well within a well-defined flight envelope.

H: In the course of flight testing the F-105 when you were at Edwards, did you lose an F-105 in flight test? The contractors lost some.

A: In the F-105B Category II testing there were none lost at Edwards, but there was at least one lost at Eglin [AFB, Florida]. Lt. Gen. Howard Leaf, USAF, Retired, then current Air Force Director of Test and Evaluation (T&E), made a high-speed ejection from an F-105 at Eglin.

H: The F-105 always seemed to be one of the better of the Century series if you compare the developmental teething troubles of aircraft like the F-100 and the F-104. One of the other airplanes you mentioned when you were starting out was the Convair F-106, which, at least in its design development, had started out originally as a relatively simple, straightforward adaptation of the F-102. But very quickly Convair found that it wasn't going to be more complex. Do you have any reflections on the F-106 experience? And did that teach you any lessons for your later career?

A: I was only involved with the F-106 for a few months working for Convair very early in the flight-test program. I recall the F-106 made its first flight on the day after Christmas, 26 December 1956. Dick Johnson was the pilot. I didn't have any further close association with the F-106 test program until a later model of the F-106. It obviously represented a very substantial improvement in performance over the F-102. It had the attendant developmental problems, in retrospect, that one might expect with that kind of a performance improvement, basically a Mach 2 airplane. It evolved into a very capable and very reliable platform. There were some similarities and some differences in the F-106 and the F-105 developments, but the J75 engine, which was common to the two aircraft, turned out to be a very reliable engine and an outgrowth of the J57.

H: And is still with us in the TR-1 [now redesignated as U-2R].

A: Yes, that's right. When you contrast the F-105 and the F-106 with the F-104 development test (DT) program, the F-104 represented a significant change in both aerodynamics and propulsion. In retrospect where you have a new propulsion system combined with a new aerodynamic concept on a single-engine airplane, one could expect developmental problems.

Concurrent Testing and Production: The Case of the F-111

H: The whole question of engine inlet matching has historically been a problem for jet airplanes, all the way up to, in some cases, the present day. It seemed particularly serious in the experience of the F-111 in the early 1960s. You left active duty with the Air Force [in 1960] and went into the federal civilian service side of flight test. Did you become involved with the F-111?

A: I was the original Air Force performance and stability project engineer on the F-111 program and was later the fighter performance engineering section chief responsible for all the F-111 performance and stability testing. I was involved directly in or in overseeing all of the F-111 performance and flying qualities testing from 1964 through 1973.

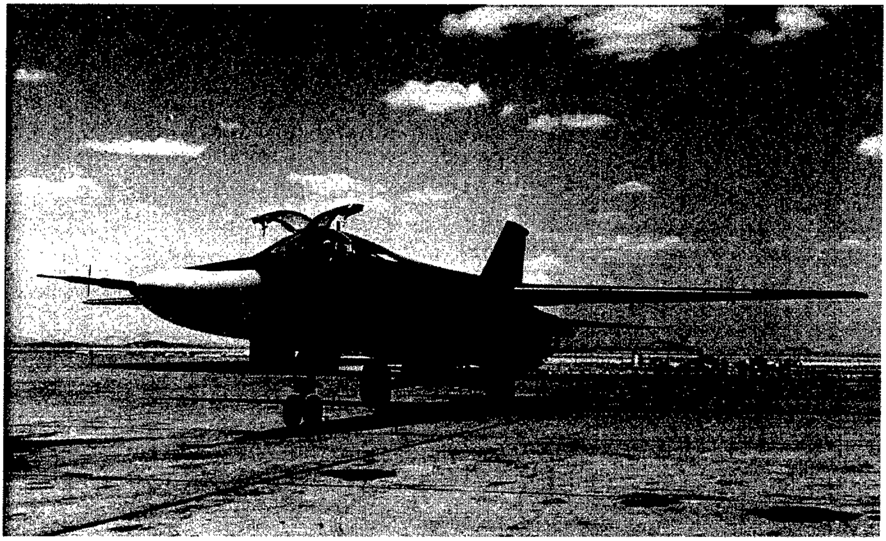
H: That was a fascinating time period because the F-111 was a program that had major national visibility, and there was a tremendous personal interest by then Secretary of Defense Robert S. McNamara in the flight-test progress in the airplane because the aircraft to a great degree had been his pride and joy, a multipurpose aircraft for both the Navy and the Air Force.* Did the fact that it had a very high-level focus on it impinge upon you in the way that you were able to do your job in the flight-test arena?

A: Yes. The Air Force flight-test team participation early in the program was somewhat limited, in fact, I would say very limited.

H: Were you constrained in terms of the kinds of flight-test areas that you could look at?

A: Col. Jim Wood was the chief Air Force test pilot, and I was the test engineer. My recollection is that he was not allowed to fly the airplane until much later than we had hoped. We were very limited in what we could look at very early in the program. It is fresh in my mind because we could have provided objective inputs as to the adequacy of the airplane and the way the test program was progressing.

*McNamara centralized much of the acquisition process within the Office of the Secretary of Defense (OSD) and implemented a complex and concurrent "total package procurement concept." He also believed the separate services had practiced unnecessary duplication in developing their own systems and was determined that the F-111—known as the tactical fighter experimental (TFX) prior to 1963—would meet the needs of both the Air Force and the Navy.



General Dynamics F–111A (formerly TFX) in the mid-1970s, with facilities of Edwards AFB Area 003 in background.

I think the F–111 was a classic example of the downside of concurrency.* I don't recall the exact number, but there were 130 or 140 F–111As produced [actual number: 141 plus 17 for research, development, test and evaluation (RDT&E)]. They were into production well before the flight-test program had progressed to the point where they could stabilize the configuration of the aircraft. It was a classic example of a highly concurrent program with production well under way before the flight-test program had identified the major deficiencies.

H: There was a plan in the early 1950s called the Cook-Craigie plan** that encouraged concurrency. Some have looked at the story

*Concurrency refers to the closely timed development, testing, production, and fielding of weapon systems in which the various phases overlap or occur simultaneously rather than proceed step by step at a more deliberate pace. The successful application of concurrency in the Air Force's crash program to field the first generation of intercontinental ballistic missiles (ICBM) led to its adoption as the favored acquisition strategy for manned aircraft as well—with less favorable results. For the definitive account of the development of concurrency, see Jacob Neufeld, *Ballistic Missiles in the United States Air Force, 1945–1960* (Washington, D.C.: Government Printing Office, 1990).

**In 1954 Lt. Gen. Orville R. Cook, the Air Force Deputy Chief of Staff (DCS) for Materiel, and Lt. Gen. Laurence C. Craigie, the DCS/Development, agreed on the weapon system concept by which a single project office in the Air Research and Development Command would manage all aspects of an acquisition program prior to a production decision, when the responsibility would transfer to the Air Materiel Command. This and other developments in Air Force acquisition management are summarized by Michael H. Gorn in *Vulcan's Forge: The Making of an Air Force Command for Weapons Acquisition (1950–1985)* (Andrews AFB, Md.: Headquarters Air Force Systems Command, 1989).

of the F-102 as a classic and early example of how the weaknesses of that concurrency approach resulted in the Air Force acquiring a full fleet of aircraft, many of which up to, say, forty or fifty of the initial productions models bore very little resemblance to one another. They had air dynamic configuration changes, structural changes, propulsion differences, and all that sort of thing, yet it seems that there was no lesson really learned here. This approach then was also followed with the F-111. Would you elaborate a bit on your philosophy of concurrency and when the approach is appropriate and when it is inappropriate.

A: I have seen the pendulum swing from total package procurement concurrency under McNamara, typified by the C-5 and the F-111, to the emphasis on prototyping by Mr. Packard* which resulted in the A-X competitive flyoff [A-9 versus A-10], the lightweight fighter [YF-16 versus YF-17], and the advanced medium STOL [short takeoff and landing] transport [AMST—YC-14 versus YC-15]. The pendulum then swung back in the direction of concurrency with the B-1. Many people make the argument that in today's environment, with the improved suite of tools that we have available—wind tunnels, simulators, and computational fluid dynamics—that flight testing is—at least in the aerodynamics, propulsion, and flying qualities areas—a fine-tuning process. I would agree that there are fewer major surprises today than there were in the past. Having said that, I think a limited but very aggressive flight-test program on a production-representative article is essential. That can be done easily in a period of a year or a year and a half, based on the lightweight fighter, A-X, AMST, and YF-22/23 [advanced tactical fighters—ATF] experience. In each instance in a period of roughly a year, they completely opened the flight envelope and explored aerodynamics, propulsion, and flight worthiness. It has been repeatedly shown that we can, within a period of a year, do a very comprehensive test of a platform. I think that in today's environment, with the possible exception of a cargo or a bomber type aircraft, that some sort of DEM/VAL [demonstration and validation] effort is definitely needed. The program needs to be structured to do it on an accelerated basis.**

*David Packard, the acclaimed cofounder of Hewlett Packard Corporation, who served as Deputy Secretary of Defense from 1969 to 1971 and led the influential Presidential Commission on Defense Management Reform from 1985 to1986.

**For more on Mr. Adolph's views on enhancing T&E in general and flight testing in particular, see appendices A and I.

H: At the level that you were in testing in the 1960s, did you feel that the concerns of testers were adequately taken into consideration by system program offices [SPO], and beyond that by the acquisition leadership within the service? Did you feel that flight testing should have been more represented in higher level decisions?

A: Let me go back and try to characterize it as I recall it. Again, to go back to the late fifties and perhaps the early sixties, the role of the government in developmental flight-test—as I mentioned in conjunction with the F-105 and the F-100—was largely characterizing the performance and the flying qualities or the weapons delivery characteristics of the platform within an envelope that had already been defined and cleared by the contractor, the old phase system of testing, phase one through phase five. The F-111 program was the first new development of an aircraft that I was involved with under the category system of testing (see Fig. 1).*

The F-111 was also the most highly politicized flight-test program of that era. The objective was to build an aircraft that would satisfy the needs of the Air Force in a tactical and strategic bomber role, satisfy the US Navy requirement for an outer air battle air superiority aircraft, and meet a United Kingdom and Australian requirement as well. After the Navy pulled out, we still had the Australian test pilot (Ron Green) and test engineer (Harry Walton) with us at Edwards for a period of four or five years. The role that the test center performed was quite different in the case of the F-111 because we were involved, at least as observers, from the outset. It was a frustrating experience. The role of the flight-test center as a government test agency in a program of that type had not crystallized. There were a number of developmental problems that arose. The problems were obvious to us, but we (the government test community) were constrained from really articulating those problems.

H: There were two sets of congressional hearings on the F-111. One set was in 1963, largely concentrated on the contract award. If memory serves me right, the second set of hearings was in 1969, running perhaps into 1970, and it was on the evolution of the program of flight-test deficiencies and then the Combat Lancer experience, where six F-111s went over to Southeast Asia and three were lost [in only five weeks]. Did you folks ever get involved at Edwards

*The evolution of Air Force test procedures can be traced through numerous editions of Air Force Regulation (AFR) 80-14, *Test and Evaluation*, published from 11 September 1951 through 3 November 1986 and now encompassed in Air Force Policy Directive 99-1 and its implementing Air Force Instructions (AFI), especially AFI 99-101, *Developmental Test and Evaluation*, 22 July 1994, and AFI 99-102, *Operational Test and Evaluation*, 22 July 1994. For a study of flight test through 1980, see Larry G. Van Pelt, "Flight Test Concept Evolution," USAF Air War College Research Report MS 120-81, April 1981.

PHASE TESTING
1951–58

	TYPE	ORGANIZATION	AIRCRAFT
I	AIRWORTHINESS	CONTRACTOR	PROTOTYPES
II	CONTRACTOR COMPLIANCE	AFFTC	PROTOTYPES
III	DESIGN REFINEMENT	CONTRACTOR	PROTOTYPES
IV	PERFORMANCE & STABILITY	AFFTC	PRODUCTION
V	ALL WEATHER	AFFTC	PRODUCTION
VI	FUNCTIONAL DEVELOPMENT	AFFTC/LIMITED USING COMMAND	PRODUCTION
VII	OPERATIONAL SUITABILITY	AIR PROVING GROUND COMMAND/ USING COMMAND	PRODUCTION
VIII	UNIT OPERATIONAL EMPLOYMENT (1956–58)	OPERATIONAL UNIT	PRODUCTION

CATEGORY TESTING
1958–72

	TYPE	ORGANIZATION	AIRCRAFT
I	SUBSYSTEM DEVELOPMENT INITIAL AIRWORTHINESS	CONTRACTOR/ AFFTC	PROTOTYPE/ EARLY PRODUCTION
II	SYSTEM DEVELOPMENT (INTEGRATION OF SUBSYSTEMS INTO COMPLETE SYSTEM)	AFFTC/ CONTRACTOR	EARLY PRODUCTION
III	OPERATIONAL	USING COMMAND	PRODUCTION

COMBINED TESTING
1972–Present

	TYPE	ORGANIZATION	AIRCRAFT
I	DEVELOPMENT T&E INITIAL OPERATIONAL T&E	AFFTC/ AFOTEC	PREPRODUCTION
II	OPERATIONAL T&E DEVELOPMENT T&E	AFFTC/ AFOTEC USING COMMAND	EARLY PRODUCTION

Source: AFFTC History Office.

Figure 1. Major Types of Air Force Test and Evaluation

lost [in only five weeks]. Did you folks ever get involved at Edwards in the congressional hearings aspect of it?

A: No, we were not involved. Participants were from the program office and the principals in the Air Force R&D staff. Of course, we followed it to the extent we could, but we did not play in it. Again, going back to the program, the first year or two were frustrating for the reasons I mentioned. The Navy subsequently pulled out of the program. By the way, the F-111 was my first experience with a joint Navy/Air Force program. We had joint test planning working group meetings with the Navy and Grumman (who was the Navy prime) at Pax River [Patuxent River Naval Test Center in Maryland] and Calverton [Grumman's test facility on Long Island, New York] as well as GD [General Dynamics] Fort Worth. At the time the Navy was trying to pull out of the program from the outset, and we could never reach agreement on many aspects of the flight-test program.

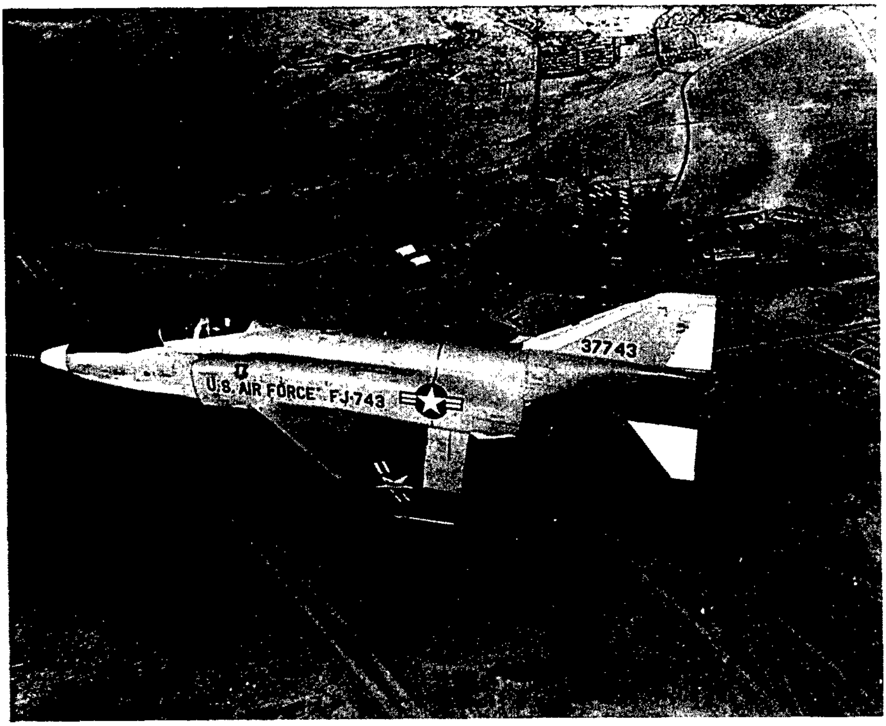
H: Was that because they wanted to go the route of the VFX, which became the F-14?

A: Yes, it was a learning experience for me in that it was my first brush with some of the—I don't want to use the word "political" but can't think of a better one—pressures other than technical that get involved in shaping and structuring a test program.

H: The external pressures, if you will?

A: The external pressures. After the first two years of the F-111 program, the AFFTC center settled into a more classical role. Once we got a production-representative platform, we went on to characterize the performance, the flying qualities, and the various system test programs.

I later on became the fighter engineering section chief, reporting to Don Smith, who was the chief of the Performance and Flying Qualities Flying Branch. Dick Hildebrand was the engineer on the F-111 performance program. Pete Hoag was the project pilot. Jim Papa, a lieutenant at the time, was the stability and control project engineer, and Dave Livingston was the project pilot. As I mentioned, we had the Australians working directly with us, and Harry Walton and Ron Green were the engineer and pilot, respectively. All of that progressed well. The next significant event was the high angle of attack (AOA) and the stall and spin testing. At that time we were heavily involved in F-4 and F-111 stall/spin test activities.



McDonnell F–4C Phantom over Edwards AFB in the early sixties.

The Quest for Improved Aircraft Performance: Departure and Stall/Spin Testing

H: Was that the Jerry Gentry episode?

A: Yes. He and Colin "Mac" McElroy ejected from the F–4. My recollection is that, within a year, Pete Winters and Pat Sharp ejected from the F–111 during a spin. Stall/spin testing as a discipline was really going through an evolution during that period. The high AOA work done on the F–4 and the F–111 resulted in the need, in our minds, to drastically alter the high AOA design and test criteria. I can remember numerous discussions and arguments on the subject. In earlier programs, contractors defined the envelope and then the government evaluated high AOA characteristics within the cleared envelope. With the F–4 and F–111, the government got involved for the first time with envelope expansion.

At the time, spin test criteria were written based on the concept that you would force the airplane into a fully developed spin, with full pro-spin controls for three full turns, then apply recovery controls, and then be able to recover. This philosophy was okay for a trainer but didn't make sense on a fighter or an attack aircraft where recovery prior to a fully developed spin was preferred. The philosophy was evolving and two of the key engineers in this were Burt Rutan (F-4 flight-test engineer) and Pat Sharp (F-111 flight-test engineer).

H: Rutan was a flight-test engineer?*

A: Yes. Burt was the lead flight-test engineer on the F-4E.

H: As a civilian?

A: As a civilian. Pat was working for Burt for a while on the F-4 and then became the flight-test engineer on the F-111 High Angle of Attack Program.

At that time we all began to question why would you want to force a high-performance airplane into a spin? Basically, what we wanted to look at was first departure resistance and then be able to recover from an incipient spin rather than allowing or forcing the airplane to progress to a fully developed spin.

H: A pre-spin condition?

A: Yes, a post-stall/pre-spin condition. I can recall meeting with the Navy on this. The contractor [McDonnell] F-4 spin test pilot was Jack Krings.** I also recall meeting at Pax and Langley AFB, Virginia, with the National Aeronautics and Space Administration (NASA) spin tunnel expert, Jim Bowman. Jim was extremely helpful to us on all our spin programs.

It was a major cultural change to adopt the philosophy of spin avoidance and recovery during the incipient phase of departure. We finally recognized the need to emphasize departure resistance in design and testing. High angle of attack test programs needed to focus on exploring departure and incipient characteristics rather than applying full pro-spin controls for several turns. First priority was departure resistance and then recovery from incipient spins. Full pro-spin controls on a prolonged basis should be explored only on an exception basis and then only for trainers. It was a total change in the philosophy, which involved the Air Force Flight Dynamics Lab and engineers of the Aeronautical Systems Division

*Burt Rutan later gained fame as the designer of the globe-circling *Voyager*, flown nonstop around the world by his brother Dick Rutan and copilot Jeana Yeager.

**John E. Krings, who served as DOD's first congressionally confirmed Director of Operational Test and Evaluation (DOT&E) from 1985 to 1989.

[ASD]* as well as the Edwards test community. In fact, the Air Force sponsored a stall/spin symposium at Wright-Patterson AFB (WPAFB), Ohio, from 15 to 17 December 1971. The proceedings from that symposium contain the most comprehensive collection of high AOA information in existence.

I was the Air Force project manager for development of what became the Air Force spec—later MIL-S-83891 (USAF), "Stall/Post Stall/Spin Flight-test Demonstration Requirements for Airplanes." Sharp, Burt Rutan, and McElroy were the people who provided the engineering expertise—three very bright young men. We developed the new spin spec as a result of the F-4 and F-111 experiences. The philosophy of departure resistance was embodied in the design of the F-15 and the F-16. That was a significant change in outlook and was my second major exposure to design criteria. In the late 1960s I had been involved in the revision of the flying quality spec to what became, MIL-F-8785B, "Flying Qualities of Piloted Airplanes." I was the AFFTC's representative for the revision to 8785. That was early in my career; I didn't question basic design criteria. Participation in the spec revision gave me a better appreciation for the technical underpinnings, if you will, of spec criteria. In the case of the flying qualities specification, much of it had been developed by what is now Calspan, formerly Cornell Aeronautical Labs. Bob Harper and Chick Chalk were two of the principal Cornell participants in that effort. That was the first time I really thought about test philosophy and adequacy of test criteria. We began to ask ourselves some very basic questions.

H: That seems comparable to the old problem of inertial coupling early on in the 1950s. The fifties were a period for development of stability augmentation systems. It seems the 1960s were a time when—and the F-111 is a good case study—we learned to appreciate the problems of things like engine inlet matching with the airframe and the early problems of high AOA, which would become much more important in looking at the next generation of fighter aircraft. We discovered that our investment in air-to-air technology and the kinds of aircraft that we had thought we needed were not all that appropriate for the war in Southeast Asia. Even before the war in Vietnam, the Air Force recognized this by having to buy large quantities of the Navy-developed F-4.

*Now known as the Aeronautical Systems Center (ASC). ASD and its predecessors at WPAFB have been the primary aircraft development organization for the U.S. Air Force since 1947 and for the U.S. Army from 1917 to 1947.

The late 1960s were a time in which there was a great deal of concern in Air Force acquisition on what the next generation of fighters should be. That, of course, resulted eventually in the "fighter mafia" who spawned the F-15, and then a splinter group that triggered what became the F-XX, which emerges as the F-16 thanks to the lightweight fighter competition. From your perspective at Edwards in flight testing, was there any pressure coming from the flight-test community that one would say resulted in this changing climate of thought toward much more agile, much more superior performing air-to-air airplanes typified by the F-15?

A: Yes, and there was a tremendous amount of dialogue resulting from the F-4/F-111 experience. There were ongoing discussions in those days between the flight-test community at Edwards and the engineering community at the Aeronautical Systems Division. This dialogue resulted in a much better understanding of design criteria and test criteria. Earlier in my career I never thought about the adequacy of design criteria and the connectivity of design and test criteria. This remains a major issue even today, which I will discuss next. When I was involved in the F-105 program, we tested against what was the flying qualities specification at the time. I had assumed that was a test spec. It wasn't until I got involved in the update to 8785B that I really appreciated that the people in the engineering community at Wright-Patterson and Cornell Aeronautical Labs and, indeed, the designers viewed that document as a design spec. That is what they were; they were really primarily design specifications with some testable measures of performance incorporated.

This was one of the areas of discussion when we were updating 8785B. I believed that it was important to make certain that design criteria were testable. The requirements needed to be stated not only in terms that were meaningful to designers but contain testable measures as well. I am drifting off on a tangent here, but I want to make a point. It is the same issue we are talking about today in terms of overall user requirements definition and how you translate a using command's requirements into, first, design criteria that can be distilled and put in a spec and then later translated into some measure of effectiveness that can be used for test purposes. That same measure of effectiveness needs to be used in cost and operational effectiveness analysis. I think we have come full circle on this issue.

H: It was almost a glimmering of what would be.

A: Yes, and it's the same issue today. The manufacturer still needs flying qualities design criteria, propulsion design criteria, etc. Those are system level performance criteria, but we also need macro measures of effectiveness to evaluate improved capability in terms of weapons effectiveness. Quantifying improvements in weapons effectiveness is very scenario-dependent.

User Requirements and Operational Testing

H: It always seems that there has been tension between the test community and the user commands. At one extreme, the user command may feel that it should determine the test criteria of what the aircraft is designed to do. On the other extreme, the test community could be accused of saying basically, "here's the plane that we think you want." In your experience with the F-105 earlier and then the F-111, how did you find your relationship between the test community at Edwards and TAC [Tactical Air Command], the using command? The reason I say that is, with the F-100 experience in the 1950s, it hadn't been all that good.

A: I think that's a good question. I wasn't aware of any dialogue in the early days between the DT community and the using command. Operational testing (OT) of the F-105 was going on at Eglin AFB, Florida; at that time Eglin was the Air Proving Ground Command.* We had no dialogue with the using command in developing our test plan. During the F-111 program, we had using command representation almost from the outset. I don't think we called it a combined test force initially, but it was the first time we had using command representation in the test planning process from the beginning. We had TAC representation, Col. Henry Brown and Herb Brightwell, who was the first fatality in an F-111 accident, a landing accident—a very tragic accident. Herb was assigned to the test force.

The government test process had evolved to where we had using command representation. There was, even on the F-111, dialogue between developmental testers and the the using command. I later moved over to systems test. There was probably more dialogue between the system test organization and the users. With some exceptions, for the most part we still do not have enough operationally oriented testing built into our DT programs; we are still too stovepiped. We have combined test forces—the B-2 and C-17 and

*The Army Air Forces created the Air Proving Ground at Eglin Field in 1941 and upgraded it to the Air Proving Ground Command (APGC) in 1942. After a brief hiatus following World War II, the APGC was recreated as a major command of the new USAF in 1948. Reflecting a deemphasis on conventional combat and operational testing after the Korean War, as well as the drive for concurrency in weapon system acquisition, the APGC was disestablished in 1957.

so forth—but we still need to do a better job of operationally oriented testing earlier in the process. That remains one of the big areas of potential improvement. We also need to work more closely with the requirements community. AFOTEC [Air Force Operational Test and Evaluation Center], created as AFTEC, is supposed to be a surrogate for the user. There is now an Air Force, Army, and Navy agency that is a surrogate for the user, but they may not always represent the user adequately.*

H: Because they have moved beyond it?

A: Because they have, and the user's needs change. We have combined the test functions, but we still need to maintain close and ongoing contact with the actual user community.

H: The Navy seems to rotate their flight-test personnel back into the fleet and then back into the RDT&E side of the house. Does that seem to be a better model?

A: It's a different model; I don't know that it's better or worse. On the down side, the Navy has not historically integrated their testing to the extent that the Army and Air Force do. One criticism is that they tend to repeat tests. I think the key is allowing people to work together and still maintain some degree of independence and autonomy in reporting. That's an area that we need to continue to improve upon. The operational test community in the services is being driven and driven more heavily by the Director of the Operational Test and Evaluation Office in OSD.** The operational test community in the services should not be more concerned about being responsive to DOT&E than they are about being a surrogate for the user. Organizations that were set up to serve the user's interest should not be subjected to some of the pressures they currently have to deal with. My comment here applies to DOT&E and congressional pressures which I will discuss later as well as the service operational test agencies.

*In 1970 President Nixon's Blue Ribbon Defense Panel, headed by Gilbert W. Fitzhugh, recommended that each service have an independent operational test agency (OTA). The Navy's Operational Test and Evaluation Force (OPTEVFOR), which had existed in one form or another since World War II, became DOD's first independent OTA in 1971. One year later the Army created the Operational Test and Evaluation Agency (OTEA), which was expanded into the Operational Test and Evaluation Command (OPTEC) in 1990. Reflecting entrenched opposition, especially within Systems Command, the Air Force did not create an independent OTA until 1974 when it formed the Air Force Test and Evaluation Center (AFTEC)—renamed AFOTEC in 1983. The Marine Corps Operational Test and Evaluation Activity (MCOTEA) was created in 1978. See Lawrence R. Benson, *History of Air Force Operational Test and Evaluation (OT&E)* (Kirtland AFB, N. Mex.: AFOTEC, December 1992).

**Although creation of a separate OSD office to oversee operational testing had been debated for a decade, the DOT&E function was not created until Congress mandated it by law in September 1983.

H: They are serving the needs of the process but not serving the needs of the customer?

A: Yes. We have to emphasize a continuing, ongoing connectivity to the user in the requirements process and as the requirements evolve, the test program evolves in a corresponding manner. We need to keep an operationally oriented focus in our testing.

H: Before we get into the specific cases that you may wish to address, would you like to talk generally about the changes in the acquisition process the DOD began to implement in the late 1960s; for example, the establishment of the DSARC process—the Defense Systems Acquisition Review Council—and the call for IOT&E [initial operational test and evaluation] by independent testing agencies?*

A: Let me address the issue of IOT&E, and then we'll discuss prototyping and fly before buy. The initial reaction at the AFFTC to the creation of what became AFTEC [Air Force Test and Evaluation Center] was the question: Is there a need for another separate test organization? The issue had to do with the developmental test world believing that the job it was performing was adequate. In retrospect, what we were doing, at least in terms of addressing what would come to be known as critical operational issues, was inadequate. From our perspective at the time, (a) we questioned the need for the formation of AFTEC and (b) there were later tremendous turf issues. I happened to be involved in it because I was working the engineering aspect of the issue, which was what should be defined as developmental testing and what should be defined as operational testing.

Again, I am characterizing it from the perspective I had at the time, which was that AFTEC was trying to carve out a role for itself which overlapped into areas that were developmental. In retrospect I think the role that they carved out was proper and, perhaps, looking at the benefit of twenty years of watching operational testing evolve, didn't go far enough. The whole idea was that we would create an organization that was the user surrogate, if you will, that represented the operational community. At the outset AFTEC drew very heavily on staffing provided by TAC and SAC [Strategic Air Command] at the time, depending on the aircraft. In a sense, it was a "basket command" with matrixed resources. There were a lot of discussions about turf. We had just begun to define what constituted operational testing. From the outset we jointly developed the test plan and the AFFTC Commander, the Program Office, and AFTEC Commander jointly coordinated on the government test

*Memoranda from Under Secretary of Defense Packard established the DSARC process in May 1969 and defined IOT&E in April 1971.

plan. So the test plan was worked in concert, although in retrospect, the DT and OT parts of the plan were, for the most part, two distinct plans bound in the same document.

We need to become more operationally oriented in our early testing. The compartmentation between developmental and operational testing, which has been exacerbated by congressional legislation,* has forced the operational community to move in the direction of less dialogue and less interface. I think that's wrong. I think there should be more dialogue, more interface with the operational community in planning even the early tests. Obviously, if you are going to accelerate the process and cut down on the cost of full-scale testing, you need to use models and simulation more heavily. The operational test community has to play a dominant role in model and simulation validation. In the final analysis, those tools have to be acceptable to that community in order to significantly cut down on the amount of platform testing.

Combined Testing with Contractors

A: Let me address another turf issue that evolved. This had to do with the issue of combined testing and the role of contractor versus the government. The government, for a variety of reasons, began to play a larger and larger role in the early development, although the Air Force had a different philosophy than the Navy. If you back up to the late fifties and sixties, the Air Force conducted what were called Air Force preliminary evaluations and the Navy performed Navy preliminary evaluations. The services would accomplish an evaluation at selected points in the development cycle. The Air Force then migrated from the old phase system to the category system and two things happened. One, because of encroachment pressures at aircraft manufacturing facilities, the contractors migrated to Edwards for their development testing. In the earlier days, McDonnell used to test out of Lambert Field at Saint Louis and General Dynamics out of Carswell AFB, Fort Worth, Texas, for the early B–58 work.

H: And Republic out of Farmingdale [New York].

A: Yes, and Republic out of Farmingdale. North American moved its F–100 testing to Palmdale. There was a geographic migration from these contractor facilities to "Contractor's Row" at Edwards, and which resulted in a collocation of contractor and government test

*For example, in 1986 Congress legislated strict prohibitions against any contractor involved in the development of weapon systems providing support or data for OT&E.

activities. There were some contractor versus government turf issues in the developmental process which I will discuss later. The Navy also migrated to single-site testing at Pax River, about the same time, for many of the same reasons. The Navy, however, did not move in the direction of consolidated testing to the extent the Air Force did.

I am convinced that the combined test concept as it evolved had merit and continues to have merit. There are some problems, however. Having seen it from a different perspective over the last decade, my belief is that there is a danger in the government being too heavily involved in executing the developmental process. The role of the government in the developmental process is fundamentally one of oversight and evaluation. The government doesn't develop products; the contractor is paid to do so. The contractor has to have the latitude to do that without undue interference and micromanagement on the part of the government. When you become too heavily involved in the developmental process, you lose sight of your role as an evaluator. I think that has happened to some extent today.

The other thing that happens is the contractor is forced to rely too heavily on governmental infrastructure support and, therefore, loses some flexibility. In addition the pace of development can be slowed because of the availability of government support. Contrast a typical commercial aircraft flight-test program with a military program. I believe the government's role in the safety review area is appropriate, but I also believe there is often micromanagement in the technical review process. The contractor must have the flexibility he needs to execute the development.

H: At that point the government almost becomes a policeman, if you will. Rather than a person trying to get something accomplished, it becomes a monitor of contractor performance.

A: Absolutely, particularly in the airplane test world. I'm not going to cite specific programs, but I think in some instances in recent years the government has gone too far. The contractor needs flexibility and some autonomy early in the developmental process. The combined test concept—having a single database and a single plan which all the parties agree to—has merit. The government developmental community, the government operational community, and the contractor all analyze the data and sometimes reach different conclusions from the same database. That fundamental premise is good, but early in the developmental process the contractor needs to have autonomy. There are a lot of historic reasons, but if you look at what has happened in some of the European countries, basically you have a . . . I'm groping for the word.

H: Incestuous?

A: Where the government micromanages or runs the industry, I think you lose initiative and flexibility.

H: It's almost a socialistic system.

A: Yes, it's almost socialistic. I am getting off on a tangent here, but there is a role that government needs to play in the development process and that is serving as a catalyst for technology.

H: As a catalyst but not an inhibitor?

The Government's Role in Developing Technology

A: Yes, a catalyst but not an inhibitor. The old NACA [National Advisory Committee for Aeronautics] exemplified the proper role. You have some high-value facilities that are made available to industry and academia as well as to government researchers.

H: In fact, it is funny that you say that because one can contrast the history of the NACA, which was highly successful, with the history of NASA, which has had more than its share of problems.

A: Right now, the DOD and NASA are involved, along with the Department of Commerce, the National Science Foundation, and the Department of Energy in a national aeronautical and space facility study. This study was initiated by NASA to look at what we need to do to stay competitive in the commercial arena in aeronautics and space, a worthwhile initiative.

H: That sounds a little bit too popular.

A: That's right. The two highest priority areas are low-speed and transonic wind tunnels to support research into the next-generation commercial airplanes. We don't have low-speed facilities in the United States (U.S.) that are as cost effective and have the Reynolds number* capability of the facilities that exist in Europe. The same thing is true in the transonic area. The transonic facilities have military as well as commercial applications, but much of the application, at least in the future, is commercial, improving cruise performance at transonic Mach numbers. This is potential a two- to three-

*The Reynolds number is a coefficient used to measure the dynamics of a fluid flow, or in the case of testing scale models in wind tunnels, the density and speed of the air flow needed to emulate actual atmospheric conditions for full-scale vehicles.

billion-dollar program, a national program, where the government would invest in these facilities, both for military and commercial use. At least, that's the proposal.

A lot of people are involved, but NASA and DOD are the two big players. One of the most immediate issues is a real need to develop some new facilities. Two that have bubbled up as the highest priority are a low-speed wind tunnel and a transonic tunnel. The driver for both of these facilities is primarily, although not exclusively, commercial, but the low-speed requirement is almost exclusively commercial. Again, it's the issue of not only that capability of the tunnels but being able to put models in and out relatively rapidly, the efficiency issue again. I think this is a good example. In the past, neither NASA nor DOD focused on efficiency of operation at the outset of test or research facility design. That's one thing we have begun to focus on in the wind tunnels. With the recent major range and test facility infrastructure upgrades, we have considered efficiency of operation. I don't know how successful we'll be. We need to develop metrics, some measure of cost effectiveness. How do you turn more data around less expensively, more rapidly? People need to start thinking in those terms at the outset of a new facility development.*

H: Given how we played such a major role in shaping the whole transonic and supersonic revolution at mid-century, it's really an astonishing comment that you have just made. When you think that the United States really implemented this revolution, and now we're the ones who have a shortfall in transonic facilities. You just wouldn't expect that to be the case.

A: It really is, and I think it's a manifestation of a couple of things. I don't want to politicize it, but I have seen the last eight years of a Republican administration and the transition to a Democratic one and there are differing views. Should we invest in technology, and how do you pick the winning technologies? That's a very difficult and a very complex issue. Basically, the last Republican administration's view was that DARPA [the Defense Advanced Research Projects Agency—now the Advanced Research Projects Agency or ARPA], as an example, ought not to invest in dual-use or commercial technologies, and the Democratic view is the opposite. In areas such as high capital investment research facilities, I think it is proper for the government to make the capital investment in developing the facilities.** The facilities are then made available to commercial

*For more on test facilities, see appendices C and H.

**See appendix F for Mr. Adolph's statement on domestic technology transfer to a subcommittee of the House Committee on Energy and Commerce.

companies and academia as well as government researchers and the users pay the direct cost of operations. As I said, there are aeronautical facilities in Europe which exceed the capabilities of our facilities in two respects: one, in terms of Reynolds numbers and, two, in terms of productivity. Those two issues, more capable and cost-effective facilities, are part of this study.

H: Was any of this related to the notion that people had in the early eighties, which seems to have cooled somewhat, that we could get away from actual physical testing in a tunnel and do a lot of computerized matching with computational fluid dynamics and things of that sort?

A: Yes, and I think that's part of the problem. I think CFD—computational fluid dynamics—has been overrated; you must have a balanced approach. It's true in aerodynamics; it's true in propulsion; and it's true in any discipline. You can do so much with basically what I would characterize as physical models. We get a more powerful number-crunching capability every couple or three years, so we have been able to do more and more, but there are limits on what you can do with a model. Some of the phenomenology isn't totally understood, particularly when you get into the transonic region or the high angle of attack flight regimes and things become highly nonlinear. You can't totally characterize the phenomenology; that's one dimension of the problem. There is certainly a place for models, and there is a place for simulators. Wind tunnels could be characterized as simulators. As these tools become more powerful, you have to do less and less in flight test. We have seen the positive results of that evolution, going back to the number of aircraft that we lost in the fifties, in a typical flight-test program. We do much more of the basic and applied research in aerodynamics and propulsion today, using both analytical models and wind tunnels, than we did in the past. It is a suite of tools that has to be used in concert.

When NASA moved into the space business, they didn't pay enough attention to aerodynamics and turbine engine propulsion; they did not take a balanced approach. The second letter in NASA stands for aeronautics.

H: Something like only 5 percent [of the NASA budget] was going into aeronautics.

How about taking a look at DARPA—now ARPA—in light of the end of the cold war trying to shift its focus more toward civilian needs but still certainly as a major player in military affairs. DARPA had sponsored development of a number of technology demonstrators and technology evaluation programs. I'm thinking in particular of, say, "Have Blue" [the experimental scale model for what became the F–117 stealth fighter-bomber]. Looking at that, you have, if you will, something almost like a techno-RAND. In much the same way that RAND* serves as a policy analysis and idea organization, ARPA in many ways serves a more hard-edged R&D technology purpose. Do you see an expanding role there, or do you see that as something that poses more problems than solutions?

A: What happened, if we go back to the seventies, is DARPA filled the void caused by NASA migrating to the space business. DARPA filled it with Have Blue and the X–29. I think there is a niche there for them. They are currently taking the lead in the ASTOVL [advanced short takeoff and vertical landing] area. They have also recently sponsored an "affordable aircraft" study. That's currently under way. I would like to see them continue to play the role they are playing. I think there's a real need, whether it leads to actual flying platforms or terminates with concept studies.

H: We were talking about ARPA in the present day and the function of organizations such as ARPA in the test process. One other question that came up even in the 1970s was—because of the increasing cost and complexity of test—is there a value in pursuing unmanned test systems and methodologies rather than just manned ones? The program that I think people placed a lot of hopes on and that was basically a disappointment was the HIMAT effort [highly maneuverable advanced technology] demonstrator, which NASA had back in the late seventies and into the early eighties. Do you have any thoughts on the manned versus the unmanned role in flight test?

A: I guess I would characterize it in terms of cost-benefit from having an aviator on board. A classic example is in the hypersonic area. There is a lot of research that can be accomplished in the hypersonic area that certainly doesn't require a manned vehicle. There is a high price you pay for a suborbital or a single-stage to orbit to have a man on board in terms of the payload, safety issues, and so forth. The initial research can and should be done on an unmanned vehicle. HIMAT is an example of where people tried to do things relatively inexpensively. It turned out to be very protracted, and I'm not certain of the payoff.

Another example is the model F–15 spin effort. The manned F–15 spin program was completed before the unmanned program was

*The government-sponsored nonprofit research corporation in Santa Monica, California, established as part of Douglas Aircraft in 1945 and chartered as an Air Force think tank in 1948.

well under way. The basic assumption in the manned versus unmanned is the unmanned is less expensive, can be done more expeditiously, and will provide technological insights for a later manned program. That may or may not be the case. In the area of hypersonics, you pay a tremendous price in terms of weight, payload, and safety-related issues to explore that regime by putting a man on board. It makes sense to explore this area with an unmanned platform. In most other areas, the value added by having the flexibility afforded by an individual is worth the cost, if you will, of developing a manned platform.

Again, the two examples I have cited, the HIMAT and the F–15 spin platform, did not live up to expectations because of the issues involved with supporting (a) an unmanned platform and (b) scaling effects. If you are going to use the big wind tunnel in the sky, scale it up. As an example, Have Blue was not full scale but close to it.

H: Getting back to the wind tunnel study we discussed previously, in a much cruder form there was in the late 1940s and early 1950s the Southern California Cooperative Wind Tunnel Program designed in many ways to do exactly that. And I think it worked fairly well for its time.

A: Yes, it did.

H: Who would run these new tunnels? Would these be considered NASA facilities or would they be national facilities?

A: That decision hasn't been made yet. It could be GOCO [government-owned, contractor-operated]. It will probably be either NASA or GOCO, but it may be located at AEDC [Arnold Engineering Development Center, Tennessee] because it may be the most cost-effective location, both in terms of construction costs and operating costs—relatively inexpensive electrical power.

H: TVA [Tennessee Valley Authority] power.

A: Yes. That's one dimension. Another dimension—it's in the national interest to have these kinds of facilities. The Boeings and Pratts of the world aren't in a position to step up to funding them. Right now there are facilities in Europe that are more capable, and U.S. firms are using them to support aircraft development. We need to step up to this as a nation.

H: It's a national industrial security issue, if you will.

A: Yes, a national industrial security issue; I think people are coming to that realization.

H: It makes more sense to do this than something like the superconducting super collider.

A: That's right; you had better believe it.

Evolving Challenges in Flight Testing

H: As I mentioned at the beginning of the interview, the 1970s were a particularly rich time for flight testing because it really did mark the reemergence of the primacy of flight test in aircraft development. How did you perceive this from your perspective at Edwards, first in the engineering community and then increasingly in the test resource management and range management framework? Did you recognize that you were in the beginning, if you will, of a renaissance of test thought?

A: Very definitely; the pendulum had swung hard over from total package procurement, characterized in the Air Force by the C–5 and the F–111 in the McNamara era, to an emphasis on prototyping, fly before buy. This was epitomized by the lightweight fighter, the A–X, and the AMST. Later there was a competitive fly off between Boeing and General Dynamics for the Air Force air launched cruise missile. We had all those prototypes or competitive fly offs, and we looked at a lot of new technologies in a hurry.

Another dimension was the way those programs were structured. Each one was structured differently to some extent, but there was a common thread, and there was a very aggressive flight-test program. Aggressive but safe, because there were no airplanes lost in the accelerated flight-test program. The contractor had the latitude to go out and develop things and then the government came in and evaluated. The lightweight fighter had mixed crews all along, whereas in the A–X, as I recall, the contractor was totally responsible for the development and envelope expansion and there was a three-month evaluation period at the end. The common thread of all of those, including the advanced medium STOL transport, was a very aggressive program, where basically a complete flight envelope was opened, from an aerodynamics, propulsion, and even from a limited weapons systems demonstration point of view, in a period of roughly a year.

You contrast that to the very laborious, protracted development cycles that programs have been involved in since that time. There is a lesson to be learned, and we can do things rapidly and safely and relatively inexpensively if we structure the program in an appropriate way. In today's downsizing environment we have to go back to that. Some people would argue that with today's analytical tools and computational fluid dynamics and improved models that you no longer need to build prototypes. If you can totally define the requirement, there is probably less need to prototype an airframe than there was in the past. I believe we are in a mode now where we may not be able to totally define the requirement and don't want to lock ourselves into a 1995 or a year 2000 technology, but rather continue to move technology. The way to do that is to build prototypes and demonstrators.

H: There were several major revolutions in aerospace history. In the 1920s and 1930s we had revolutions in structures and in propulsion technology and a generalized aerodynamic revolution that we called the streamlining revolution. When people think of the 1920s and 1930s, they think of the transition from the wooden airplane to the all-metal airplane. Then, of course, at mid-century we had two revolutions: the turbojet revolution and the supersonic breakthrough.

In the 1960s onward, it seems that there were two major revolutions. One was a revolution that proceeded at a relatively slow pace, but we are seeing the fruits of it now, and that was the revolution, once again, in structures, with composite and advanced materials technology. The other revolution, which had started to some degree at the end of the 1930s with the development of flight control instrumentation, was the electronic revolution. Increasingly in the 1950s, you started seeing the emergence of what were rudimentary systems airplanes, where you had very complex command and control systems and primitive radar systems. In the 1960s we saw that evolve, not merely in terms of weapons systems capability; we also entered an electronic flight control revolution. How did you see this impacting on the flight-test process, and how did you see the role of flight testing and demonstrators as the 1970s went on?

A: I will discuss that not only in terms of flight testing but in the area of requirements definition as well. As you mentioned, if you look back at aviation, when I first became involved in the late fifties, higher and faster were in vogue. The F-102 and F-100 test programs explored supersonic level flight capability in production aircraft. The next logical step, it appeared at the time, was the ability to go to Mach 2 and that was a requirement of the F-104, the F-105, the F-106, the F-107, and the F-4. People then realized that speed in and of itself did not provide the capability needed. The next generation emphasized what was called at the time energy maneuverability, or excess thrust—lift and thrust-limited turning capability.

These were design criteria for both the F-15 and the F-16, as well as benign high AOA handling characteristics.

In terms of airplane design, people came to the realization that a high Mach number was not all that important. Offensively there was a need for increased lift and thrust-limited maneuvering capability. Defensively, as the enemy's acquisition and track radars became more sophisticated, aircraft needed to fly low to avoid detection. You saw that happen right in the middle of the B-1 program.

H: The high-altitude—low-altitude profile.

A: Yes, the B-1 originally had a sophisticated inlet design for supersonic flight. Designers then started looking at signatures, initially RF [radio frequency] but later the whole gamut of signatures. Now the questions: Where do we go from here? What's the right mix? How much stealthiness is enough?

In terms of defensive weapons, first the very sophisticated surface-to-air missiles posed a threat. A significant recent change in weaponeering is the availability of the very lethal man-portable weapons.

H: Very high energy weapons.

A: High energy weapons, Stinger-like weapons. So a fixed-wing aircraft, in particular one moving in a relatively predictable trajectory is at risk. A $50,000 weapon can put a $50 million airplane at risk.

H: In effect the weapon now becomes a maneuvering air combat vehicle. It is much more important for the weapon to have that ability, one that takes high-G maneuvering on a target, than it is necessarily for the carrier of the weapon.

Did you have the feeling as the seventies were going on that you would start to see these complexities in warfare and that these would put special burdens or challenges upon the testers? Or did you suddenly find yourself in the situation where, almost like a storm, you were suddenly in the midst of all these conflicting requirements, changes in technology, and changes in weapons effectiveness—and have to struggle to keep up?

A: It was the latter. I really didn't appreciate this until I came to the Pentagon. The people who are executing the day-to-day developmental testing out in the field don't need those insights. What is missing is a cadre of people who have insights into the interaction of requirements, testing, tactics, and doctrine. We now have some of the analytical tools—models and simulations so people can look at these issues. At some level within the services, the people who structure tests, particularly operational tests, need that kind of visibility and insight into requirements issues, tactics, and doctrine.

In my view, this is a significant void. Some of the operational test plans were written without those insights. I'm not being critical of the operational test community. I am probably being more critical of the requirements community and the requirements definition process. Requirements don't evolve and operational test plans don't change as a part of that requirements evolution. There has to be very close connectivity between operational testing and requirements definition. The test measures of effectiveness and suitability have to be updated, but you have to make sure that they are really consistent with a real-world requirement. With what has happened in recent years, all the systems that are out there today were designed against a Soviet/Warsaw Pact threat, so our simulations, our requirements, the threat assessments that come out of DIA [Defense Intelligence Agency]. So everything has to be rebaselined.

H: In many ways we were developing systems to confront an SA–2/SA–5/SA–6 environment and an air defense environment that was based on advanced MiG–23/MiG–29/Su–27 type threats, and you find yourself now possibly fighting a very different kind of war, for that matter, with blue threats.

A: Blue threats, gray threats, and things that are produced anywhere in the world and lashed up and kluged together in an infinite number of ways.*

H: For example, looking at the Iraqi air defense network, the so-called Kari system, which had French components, British components, and Soviet components all kluged together in this mix.

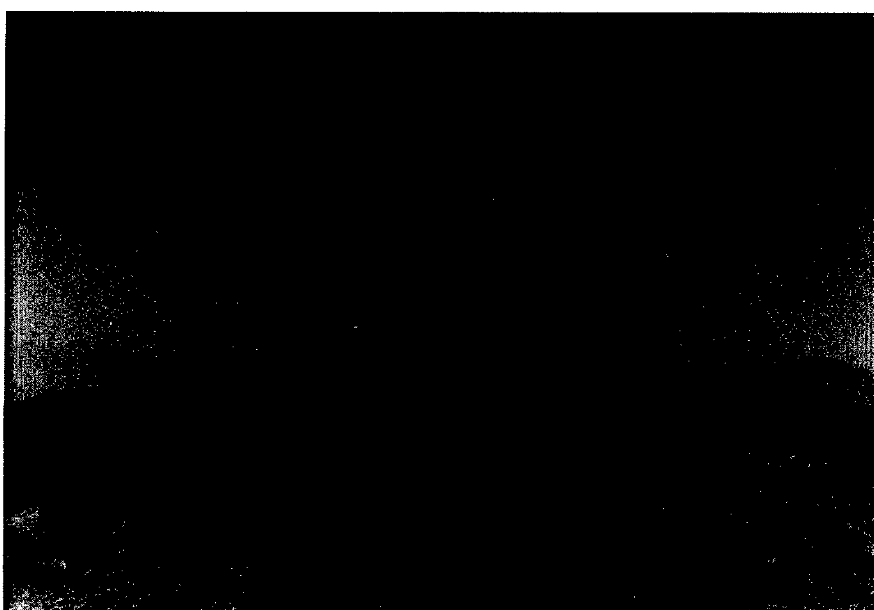
A: From a requirements point of view, scenarios may be less stressing, but they are more ambiguous. We need—for test purposes and for tactics and doctrine and certainly for warfighting purposes—to have the flexibility to put together a wide number of potential scenarios. There is a need to test and train in those environments to make sure that next-generation systems meet a wide variety of requirements.

H: I think what you've stated very well is the fact that you have to be much more concerned about the requirements process and the interplay between test and requirements when you are developing a system. Recently we have seen situations in which people have in

*In intelligence terms, red threats represented those developed by the Soviet Union or its Warsaw Pact allies, blue threats those developed by the United States or its NATO allies, and gray threats those developed by other more neutral nations.

Competitors in the A–X Program in 1972: Northrop A–9 *(top)* and Fairchild Republic A–10 *(bottom)*. Based on test results and cost estimates, the Air Force selected the A–10 as its specialized close air support aircraft.

The Advanced Medium Range Short Takeoff and Landing Transport Program: Boeing YC–14 (with HH–60) and McDonnell Douglas YC–15. Although successfully demonstrating new technologies from 1975 to 1977, the program was cancelled in favor of continued reliance on the C–130 turboprop airlifter.



some cases almost glibly stated values of performance or values of expectation for new systems without thinking through those and the implications those have for testing and the difficulty it may be to test to get that. Two examples come to mind readily: ASPJ* is definitely one and the other is the F-22, where you have the statement that the F-22 will be twice as effective as an F-15. When you get to those cases, you have very great difficulty actually finding numbers or finding measures of merit that enable you to say that. You are looking at a much more complex process.

In your experience in the seventies, as we saw this renaissance in returning to what was termed in some circles the fly-before-buy philosophy, what were the disappointments? Were there surprises? Were there unexpected things that caused you to rethink the relationship between testing and developing requirements?

A: Looking back at that time, I can't identify any real problem with requirements, but I didn't have the perspective and the visibility that I have in my current job. Certainly, the whole issue of energy maneuverability and the significance of it was not appreciated in the late fifties/early sixties. There was the push for higher and faster; then people decided that wasn't enough and that the capability to maneuver required more design attention—both thrust and lift-limited maneuvering. In retrospect there was a focus on a single issue. I guess the lesson to be learned—with the focus on speed, then energy maneuverability, then low-level operation, and then stealth—is, again back to the issue of balance. With today's focus on stealth and the issue of how much stealth is enough, the lesson to be learned is that we have to constantly keep in mind balance and not focus too heavily on one set of criteria or one requirement. That lesson was learned again in the A-12 and the TSSAM [tri-service stand-off attack missile].**

In terms of surprises, looking back on all three of the prototype test programs that, as engineering section chief, I was fairly close to—the A-X [A-9 versus A-10], the lightweight fighter [YF-16 versus YD-17], and the STOL transport [YC-14 versus YC-15]—there were really no great technical surprises. The one thing that sticks in my mind, looking back at those programs, was how smoothly the flight-test program went on all three, and the fact that it was really done in a combined-test arena. We had contractor and the government developmental and operational communities sharing a common

*Advanced self-protection jammer.

**Both cancelled because of performance and cost problems.

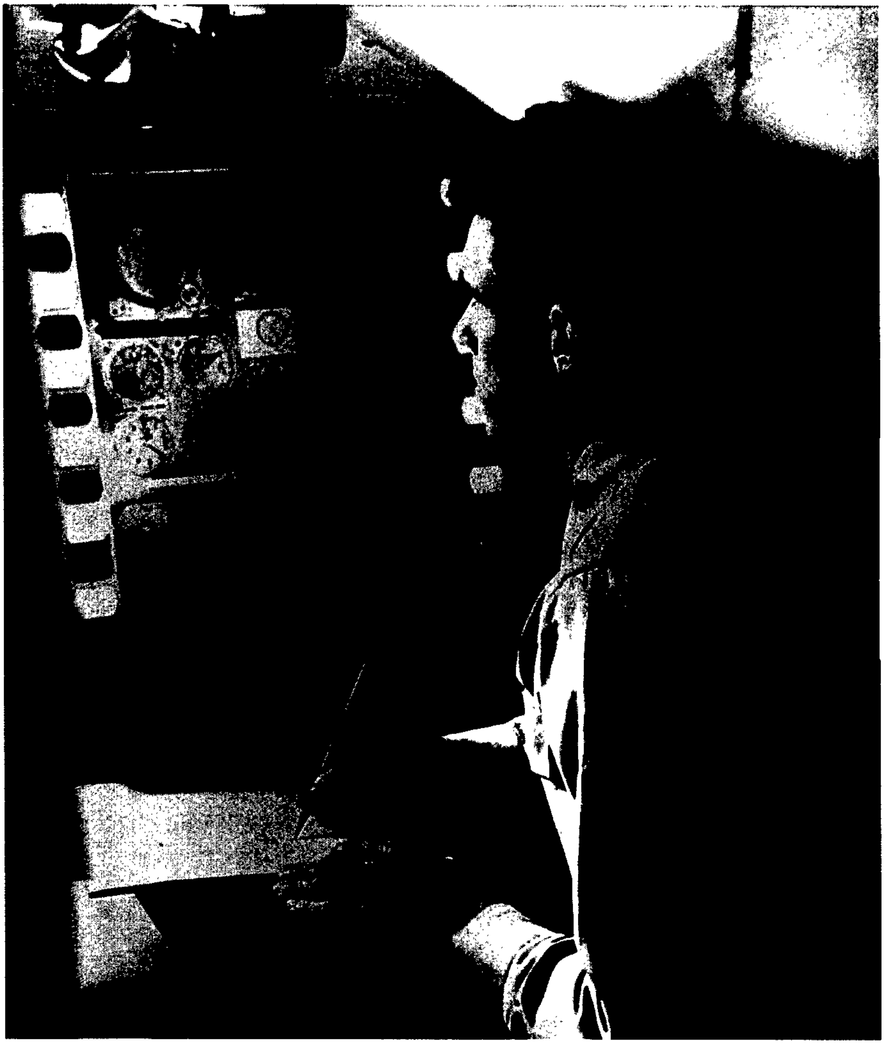
database. I think it reinforces that basic notion. It also reinforces the concept of allowing the contractor some autonomy and flexibility in planning and conducting the developmental test program. Looking back, it's the right way to look at technologies relatively rapidly and inexpensively.

The Culture of Flight Testing

H: One thing I found interesting, talking to people like ourselves who have come out of the Edwards environment, was the degree to which, as they moved to the acquisition community, for example, into what was the Aeronautical Systems Division under the old Air Force Systems Command, or then moved into the AQ [Acquisition]* community in the Pentagon, their perspective not only broadened. There was a recognition that there had been an Edwards culture that was very good but was very localized and parochial in some respects. Much to their surprise, many of these people who were very much die-hard testers—for example, test pilots, flight-test engineers, or whatever—almost against their will suddenly discovered their perspective changing. They were becoming much more sympathetic, if you will, to the environment that drove the acquisition world. Did you find yourself going through that?

A: Yes. Flight-test was almost a "religion"—I'm reluctant to use that word—but certainly a culture where testing had to be very complete, very thorough and even minor deficiencies had to be corrected. We were not sensitized to the cost of doing business or trying to do things more efficiently. Obviously, in any test business and for the airplanes in particular, you never compromise on safety. That's a given. But in other areas there is room for risk-taking. Airplane performance definition is an example. Testers like to define performance to a half of a percent, whereas a flight manual has up to 10 percent conservatism in it. I don't think we need to be as thorough in our testing as the test community would like at times. We have to decide what's important and certainly minimize technical risks and not compromise where safety is involved, but that doesn't translate into an across-the-board by-the-numbers execution of test programs as was done in the past. There has to be some balance, and

*The position of Assistant Secretary of the Air Force for Acquisition (SAF/AQ) was created in 1987 to be a service acquisition executive (SAE) reporting to the Under Secretary of Defense (Acquisition—later Acquisition and Technology) as part of the streamlining recommended by the President's Blue Ribbon Commission on Defense Management headed by former Deputy Secretary of Defense David Packard (a.k.a. the Packard Commission). Prior to then, responsibilities for managing USAF acquisition had been dispersed among the DCS for Research, Development, and Acquisition, the Assistant Secretary (Research, Development, and Logistics), and the Commander of Air Force Systems Command.



Manually transcribing test data from photo panel film using a "Recordak" in the fifties.

I think that balance is inevitable when you are operating in a budget-constrained environment.* People have to take a little different perspective than they did in the past.

Another issue is the focus on technical characteristics as opposed to operational requirements. We have to become more operationally oriented, even structuring so we can get some early

*By the mid-1990s, Air Force developmental test teams were being reduced in size and made part of SPO-supervised "integrated product teams," leaving the AFFTC in a supporting role.

visibility into operational utility. The people who spend their entire career in a developmental test culture don't really have the background to do that. I would like to see more migration of the workforce back and forth between the developmental and the operational community. I think that could happen, but it's going to take a lot of arm-twisting.

H: At the same time that you had this revolution taking place in the technology of aircraft and flight systems themselves that was giving them unprecedented ability to do certain things, for example, safe high AOA flight, we also had a revolution taking place in flight-test facilities, notably the introduction of real-time data acquisition and data analysis capabilities, typified by the development, say, of the Ridley Mission Control Center at Edwards. This at first appeared in very rudimentary form as early as the late 1940s and then had expanded somewhat in the 1950s. This led to the notion of near-real-time analysis while a flight was under way but did not really blossom until the 1970s and 1980s.

How did you see this as a tool? Did you see that it would offer the potential of reducing test time? Yet at the same time it seemed that now, since we could measure so many values, there was also the danger it could actually complicate testing. Engineers now could get answers to literally thousands of parameters as opposed to perhaps dozens or at the most hundreds.

A: There are a couple of issues there, but let me try to address them. I will bounce around a bit perhaps. In the fifties and the sixties telemetry was limited and the emphasis on the use of real-time data was for safety. Applications had to do with safety-related test areas such as flutter and high AOA testing. Real-time data was used later for what I would characterize as goodness of maneuver—displaying and manipulating data in near real time to assess the quality of the test results. The issue of safety was important, but with the ability to have engineers of every discipline looking at a strip chart and assessing in real time or near real time how well the maneuver was performed, you could go back and repeat things or fill in data voids. On some of the cargo, transport, and bomber airplanes, engineers on board reviewed the data in real time. In the commercial flight-test world, that was the way they operated routinely and continue to do so.

Regarding the inefficiency issue you mentioned, telemetry was an umbilical that tended to limit where and when you could do your testing. I was the Air Force performance and flying qualities engineer on the B–52H; we relied on onboard recording. Basically we flew all over the western United States. There were some airspace restrictions, but we were self-contained. With telemetry data acquisition, you have a scheduling problem and you are limited in the airspace you can use. That problem becomes exacerbated when you are doing low-level work because of your ability to acquire the telemetry data. The reliance on ground station support becomes a potential impediment to weekend flying at government test facilities. Basically, the more autonomous you are, the more flexible you can be.

Strides were made in the goodness of maneuver arena but only in the first-generation processing. Even in cases where telemetry is used for real-time support, most data are reprocessed later in a batch mode, which is often a very time-consuming process. We now have the technology to do things much more efficiently by processing the data to the final result in near real time, not only assessing goodness of maneuver but going the next step. The next level of sophistication is to review the originally planned tests and cut down on testing by virtue of the fact that we've acquired a certain amount of data. You do that by analyzing test results and comparing them with a model. You update that model until you get a match in a certain area. Once you've achieved a certain level of congruence between the model and the actual article, you can discontinue or cut down on the testing that you do in the future. Until we get in that mode in the use of analytical tools, we are not going to become more efficient. That's the next step. The analytical tools are there, but I'm not sure that the mind-set and the discipline is there to operate in that mode. It will take a cultural change.* Testers love to test; I know because I was one of them.

H: The 1980s also witnessed the development of a whole series of other specialized facilities. I am thinking, for example, of the kind of avionic testing that you saw emerge, things like IFAST [Integration Facility for Avionic Systems Test] and then the ultimate developments beyond that at Edwards. Would you like to comment on those?

A: With federated avionics subsystems, it became apparent that more of the integration process and more of the stimulation could be accomplished in a hardware-in-the-loop lab. You could solve many of the integration problems much less expensively and much more systematically in a ground facility where you could simulate and stimulate. Again, one problem was convincing people of the need for those kinds of tools. Another challenge was reorienting a test program to use them efficiently. With the so-called multistage

*For additional thoughts and perspectives, see appendix I.

improvement programs, which we have been in for years on both the F-15 and the F-16, these facilities really came into their own.

H: One would have not been able to do that in a previous era. We touched on this earlier, and it's worth revisiting, that there has been in many ways a cultural shift. Back in the 1950s, flight testing at Edwards was characterized by a high degree of risk and a high degree of loss, certainly in comparison to the present day. Do you think that one of the problems we see here with the interrelationship between government and industry on major research and development (R&D) projects is that failure now is almost invariably so well publicized. For example, when Doug Benefield was lost in the first B-1 accident, or when we had a loss in a program that had major national significance such as the space shuttle, or if there were the loss of a B-2 today, these events become so major in terms of their public impact. Is there a natural tendency to try to focus on safety to a point where it becomes almost burdensome to the process?

A: The short answer is yes. It is not only the loss of a platform but I think the entire developmental process has become so risk-averse that programs are stretched out unnecessarily. There's a financial

B-52H used in performance testing at Edwards in the early sixties.

analogy which probably isn't the world's greatest, but you can become so risk-averse that you invest only in CDs [certificate of deposit] and other fixed interest investments and lose the opportunity for any growth or any real expansion. I think the same thing can happen in technology. If you become so risk-averse that you don't really push technology, then you are taking the largest risk of all, and that's losing the technological edge. You have to continue to push technology; I won't say across the board, but in potentially high payoff areas. You need to do that relatively inexpensively; there is an ongoing need to build and test demonstrators, whether it's platforms, sensors, or whatever. That's the motivation behind advanced technology demos and advanced concept tech demos.

In today's world, aircraft platforms are very expensive, but there is a need to continue to get full-scale articles in a reasonably realistic environment for test purposes. You can't do everything with models and simulation. You need to build and operate the critical systems and subsystems and actually evaluate technology; it needs to be done in a way that's relatively inexpensive. The expense escalates with very protracted development processes. We have to eliminate a lot of that cost—by minimizing documentation, total subsystem quals, etc.—and evaluate tech demonstrators. In addition, more attention must be paid to development process cycle time reduction. Contrast commercial aircraft test program length with military cargo aircraft.

This is particularly critical in today's environment. If you look at the funding environment, we're not going to be fielding many new platforms, but we will continue to field upgrades. Fully instrumented platforms are needed to evaluate technologies, test-bed platforms to look at sensors. We need to pursue that vigorously as an explicit part of any advanced technology concept demonstration program. One of the problems is trying to provide something for everybody and satisfy too many people. We must rack and stack and focus on critical technologies.

To a large degree these technologies have to do with information processing and sensors. Those are two key areas that we need to continue to work very hard. Again, we need to give industry the flexibility and the wherewithal to continue to push technology. Development and acquisition oversight processes need to change as well. The advanced concept technology demonstrator (ACTD) approach is a giant step in the right direction.

H: Your test career has covered an extraordinary period in aerospace development. Certainly in that time at the various levels you've worked you must have developed your own view on how one would desirably structure a test organization. What do you think the most important structural elements are that you need within a test organization in terms of how one administers tests and relates them to other parts of a larger whole?

A: First and foremost, the developmental test organization has to be an extension of the design organization. The first word in developmental testing is "development." The test community has to be very closely coupled to the people who are responsible for the design and has to be responsive to them. I started out my career working for Convair, on the F-102 and then on the first F-106 program. We (the testers) worked with the designers when we wrote the test plans. We worked with the aerodynamics group and the propulsion group. When we developed the original test plan, there was a dialogue with them. I remember that I perceived it as an extension of the design organization.

We have now become too large, compartmented, and bureaucratic in the last fifteen years. I was involved with the YF-16 and the YF-17 lightweight fighter fly off as the government engineering section chief. In those instances, we spent a lot of time at Hawthorne [Northrop location in California] and Fort Worth [General Dynamics] working with the design engineers, the people who were working flight controls and propulsion, making sure that we were structuring a program that really provided them with the information they needed to validate predictions and develop the system. What we have today are increasingly large bureaucratic stovepipes between the contractor, the government development community, and the government operational testers. I'm convinced we've moved too far in the direction of bureaucratic stovepipes and redundancy. It unnecessarily protracts development.

The first element to change is to revert to very close coupling between design and test. The second element is to plan from the outset to facilitate testing to really be supportive of not only rapid design but design evolution. That degree of facilitation varies with the type and the magnitude of the program. If you look at an F-22 program, that's one end of the spectrum. An F-22 is going to be in flight]test for the next thirty years. The F-16 has been in flight test since the early seventies.

H: About twenty years.

A: Yes, twenty years and will continue another ten. The F-22 will be in flight test for the next thirty years, so we ought to facilitate with that in mind. There will be numerous avionics upgrades and multistage improvement programs. I am no longer hard over about the need for the ground test facility at the government test location. We need to decide where the right place is at the outset but make sure that we facilitate it in a way that supports the design community, the test community, and the service engineering (upgrade) community.

We also need to work harder at portability and transportability. Physical transportability is one dimension. In an era of distributed interactive simulation and distributed data processing, there are additional opportunities. I've said many times that I think the government test ranges have been too parochial. As an example, everybody else in the world shares computational capability. There are many examples where ranges had a saturation problem, data backlog, data turnaround. We had both at Edwards during the early days of the cruise missile test program. There was a large data backlog, both Boeing and GD [General Dynamics] data. The government had forced the contractors to do the processing at Edwards and the government couldn't process the data in a timely manner. Contract this support out and plan for surges.

H: You almost need a translator in between.

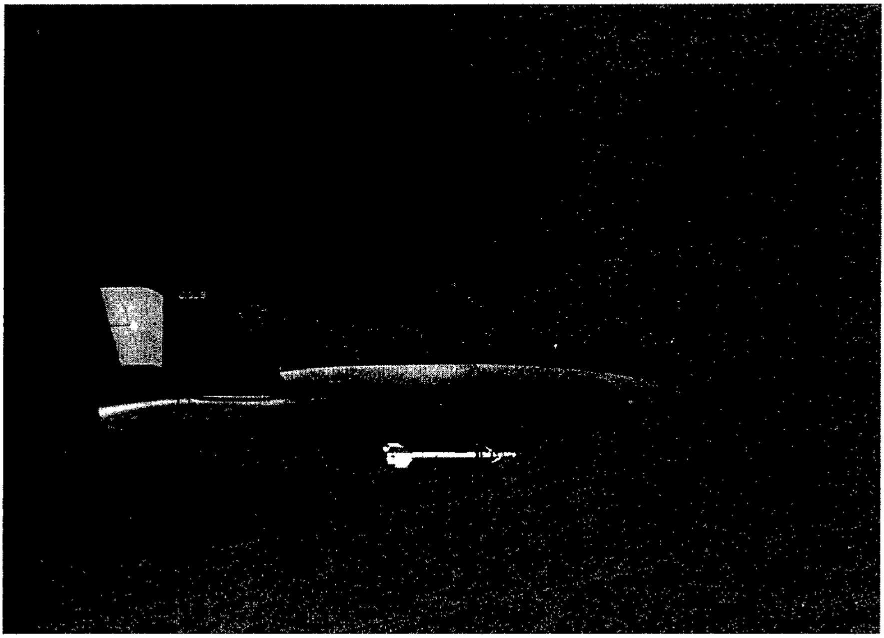
A: We must work very hard at being interoperable.

H: It seems in some cases, if we look at the acquisition process, that we take people and make them program managers who are extremely gifted technologists or in some cases practitioners, for example, test pilots, and make them acquisition officers and then SPO directors. If we are going to effect this change in corporate culture, it would seem that we would need to effect changes within the way we train our acquisition personnel. Also, would you emphasize developing test-oriented courses within the acquisition process?

A: The Defense Acquisition Work Force Improvement Act has been addressing that issue. There are requirements for program manager courses. The Defense Systems Management College [Fort Belvoir, Virginia] is already doing that. There were a number of panels set up, one which addresses systems and another test. I chaired the test panel. For every level within the acquisition workforce, there are mandatory education and training requirements.

H: Test courses?

A: Yes. There are both course and experience requirements to move from one level to another. The issue is being addressed. The concern that I have—and we are just in the early years of it now—is getting too bureaucratic and creating a lot of stovepipes and wickets. I want people to be able to move, as you mentioned, back and forth



In the Lightweight Fighter Competition of 1974, the Northrop YF–17 *(top)* lost out to the General Dynamics YF–16 but served as the basis for the Navy's McDonnell Douglas F/A–18. The F–16A *(bottom)* began flight test in 1975.

between test and systems management. People should be able to move in and out of program offices.

H: So you don't want to perpetuate a priesthood.

A: That issue is getting a lot of attention.

H: That really is a cultural shift.

A: Yes. By the way, it's a larger cultural shift for the Navy and the Army and Navy than it is for the Air Force. The Air Force historically has allowed military personnel to spend their entire careers in acquisition. You can argue the pros and cons of the Air Force approach. The people in the Army and in the Navy move back and forth between operational units and the acquisition business. The Defense Acquisition Work Force Improvement Act emphasizes relevant acquisition experience for military program managers. Part of the criticism was that the people weren't acquisition professionals. The Army and the Navy had to develop dual career tracks just like the Air Force to allow senior military to be program managers. By the way, the Army and the Navy tend to use far more civilians as program managers than the Air Force, which relies almost exclusively on military personnel as program managers and program executive officers.

Migration from Contractor to Government Test Facilities

H: As has been mentioned, in the late 1970s, we were deep in the midst of the whole force restructure issue, and programs like the F–15, F–16, and B–1 were well under way. What was the most significant change in flight test in that time period?

A: The late sixties and early seventies were characterized by the migration from contractor test facilities to government test facilities. This transition was precipitated by two things: one, technology; I will talk to that issue later in the context of integration facilities and other high-value test facilities; and the other one was encroachment at contractor test facilities. As an example, McDonnell used to conduct a significant amount of their flight testing of the F3H and the early F–4 out of Lambert Field [Saint Louis, Missouri]. KC–135 and early model B–52 test flights were flown from Boeing Field in Seattle, and Convair (General Dynamics) conducted a significant part of the B–58 test program from Carswell AFB. In the old days, Rockwell (North American at the time) flew out of LA International and Northrop was still flying out of Hawthorne [California] in the F–89 days. That flight-test activity migrated to government test facilities in the sixties and early seventies.

In aircraft design technology, we evolved from what I would characterize as a number of relatively independent hydromechanical systems to the beginnings of what I would call federated analog avionics systems. It became apparent that system integration laboratories were becoming increasingly important in the development of platforms as design tools as well as test adjuncts. Prior to that time, the one classic "integration facility" that has always been around the airplane business was the "iron bird," used to develop flight control systems and the primary and secondary hydraulic systems. The next step in the evolutionary process was the system integration laboratory. These took a variety of forms and were at a variety of locations. It became apparent that with federated subsystems that the tester and, more importantly, the designer needed access to a system integration laboratory-like test facility to make the design, development, and test process more efficient. The great debate at the time, at least on the part of the government, was the role that these facilities would play in support of test activities.

In retrospect the one thing that we testers didn't realize at the outset—but have become increasingly aware of in the last fifteen years—is that these facilities are not only test tools but are developmental tools as well. Part of the mistake that we made was that our perspective was too limited; we viewed these facilities as strictly government test tools. We ended up with a fragmented approach where we had system integration laboratories at contractor facilities, and we had a test facility at the test site. There was a third facility at the depot, the logistics center, to support configuration management and upgrades after the platform became operational for Air Force aircraft.

The Navy took a different approach, starting with the F–14. They used one system integration laboratory to support both contractor development, government development, and configuration management after the aircraft became operational. The F–14 integration facility is at Point Mugu and the F–18 at China Lake.

I think, in retrospect, you need to encourage, if not force, both the government and the contractor to use the same facilities early in the process. I believe the facility can evolve from primarily a contractor developmental tool to a test and configuration management tool if the planning is done properly. With highly federated systems, you need to plan from the outset to have an integration laboratory that is used by the integrating contractor.

H: If you have a contractor as the integrating contractor, should that be a contractor outside the scope of the rest of the program, so to speak? In other words, not the prime contractor but a separate independent contractor.

A: For aircraft and missile development programs, with rare exceptions, the contractor responsible for the platform has to be the integrating contractor. There are exceptions, but for development of a new platform, the platform manufacturer has to be the integrator. We've gone now from analog to digital systems, and we've gone from federated digital systems using the data bus to what I would call integrated systems in the F-22. You need to plan from the outset to facilitize a system integration laboratory, certainly to support the development. That was a lesson learned on the B-1* and on the C-17 as well.

Because of cost and configuration management issues, you often can't afford the luxury of having separate facilities to support development at the contractor location, support testing, and support depot activities. If you are only going to have one, the right place to do it may be at the contractor facility, which means that the government test community must have access to it. I have changed my mind about forcing it to be at the test location. There are, however, geographic considerations. If the contractor manufactures the airplane at Palmdale, you can have the facility at Edwards. If the aircraft is built in Marietta, Georgia, and tested at Edwards, location is an issue. The Navy has gone one step further; they have planned for the whole life cycle, using one facility at the Navy test support facility. The Navy characterizes these facilities as full spectrum—lab, test, and service engineering.

There are a variety of solutions, but I think the key is to plan from the outset to have a fully equipped system integration laboratory and to use it not only to support the development but test and configuration management and upgrades when the system becomes operational. If you look at what happened with JSTARS [E-8 joint surveillance target attack radar system], they have one facility at Melbourne [Florida] and they will build a separate facility at the Warner Robins Air Logistics Center [Georgia] a few hundred miles away. In retrospect, it would have been prudent to plan for one facility at the outset and to use it to support both development and operations and upgrades after the system was fielded.

*The Air Force B-1B SPO served, in effect, as the integrating contractor.

The first Rockwell B–1B is shown here releasing a B–83 bomb with an F–111 serving as a chase plane.

H: There seems an analogy one can make to the operational world too: that having the technology to do something is not enough; you must know how to use it appropriately. Andy Marshall has an example in which he says that many nations came up with tanks but only one nation before World War II really understood how to use them: Nazi Germany. The message I'm getting is that we developed a superb analytical tool, in the specific case of IFAST, and then found that the result of our initial use of it was flawed because of the integration issue that you just raised. So it is not nearly enough to develop the capabilities; one has to know how to use those properly within the development process.

In your experience, have you found that the old issue of technology push or requirements pull has played a major role in the development of our analytical tools and capabilities to assist in the acquisition process? For example, to me it seems that one would not have been able to undertake an IFAST type of approach at an earlier time. The technology to do that had to come of age.

A: Obviously, the technology showed up in terms of the subsystems and the platform, but what you really need to support that technology is a real-time simulation capability so that you can extract. You can either have models of subsystems or the actual hardware in the loop. The VAX line of computers and similar number crunchers were among the technology enablers to be able to allow that kind of real-time facility to be effective.

Importance of the Private Sector

A: I want to talk about philosophical issues for a moment. I'm concerned about the direction that we've headed in the last ten years in that we are continuing to build barriers between the government and contractors. We have some narrowly defined stovepipes relating to the government's role in the test business, developmental and operational testing. We need to make greater use of models and simulation tools to support testing. To do that, we need cultural and legislative change. There are some procedural issues that need to be addressed as well. In addition there are technology issues which have to do with the validation, verification, and accreditation process. All these issues need to be addressed in the planning process.

There is another related philosophical issue. In the final analysis, contractors are really responsible for designing, manufacturing, and developing military equipment to meet service operational requirements. With the old arsenal and lab system in the Army and Navy, there were some substantive in-house developments in the past. But the aviation community has always relied very heavily on the private sector and I think properly so. I believe you need to rely on the private sector, but when you have these very expensive government integration facilities, you have to make them available to the contractor to support the development process. We haven't worked that issue nearly enough.*

Let me make a tangential point here while I'm thinking about it. I was involved several years ago on a laboratory commission.** We had a lot of debate about the role of government laboratories. If you look at it from a historical perspective, the Army and the Navy have had arsenals and labs since the 1800s doing in-house R&D. The issue of a full-spectrum laboratory and heavy government involvement in guiding the development was a topic of debate between the Navy and the Air Force members of the group. The Air Force contracts most of it out, a result of having been a relatively new orga-

*How to share the costs for these facilities is one of the more difficult issues.

**Mr. Adolph chaired the Federal Advisory Commission on Consolidation and Conversion of Defense R&D Laboratories. For his statement to a Senate subcommitee, see appendix D, and for the executive summary of the Commission's report to the Secretary of Defense in 1991, see appendix F.

nization with rapid growth. You can make judgments about which way is the proper way, or even if there is one best way. One measure is results. That's probably the best metric. I have reached a conclusion and my philosophy is intertwined in this: You need to rely heavily on the private sector; that's where the incentive and detailed design and manufacturing expertise must reside. You need to have some in-house government expertise, but when you get into micromanaging and trying to do too much in house, it has not proven to be very effective in today's high-tech world. As a footnote, I want to make it clear that I am not criticizing government employees in general.

H: We have a historical case in the Air Force and the Navy looking at the engineering division in the 1920s at McCook Field [Dayton, Ohio] and then the Navy with the naval aircraft factory in Philadelphia. While both those organizations did some very interesting work and came up with some very interesting concepts and ideas that the private sector was able to run with very successfully, when both those centers turned to actually designing the airplanes, the aircraft that came out were in almost all cases inferior to what was coming out of the private sector. They were able to generate technology, but they were not able to generate operational systems.

A: My bias is to let the private sector take the lead. I've said this many times: the government does not design and develop systems; the government is responsible for managing it. Fundamentally the government is a customer. When we get too heavily involved in the process, I'm convinced it's to the detriment of the product. I think there have been occasions when we've done that—micromanage—both in the development business and in the test business. The government's role in development is primarily one of evaluation; we need to put our resources into evaluation. If you look at what's happened, I think the government is too involved in test execution and committing too many resources to test support. The contractor is paid to develop the product and execute the development test. We need to create an environment where the contractor can do that unimpeded. We need to have visibility, insight into the process to make sure that it is progressing, but I think we have evolved to a point where we are far too heavily involved in micromanaging development and testing.

H: We've had some notably successful examples of technology demonstrators that have pointed toward successful operational systems and some of these were quasi-operational themselves. I am thinking in particular in that case of the lightweight fighter we discussed earlier. What do you see as the most desirable role of the

The integrated facility for avionics system testing building at Edwards *(top)* and one of its simulator rooms *(bottom)*.

government in terms of supporting the development of technology demonstrators? What should be the relationship among the government and industry and the test community?

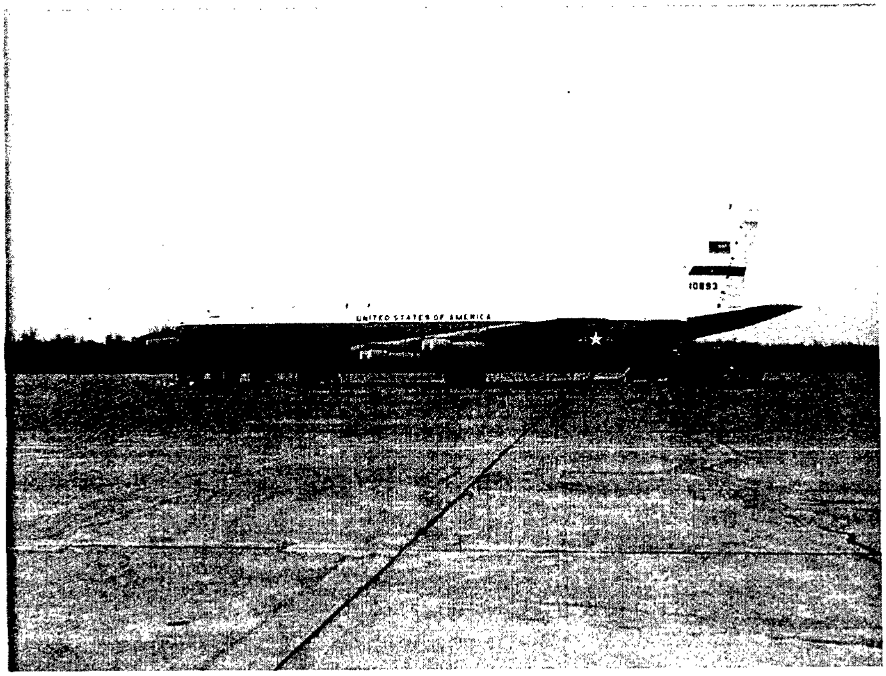
A: The government needs to let industry have more autonomy and a lot of flexibility. In the A-X program—the A-9 and A-10—the contractors were given a period of time to expand the envelope and then at the end of that period the government made an evaluation. In contrast in the lightweight fighter program, the government was involved in evaluation from the outset. In both instances, we ended up with a complete technology demonstration in a very short time period. I would move in the direction of giving the contractor more autonomy in development. I believe that was done in the case of the F-22 and F-23 DEM/VAL. The government provides the test facility but doesn't get involved in micromanaging all the developmental flights. The issue of safety has been the big hammer. Even in that area there has to be some risk-taking. The contractors realize the consequences of the risks they are taking as well as anybody. The government has to provide a lot of flexibility to the contractor, except for a period of time where they evaluate the product.

The classic Kelly Johnson approach [as head of Lockheed's Skunk Works"—now formally known as the Lockheed Advanced Development Company] was: tell me what you want and I'll show up two years later at Edwards with it. That is perhaps slightly overstated, but all of the Skunk Works programs were very lean; there was some ongoing government involvement, but it wasn't the high degree of micromanagement that exists in the major programs. In the Skunk Works programs, they brought in a product essentially on schedule that met the requirements. There were some losses in all of those programs: the U-2, the F-117, the SR-71. They took some risks and moved fairly fast through the test program and ended up losing platforms as the result, but that's the price you pay for moving technology along relatively rapidly. The loss of a platform should not place a program at risk. The negative consequence of total risk aversion is that you don't move technology along fast enough.

H: We have talked almost exclusively about government test here. We have talked about the relationship with contractors but in the sense of contractors who are major contractors like the airframe suppliers. What do you see as the role here for small privately owned test organizations? I'm thinking of Tracor, Flight Systems, and companies of that sort. Are these something we will see proliferate? And if so, is that going to be difficult for the government to relate to or to provide oversight to?

A: I don't think they will proliferate. There is a market niche; it's the relatively inexpensive support of tech demos and test beds.

H: You can even throw the 4950th* in that.



An NC–135 aircraft of the 4950th Test Wing at Wright-Patterson AFB, Ohio.

A: In my opinion, the large, expensive test beds that the 4950th operates could be provided more cost effectively by contractors. For your information, we recently commissioned an independent study of how to manage test support aircraft more effectively. I tasked them to look at the advantages and disadvantages of a GOCO operation and how to manage that test support fleet more effectively. I am personally convinced that many of the test-bed aircraft could be managed more cost effectively if it were a contractor operation. A contractor would not keep all these platforms independently and would just provide an appropriate aircraft for the duration of the test.

H: Of the program?

*The 4950th Test Wing at WPAFB, which moved to Edwards AFB in 1994 and merged with the 412th Test Wing. For a history, see *Against the Wind: 90 Years of Flight Testing in the Miami Valley* (WPAFB, Ohio: Aeronautical Systems Center, 1994).

A: Yes. So we're looking at that very issue. Tracor, the National Test Pilot School, Flight Systems—I think there is a role for them. A classic example is the Calspan effort on the variable stability T-33 and the B-26. Basically, one pilot and one maintenance person would travel around the country and provide programmatic support and training support to the Navy and the Air Force. You couldn't begin to support that type of aircraft in the government infrastructure with one or two people.

H: It was in effect a traveling road show.

A: As noted earlier, I'm a believer in the private sector being incentivized to be more cost effective than a government operation. In the test support fleet area there's a lot of room for improvement by contracting some of it out, particularly in a period of highly fluctuating workload.*

H: Looking back, based on your years of experience in the field, what do you see now as the top one or two challenges that we have to address in the test business, given the fact that we are dealing now with such sophisticated technology with such long lead times in terms of development and with very uncertain and almost constantly shifting operational requirements?

A: There are two challenges. One challenge we face in today's environment is infrastructure downsizing. We will address that later.

From a technology point of view, we have a very large and unwieldy bureaucratic process which is geared toward major system development. We have to reorient that process to becoming more efficient, focusing on efficiency at the subsystem improvement level. I've said this many times in many talks: process efficiency has not been seriously addressed by the government development or test community. People don't even think in terms of cycle time reduction. They have always had, at least at the major system level, all the funds they needed to execute any reasonable T&E program. [Until the early 1990s] money was never really an issue, at least to the extent that anybody ever thought seriously about test efficiency with the major systems. With some of the smaller laboratory programs, it's a different story [with money a major issue].

On major programs, the orientation in the past has been to measure everything, record everything at a very high data rate, and process just about everything you recorded. We've become inundated with data, but we haven't focused on information. The focus needs to

*AFFTC began downsizing the fleet even before its transfer and in 1995 was looking closely at contractor operations.

be on efficiently providing the requisite information to assist in a timely decision-making process. Very few people really have that orientation, and a major attitude adjustment is needed. Efficiency means being able to scale things up or down to meet a specific need and not burden people with a very large infrastructure when they only need a small piece of it. Improve customer orientation; reduce cycle time.

Integration of Test Ranges

H: In the 1960s, there was a project called Have Edge that evolved into the notion of a Continental Operating Range [known by the acronym COR, pronounced "core"] and this was eventually shut down in 1974 by Congress. Were you involved in that activity at all from the Edwards perspective, or did you have any views on it?

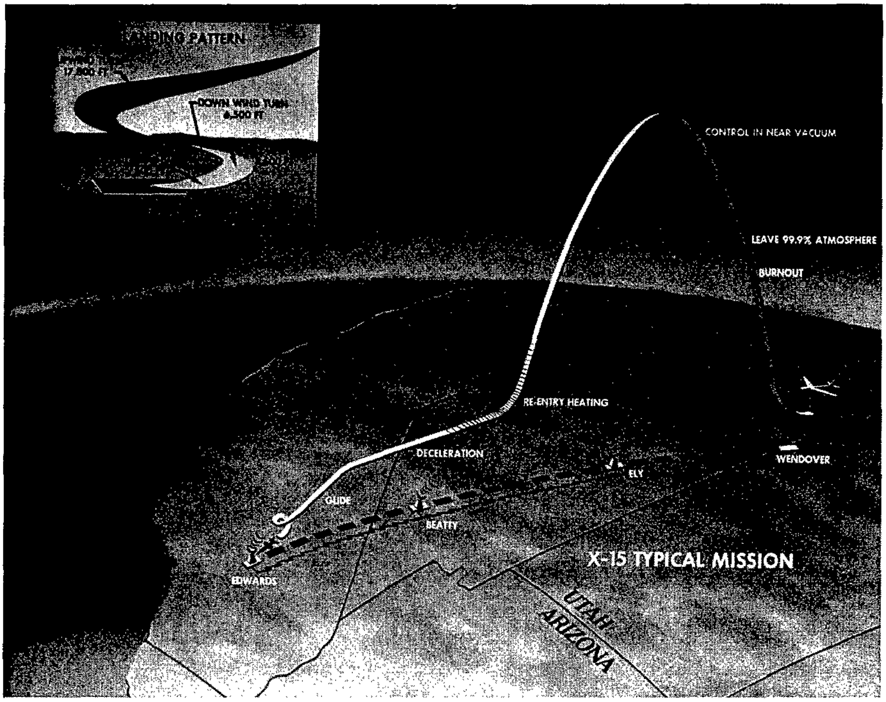
A: I wasn't involved. Prior to the late 1970s, I was involved in aircraft testing as opposed to test range and test facility development. I was aware of the COR effort. I later became interested in the concept, particularly since I've been back at OSD.

The Continental Operating Range was the first serious attempt to get an integrated range by the Air Force. It was managed by an organization called the TESPO [Test and Evaluation System Program Office, assigned to the Air Force Special Weapons Center at Kirtland AFB, in the early eighties]. They were looking at linking together the southwest training and test ranges. The effort was abandoned for reasons I am not quite aware of.*

In the early nineties there began a new effort to link the southwestern U.S. test and training complexes together to support both test and training activity.** Technology is now available to link simulations in a distributed fashion. The Navy could operate out of Fallon [Naval Air Station, Nevada], fly over a Nellis range, and then display the data back at Fallon or fly over the Utah Test and Training Range [UTTR]. There are a number of ranges that are instrumented, but the problem is they are not compatible and they are not linked. From a training perspective, it makes a lot of sense to do that. From a test perspective, the ranges are linked to some degree (see Fig. 2).

*Congress withdrew funds from the FY74 defense budget for development of the Continental Operating Range (which some saw as a competitor with existing ranges, particularly that at Eglin AFB in the Florida panhandle).

**In 1984 AFOTEC had initiated another ambitious range project to develop a comprehensive electronic combat test capability (ECTC) for the Air Force and the other services. Work on this complex, which would have established a Soviet-style integrated air defense system (IADS) centered at the Utah Test and Training Range, died with the first round of post-cold-war program cancellations in 1990.



Source: AFFTC History Office.

Figure 2. Typical X–15 High-Speed Test Profile in the Early 1960s

As you mentioned, that started with what NASA developed as the "High Range" for the X–15—linking the Ely, Nevada, site and Beatty, Nevada, and on down to Edwards. When the X–15 program terminated, NASA turned Ely over to the Air Force, and the Air Force ran and expanded it, with a link from Edwards over to the Pacific Missile Test Center [headquartered at Point Mugu, California]. It was expanded in the late 1970s to support cruise missile testing [AGM-86A and AGM-109] with a link from the Ely site to another site at Goshute, on the Nevada-Utah border in the Goshute Indian Reservation, and then two more sites on into Hill Air Force Base. This entire range was expanded and renovated to support the first cruise missile testing. It later supported low-level testing of the B–1 and B–2.

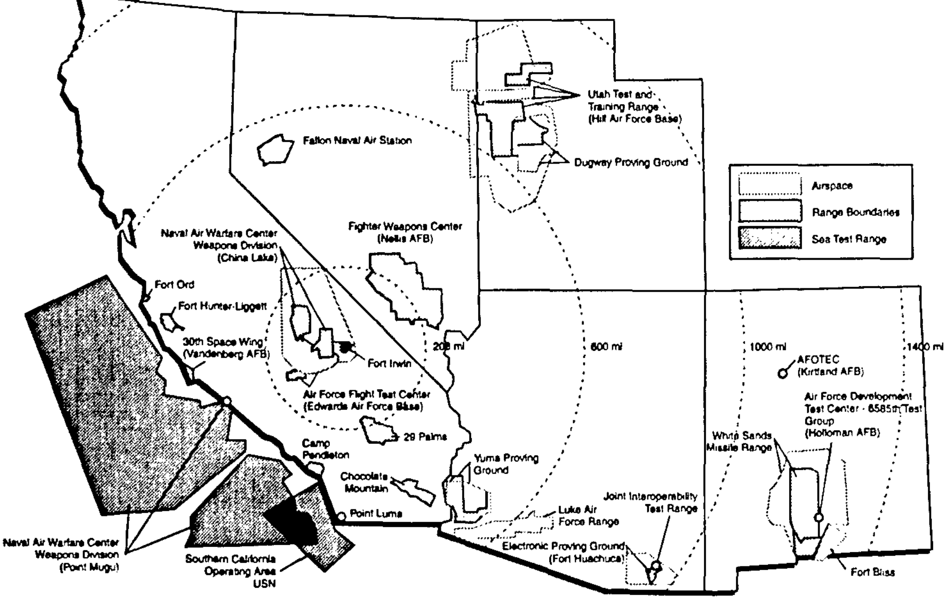
H: To me in the early 1980s, that was something that came as a real surprise. In many ways the most interesting aspect of the cruise missile program was not so much what was happening with the cruise missile itself but the fact that the cruise missile program had spawned the emergence of an interlinked range complex in the southwest. It went literally from offshore in the Pacific Ocean all the way into the Utah Test and Training Range. That had profound implications then for, as you have just said, expanding testing to cover a wide range of other programs, strategic, tactical, training, or whatever, or even to almost run a theaterwide air operation.

A: That's really the issue: Where do we go from here? The last big catalyst for the interlinking, as you mentioned, was the cruise missile. Simulation of theater operations in training will be next catalyst.

H: So the linking of the ranges, beginning in the late 1970s and into the 1980s, gave you, in effect, the basis for what we have now—as imperfectly as it may be—nevertheless an interlinked series of ranges in the southwest. The prospect now is perhaps several orders of magnitude increase in efficiency in doing test operations.

A: Nothing has been done since the cruise missile testing to enhance or exploit the interrange capability until very recently. There have been a couple of catalysts. On the training side, the Navy and the Air Force need to have compatible display and debriefing systems so that they can fly air combat training missions from Fallon Naval Air Station, Hill Air Force Base, and Nellis Air Force Base, using ranges other than their home range. That's being worked on the training side. The Army's national training center at Fort Irwin [California] also needs to be linked into this training net (see accompanying map).

Southwest United States
Test and Training Range Complex



Source: Office of the Under Secretary of Defense (Acquisition & Technology).

On the test side, there is a requirement to link ranges for tests and joint exercises. If you look at the geography and topography, you could have carrier-based operations in the offshore ranges. Naval aviation strikes could be launched against targets at the Nellis complex, Fallon, China Lake, or the Utah Test and Training Range. Air Force operations could be from a number of training or test ranges, depending upon the scenario. You could provide topography, targets, and threats for a wide variety of combined exercises and joint tests. That's been done to some extent for years with the Red Flag and the Green Flag exercises, and joint warfare interoperability demos. From a test perspective, for what we call joint tests, we have looked at how we might better use today's simulation technology to augment field tests. The challenge is to augment actual test and training assets through simulation of additional threats.

H: Like an IADS [integrated air defense system].

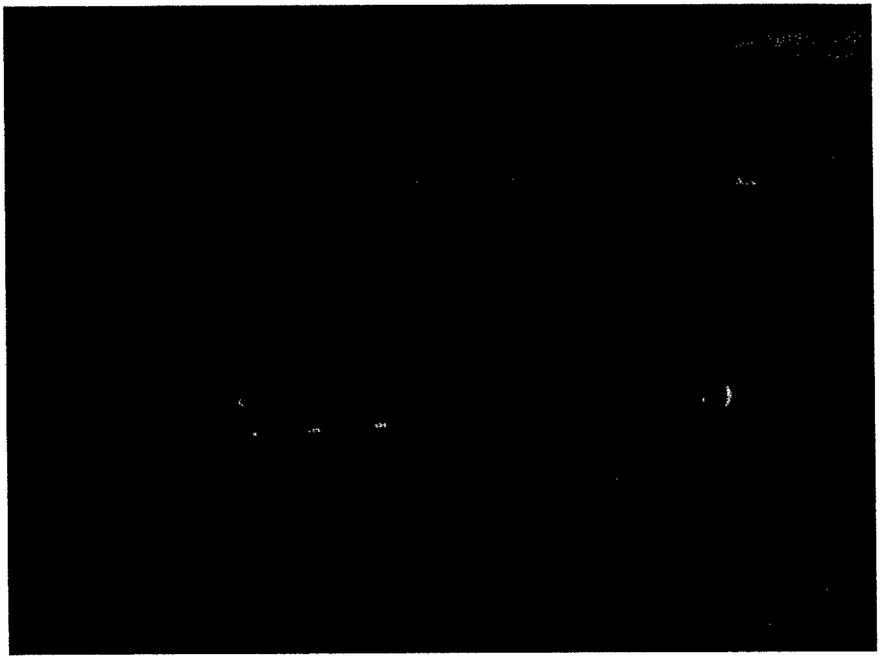
A: A dial-an-IADS capability—as examples, a North Korean scenario or an Iraqi scenario. Pick a scenario and simulate the requisite threats. The aircrews would not be able to tell the difference between what actually existed on the range and what was being uplinked from simulations.

H: One of the problems now we are beginning to see with Red Flag [large-scale air exercises centered at Nellis since the mid-1970s] is that, at least from the perspective of the airborne warning and control system (AWACS) people I've talked to, it's predictable. If you have been there once or twice, you know what you can expect to see, where it's going to come from, and what the threat magnitude will be like, and after a while you begin to lose that edge that's given to you by encountering unexpected situations.

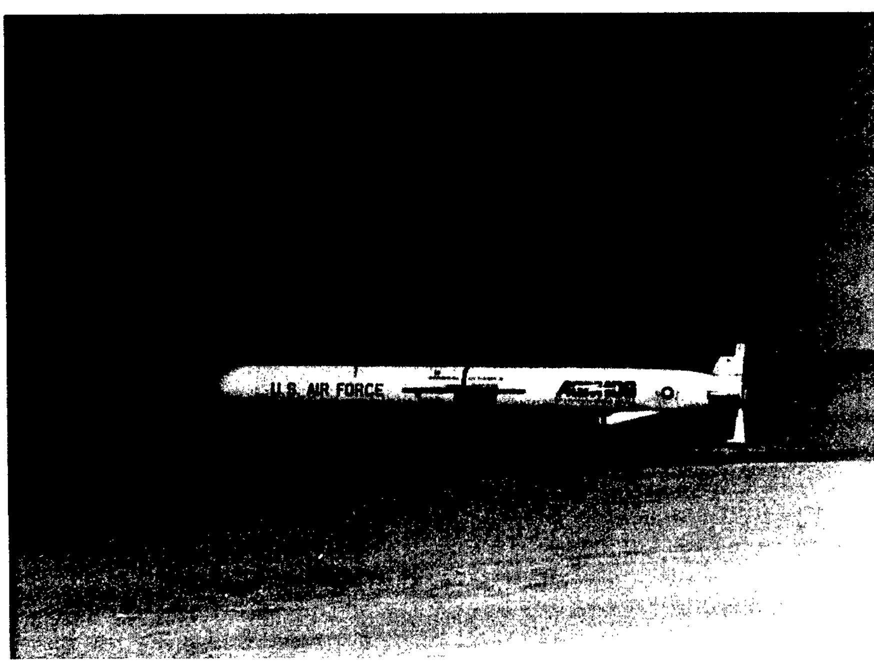
A: Absolutely.

H: You have the surrogate capability off the California coast of substituting for the Red Sea or the Persian Gulf and you could have a Luke—UTTR operation that would emulate a King Khalid versus Baghdad airfield complex scenario. As you said, you can dial in whatever you wish.

A: That's the direction technology is moving. There are a number of impediments. The training people are concerned, and justifiably so, about encroachments on the part of the test community. Testing tends to out-prioritize training, so there are some organizational and some cultural issues to work, as well as some technical ones. The utility of the so-called distributed interactive simulations for training is relatively well understood. The value for testing, particularly developmental, is more limited, and there are technical



Air Launched Cruise Missile Competition: The Boeing AGM-86B *(top)* and the General Dynamics-Convair AGM-109 *(bottom)*. Testing them required integrating several test ranges in the western United States.



problems that must be solved. The same tools that are used for training can be used to some extent to drive and refine requirements because you can do the force-on-force, many-on-many scenario. Testing, at least in the developmental world, is a one-on-one or a one-on-few activity. When you migrate to operational testing, you sometimes get involved in a few-on-few scenario and occasionally a many-on-many scenario. Only with some of these very large joint exercises do you get the actual many-on-many joint scenarios that you encounter in a warfighting situation.

H: What's interesting, I think, is that historically we have seen testing as distinct from training, as distinct from force structure issues, as distinct from operational issues, and what you now have with the distributed network like this is the ability to bring all of them synergistically together and to teach that person in Air Force Materiel Command (AFMC), for example, that his concerns as a tester are not all that dissimilar from the person in the new Air Education and Training Command (AETC) or the person in the Air Combat Command (ACC) who is trying to work together to resolve issues in projecting military power or determining what system they want for future combat.

A: There is an additional dimension. It also helps in the requirements definition and refinement process. One of the big problems I've seen in the DAB [Defense Acquisition Board] reviews, is the issue of requirements definition. These tools provide the capability to link the requirements, tactics, and doctrine and to iterate them back and forth to an extent that we've never been able to do in the past. It's really a question of how you do that organizationally, how you use the tools, and who is responsible for the development of the tools.

The test community is the primary driver for high-fidelity range instrumentation and high-fidelity models and simulations. The test community must either take the lead or play a major role, but they will have to expand their horizons. As you said, things are somewhat stovepiped now between test, training, and requirements definition. We've got to get rid of some of those stovepipes and look across them so that we are using the same tools where appropriate to support all of these activities. During the previous attempts to establish a comprehensive range in the 1970s and 1980s, interservice cooperation was certainly not at the level it is today, nor were the technologies available to facilitate interservice operations.

There has been a remarkable transformation in the last ten years at ranges such as the White Sands Missile Range [New Mexico]. We were both out there three weeks ago to give talks at the Inertial

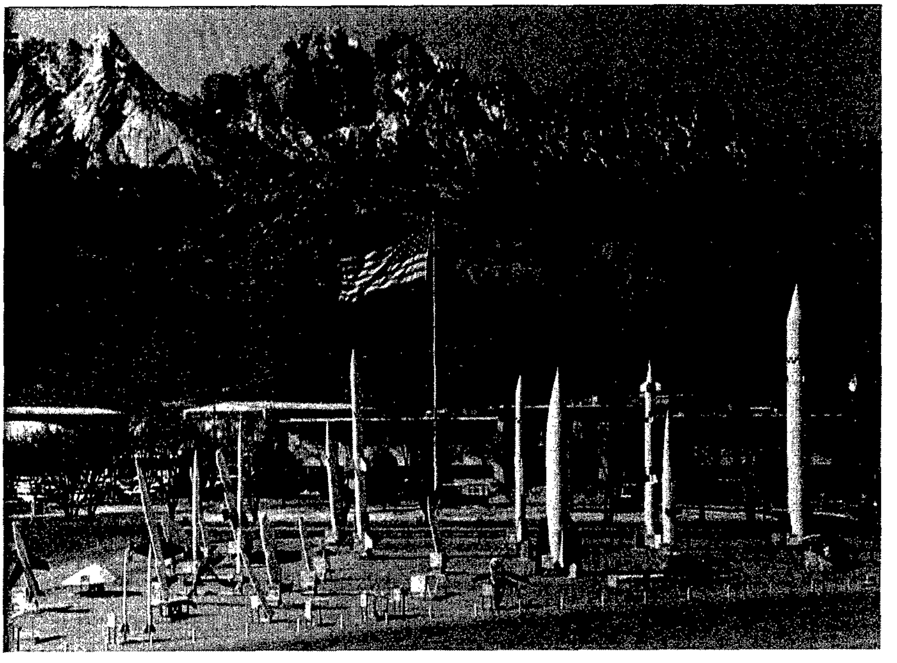
Guidance Symposium at Holloman Air Force Base. I consider the White Sands complex the prototype of where I see the ranges going in the United States. First, there is a combined interservice test and training range. Cooperation exists between the Army, the Air Force, the Navy, and the Defense Nuclear Agency [DNA]. The Defense Nuclear Agency, in the northwest quadrant of the White Sands reservation, has a sophisticated array of instrumented hardened targets. That's where they did the above-ground large conventional tests as surrogates for nuclear explosions. Instrumentation to measure shock and blast effects is available. The test complex supports U.S. test activities and NATO as well. I was out there several months ago for a large above-the-ground test; they allowed us to get reasonably close. I don't recall the exact overpressure where we stood, but you could see the shock wave coming at you and feel the pressure pulse as it passed by. It was like a schlieren photograph, the same thing you've seen on the desert from an airplane low-level sonic boom.

B: They call that part of White Sands the Stallion Range.

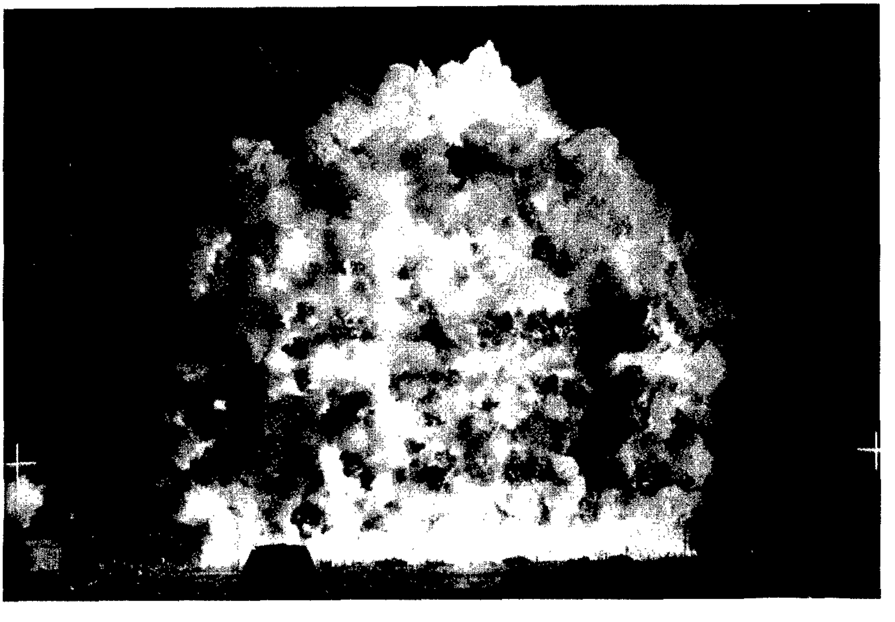
A: Yes, the Stallion Range Area. They are currently building a large-blast thermal simulator test bed. This is housed in a large building which is now under construction. By the way, our OSD T&E office cosponsored the facility. It was originally a CTEIP [central test and evaluation investment program] funded activity, but because it was primarily military construction, it was later moved to a DNA line. You can systemically test shock and blast effects in a controlled environment. It's analogous in some respects to a shock tunnel or flow down tunnel but to look at destructive effects rather than to acquire aerodynamic data.

The Army has accomplished its surface-to-air testing over the years at White Sands: the Stinger, the man-portable Stinger, and the Avenger, which is the Stinger mounted on a Humvee, the Hawk, Patriot, and a variety of other weapons. There is a high plateau there called North Oscura Peak. Aircraft can fly at low level and up toward that peak to evaluate sensors in a lookdown clutter environment.

In addition, there is the Air Force test activity at Holloman Air Force Base. They have the target drones there and they also accomplish radar cross section measurements at the RATSCAT [radar target scatter] and RAMS [RATSCAT advanced measurement system] facilities. The inertial [guidance] test facility is there as well. The Air Force has used Holloman as a test and training facility for years. Now, of course, the training activity for the F-117s is there. To the



Missile Park at White Sands Missile Range, New Mexico, with Organ Mountains in the background. . . . Conventional explosives used to simulate blast effects by the Defense Nuclear Agency at White Sands' Stallion Range in the early eighties. Note how the explosion dwarfs the vehicles parked nearby.



south, contiguous to White Sands, there is Fort Bliss, Texas, the home of Army air defense since World War II.

It is also the site of two permanent German training activities in the United States: The German air defense activity at Fort Bliss and the transition or checkout in the F–4 for the Germans now at Holloman. A Navy captain is vice commander at White Sands. The Navy has a ship site where they test surface-to-air and surface-to-surface ship protection systems. In addition, there is a small NASA facility there as well. In summary, I would characterize White Sands as the prototype of the future in terms of a multipurpose, multi-function, joint service test and training facility.

H: That point would have really surprised me before I had gone out to White Sands. On that trip, I had a vacant morning and I was able to get down to White Sands from Holloman and got a range tour. I had always thought of White Sands as merely a missile range, but it is a lot more than that.

A: Let me move on to another issue. Budget pressures and the need to cut down on infrastructure costs is one driver for consolidation. Another driver is technology.

Let me address the technology driver. Aircraft platforms have migrated from what I would characterize as designs of the fifties and sixties, which were largely independent hydromechanical subsystems, which were mechanically or hydraulically linked to today's highly integrated systems. Many of the same sensors supply information to all of the subsystems—flight control, fire control, engine—through a data bus. So it's a highly integrated platform and that integration now includes the weapons. Designs went to conformable carriage to minimize drag or internal carriage for signature reduction purposes. Historically, we did a lot of the development at the platform level and the weapons level separately. Those activities were separate and distinct functions. The distinction is far less clear today. The distinction between weapons, offensive and defensive, and platform work is less pronounced than it was with earlier generations of technology. That's one thing that will drive, over time, systems to be tested as an entity.

Another factor is we've gone in the last twenty years from high and fast to flying at relatively low level to evade detection. We need the varied topography to test platforms and the terrain avoidance systems. For all those reasons, I think we see a further migration of test activities to the southwest.

There is another consolidation driver, the EW [electronic warfare] area. We can no longer afford to upgrade all of the capabilities we now have. The issue becomes classified if we get into specifics. To acquire either real or surrogate double-digit air defense threats is a very expensive proposition. Suffice it to say that we are no longer going to be able to afford upgrades at several locations.

H: And spread the cost burden around.

A: For all the above reasons, I see a migration of aircraft and weapons test activities to the southwest. Probably the most emotional component of the issue is the direction the Navy takes vis-à-vis Pax River. My vision is that the high-performance aircraft will migrate to the southwest. I am adamant on the point that it has to be done in a way that protects the Navy's equity. I think that can be done by basing naval high-performance activities at China Lake. That doesn't mean the demise of Pax River by any stretch of the imagination, far from it. I believe the Navy has come to this realization and they are now integrating the Naval Air Warfare Center, the lab function, at Pax River. They are moving NAVAIR [Naval Air Systems Command] headquarters there [from Arlington, Virginia].

It's a facility with a lot of capabilities and talented people. [Pax River] will continue to prosper but in a slightly different role. It is now the southeastern boundary of the greater Washington-Baltimore metropolitan area and encroachment pressures will increase. If you move the high-performance aircraft westward, then you move the Navy test pilot school. I think the two test pilot schools will be integrated, with command rotating between the Navy and the Air Force. For the reasons I've already mentioned, most of the Navy's weapons testing is now done out at China Lake. I think there will be a further migration of high-performance test activity. It's going to be a traumatic move, but I think it will come about.*

H: Low-speed aircraft work would still be functional at Pax River. I think you mentioned at one point VSTOL [vertical/short takeoff and landing].

A: Antisubmarine warfare systems and VSTOL. One of the two candidate platforms under the joint advanced strike technology (JAST) program is a short and vertical takeoff and landing platform. A decision hasn't been made, but that may well become one of the platform technology demonstrators. It's going to be interesting to

*Although various studies, including the "Report on the Roles, Missions, and Functions of the Armed Forces of the United States," issued by the Chairman of the JCS in February 1993, have recommended increased use of the ranges in the southwest, the Navy and the Maryland congressional delegation have continued to resist the idea of removing any significant aircraft testing capabilities from Patuxent.

see where that test activity is undertaken, whether it's at Pax or elsewhere.

Let me make another point. We have talked about the test component. I want to address training. Test and training activity, at least in the aviation business, has been integrated for years. Point Mugu and Eglin are prime examples of that, where test and training activity is collocated. It coexists, as I mentioned earlier, at White Sands. The name of the range in Utah is the Utah Test and Training Range, although the reality is there is very little test activity there. Less than 5 percent of the missions are test missions. The R-2508 airspace, which encompasses the Navy ranges at China Lake, is shared by the Navy and the Air Force for testing. It is also used by the Navy for training out of Lemoore NAS. Those offshore Pacific missile ranges are used for test as well as training. So, in the aviation area there is joint testing and training today.

In the training world there are incompatibilities between the Air Force and Navy air combat maneuvering instrumentation. An aircraft can't fly out of Fallon, as an example, against the Nellis range and use the Fallon display and debriefing system. The Air Force and the Navy ought to be able to fly on each other's training ranges with compatible instrumentation. The Air Force has two different training instrumentation systems. The one in Utah, a multilateration-based system, the high accuracy multiple object tracking system (HAMOTS), is compatible with the unique test system that is there. It is incompatible with the training systems at Nellis and Fallon. There is a need to achieve compatibility so that the Navy and the Air Force high-performance aircraft out there can be flown on multiple ranges for a variety of training experiences. There is an initiative to make the next-generation air combat training system be compatible and interoperable between the Air Force and the Navy.

One cultural problem is integrating the Army into this, starting with the Army National Training Center at Fort Irwin [California]. The Army is very concerned about integrating test and training. Their concern has to do with their ability to accomplish the training objectives, being out prioritized by the test world. The Army has resisted, to a greater degree than the other services, being integrated into a large test and training environment. I think it's inevitable in the southwest. We're now beginning to look at electronically linking the ranges. That was the tasking from Secretary of Defense Les Aspin in April 1993. The test ranges are already linked. Linking the test and training ranges facilitates acquiring test data as well as training data. Our office sponsors joint tests. We now piggyback, and have for years, many of the joint tests on the Red and Green

Flag exercises at Nellis simply because all the assets are there, both the Navy and the Air Force aircraft. The aircraft assets are there, so you can extend a week or two for joint tests. The most recent one was the joint air defense operation, joint engagement zone [JADO/JEZ] test at Nellis.

Those activities, because they are so resource intensive, can supply a database in the future for test purposes as well as training. In addition, the information can be used to verify and accredit simulations. People say we need to train the way we fight; that's certainly true. I don't want to say the training exercises are unscripted free play because there are a lot of constraints and to some extent they are scripted, but to acquire test data, you may want to have a little more scripting and instrumentation.

H: A more laboratory-type environment.

A: Yes, more of a laboratory environment, but having the instrumentation available so that you can capture the results and use them to update the simulations that now exist in the various battle labs. We are going to link things electronically for that reason.

I don't see, with a couple of exceptions which I will cite, a lot of near-term potential for enhancing the initial developmental testing because it's more of a one-on-one subsystem level activity. I think the one area where you can is by integrating a digital simulation of threats with the actual hardware we have on our open-air ranges. You could better replicate a threat density. That's one that readily comes to mind. In the other areas it's difficult for me to see how the linking will greatly facilitate testing at the developmental level. At the operational level, I think it's going to be a tremendous boon. Also, I think as things migrate geographically for a wide variety of reasons, some of which I've already discussed, it will make our ranges more flexible and useful.

The Navy has moved in the direction of warfare centers. One of the four warfare centers is the Air Warfare Center. The Navy vision is a full-spectrum capability which collocates geographically the labs, the test and evaluation function, and service engineering support. They've done that in the Air Warfare Center at Pax, the Undersea Warfare Center at Newport, Rhode Island—they are migrating from New London to Newport—and the Surface Warfare Center.

In the case of the Naval Air Warfare Center, this includes labs, the T&E function, and follow-on service engineering. Obviously, you have collocated the technical base, but service engineering goes on wherever the fleet is. So the activity is worldwide, but you've collocated the technical base. You don't need to do all the testing there, with today's technology. As I mentioned, some of the Navy high-angle-of-attack testing has been done at Edwards for years to take advantage of the lake bed and airspace. It is an example of distributed testing but continuing to concentrate the analysis. I think that's a model that has increased potential with today's data acquisition and transmission technology.

Another opportunity that will come from electronic linking is more immediate access by designers and program managers to results. That's a mixed blessing because data will be more readily accessible. I think, on balance, that will be good because there won't be the time delays and the past penchant to massage the information. I'm not saying that in a pejorative sense. Real-time systems don't allow you to do the QA [quality assurance] before other people have it. Back in the old days, people not only did a QA check on the data but they were very reluctant to release data until they did a lot of preliminary analysis. On one hand, people are going to have more rapid access to data, but the downside is that people who aren't trained in data reduction and analysis are going to be faced with the situation where they are going to need some internal analysis capability for the information to be made useful.

By the way, in the spring of 1993 Secretary of Defense Aspin directed OSD and the services to streamline the T&E infrastructure. The objective was to achieve management efficiencies by better integration of facilities. The results from that study are documented in a briefing I gave.*

Improving Test Systems and Instrumentation

H: You have put the southwestern range complex into a much broader context than merely just looking at "a few ranges down there getting together." Now we can see the real rationale behind it. To a degree, we have really talked about this in some of the things we have alluded to here, and it leads into it and that's the efficacy of T&E resources. Jack Krings used to complain when he was Director of OT&E that investments in T&E resources had fallen far behind that of the acquisition programs. Do you basically concur with that view? Given the changes we've seen in the world, the end of the cold war for example, do you think that's still legitimate? In your particular case, how have you addressed this issue? What have you felt about it?

*See appendix H for "Test and Evaluation Facilities Fast Track Study."

A: Jack started an initiative to create a central investment fund so that the OSD T&E offices could invest in major upgrades across the ranges. He believed the ranges were falling behind and that the systems that we were attempting to test had capabilities that exceeded our ability to measure and evaluate those capabilities. I certainly agree with that, and I supported Jack in this effort. Mr. [Robert A.] Taft [III] was the Deputy Secretary of Defense at the time. Jack was successful in convincing Mr. Taft that we needed a central test equipment line. Our office eventually became the repository of those funds. I won't go into all the history, but what has evolved is that there is roughly twenty million dollars of that fund that is used annually for relatively short-term upgrades in the operational test arena. The rest of it, about $100 million annually, goes for capital investments in generic capabilities that support both developmental and operational test. I think it has been very successful.

Let me mention in my opinion a few of what I think are the major successes of that endeavor. The first, of course, is the GPS [global positioning system] range application systems. The seed money for that came from OSD prior to the establishment of the central fund. That effort has been managed very capably by Tom Hancock at Eglin AFB over the years. It was a prototype of a generic investment that impacted the entire test community. Some of the early programs that were funded by the central test line include the common airborne instrumentation system (CAIS), which is an instrumentation system to be used by the majority of Army, Navy, and Air Force aircraft. I say majority; there will be a 10 percent high end recording requirement for a special system. This is a primary pulse code modulation system. The Navy at Pax River manages the program.

H: So CAIS has basically replaced this earlier AFFTIS [Air Force Flight Test Information System].

A: CAIS is an improved next-generation version of the AFFTIS and is now being managed by the Navy at Pax River, with Air Force and Army people participating. It's a joint program office in effect. CAIS is now well along in the development. It's a system designed to meet high-performance fixed-wing aircraft requirements and rotary-wing vehicles such as the Comanche. It will be used with it, the F–22, and on the F–18E/F. It will be the mainstream instrumentation system for aircraft flight testing for the next thirty or forty years. The design is a modular building-block approach.

H: This is a major development. When you really think about it, you are setting something up that's going to be with you for almost half a century.

What I find historically bothersome is that people fail to learn these lessons. If you cut back on a system that's immature when you have the money and lose ten years, and then you try to develop it in lean times, think of the problems you have made for yourself.

A: This is a particularly sensitive issue with me. We have learned so many times that you don't want to develop an instrumentation system at the same time you develop an aircraft, yet we have done that time and time again. I can cite many programs. As an example, I was involved as project engineer of the B–52H. That was the first actual test program at the AFFTC to use a pulse duration modulation instrumentation system. It was the first use of that type of system by Boeing and the Air Force on a large aircraft. Not only was the instrumentation system new but the pressure transducers were as well. We had so many problems with the instrumentation system that we had to revert to a backup old photo panel system to measure pressures. The pressure transducers were about half as big as a coffee can. They were a high-volume transducer, or at least their internal volume was high, and therefore there was a lot of lag. Dynamic data in a B–52 isn't all that dynamic. Still, the transducers were totally inadequate for transient measurements.

H: Totally behind what was happening developmentally.

A: Absolutely. We had to revert just to acquire the steady-state data. We put a photo panel in parallel and with the long lines that you needed for a photo panel, you can imagine running a pressure line from an outboard modern engine on a B–52 to some central location.

H: Some seventy or eighty feet.

A: Oh, yes. So the lag in that instrumentation was out of sight. The instrumentation was unsatisfactory for defining non-steady-state characteristics.

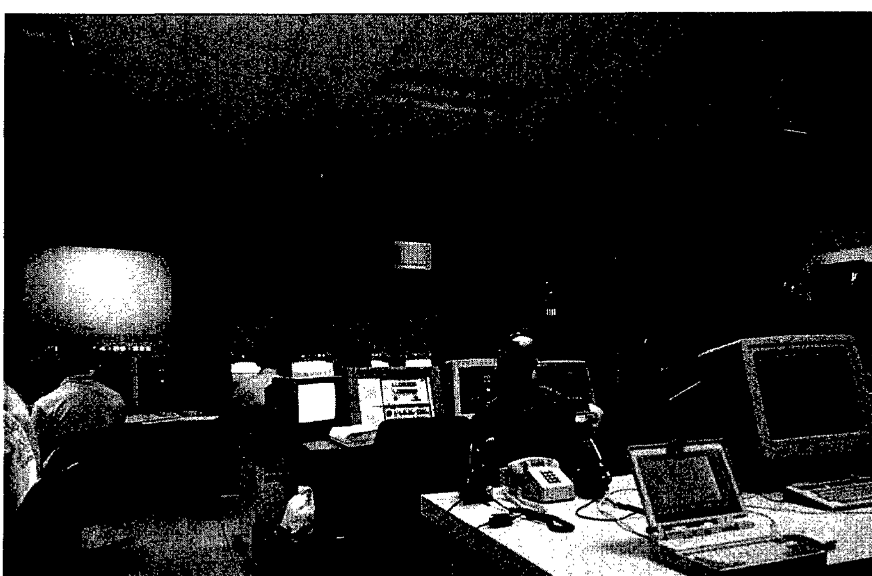
Back to the original point, I can name several programs where we had similar experiences. Contractors would use their own unique data acquisition system. The process of debugging the data acquisition system went on for the first two years in the flight-test program. It was totally unsatisfactory. Another problem in those days was the need to reinstrument an airplane when it transferred from Edwards to Eglin because of the lack of a common instrumentation system. Air Force Systems Command headquarters became very concerned, and that was the catalyst for the AFFTIS. You don't need to have a total commonality, but there has to be common interfaces so it is functionally compatible. More than one vendor can produce components as long as there is a functional commonality. There are two or three prime producers of aircraft instrumentation, aside from the various sensors, because there are numerous companies producing sensors. You don't want to be locked into the next thirty years to any one company.

H: In a somewhat related way, the point you've just raised touches on another one and that's the whole national competitiveness issue. With the drawdown, many of these smaller, specialized suppliers may not any longer be out there, or if they are, they may not be able to devote a significant effort to just one customer. You can find yourself depending on a system that, all of a sudden, is nonsupportable; it's a tricky business.

A: Yes, it is.

N: Mr. Adolph, you made the point earlier that the CAIS system was modular. Do you think that will enhance the ability to upgrade it as new technology comes along? For example, Bob Webb at Eglin is developing one thing he calls "peel-and-stick" technology, where you can literally peel off and stick onto the skin of an airplane a sensor of some kind.

A: Very definitely. By the way, Webb was the senior Eglin participant on the AFFTIS system. Modular components do two things. First, as technology moves ahead, we'll be able to miniaturize individual components without having to redesign the whole system.



C–17 room in the Ridley Range Control Facility at Edwards in 1993.

Second, modular systems are readily expandable. If you have a distributed system and you test a big airplane like a B–2, the system is modular and expandable so that you can locate and distribute components throughout the platform.

It allows for design flexibility. It also allows us to have a depot level of support infrastructure that cuts across all three services. I'm going to touch on that later on in conjunction with the whole test support infrastructure issue. In the past, we would have a pool of instrumentation spares at every facility. You need some limited change out capability close by because if you preflight the instrumentation and you find a problem, you can take it into your shop and change out the defective component. You still need that quick-reaction capability at each of the test locations. Depot level repair of all these modules can and will be consolidated across the services and you'll have central supply. Again, back to this commercial practices business, you can ship things overnight commercially, certainly anywhere in the United States, so we don't need to lay in a high level of spares at every test facility anymore. There is no need to do that.

I want to expand on that thought. It applies not only to the airborne instrumentation but it applies to most of the other systems that we use out there to support test facilities. We need to aggregate depot level instrumentation and range system maintenance, aggregating that across the services. That has been done for years with instrumentation radars, the FPS-16s. There has been one support contract. The system was built by the RCA Corporation and RCA had a sole-source contract for a number of years. That contract was later competed, as it should have been.

There's legislation proposed this year—I don't know if it has made it through all the wickets—that depot level maintenance of the electro-optical equipment be consolidated—that is, the optical trackers we have at the ranges. I think that makes good sense. The maintenance should be competed. Another area that should be aggregated across the MRTFB [master range and test facility base] is upgrades to mission control rooms. There have been at least three generations of upgrades. The first generation was the Grumman system, which was developed under Navy auspices by Grumman at Calverton.

H: This is the old automated test system, the ATS?

A: Yes, the ATS. It was later cloned at Pax and Edwards and a number of other places.

H: That seemed to be a milestone system in its own way at that time.

A: Yes, it was. It was the first generation relatively automated real-time system. There were a couple of other generations developed, one of which was IFDAPS [integrated flight data acquisition and processing system] developed at Edwards and later cloned at several other ranges. A later upgrade was developed for the self-contained B–2 facility at Edwards. If you're displaying PCM [precision countermeasures] data at Eglin, Pax, Edwards, Mugu, China Lake, or White Sands, the displays that you use for time histories, pressures, temperatures, rate data, are the same. Basically, it doesn't matter whether it's an aircraft or a missile. The government has paid for a lot of those displays and display software over and over again. There's a move afoot now to use some commercial practices in this area, and I think that can and should be done. The oversight and the management of that is best done centrally so you can look across the whole MRTFB. There is nothing unique about measuring pressures and temperatures and rates using PCM systems.

I believe the acquisition of those kinds of systems should be centrally managed. When I say centralized, I don't mean that you totally do away with the organic capability at all the field ranges. You need some limited capability there because there are minor ongoing upgrades that are really worked in over the weekend or on evening shifts. You need some in-place on-site capability to do that.

Let me mention a couple of other systems that we are developing. The OSD office is funding a "Smart Munitions Test Suite" in conjunction with the Army at White Sands to support the next generation of smart munitions testing. The upgrade will support testing of munitions such as the Army's brilliant antitank (BAT) submunition. It became obvious during the initial test planning for the BAT that the existing instrumentation was inadequate for testing wide area smart submunitions—to be able to track a number of submunitions simultaneously. This upgrade consists of modifying the multiple-object tracking radars that now exist at White Sands, in addition to some changes to electro-optical systems to give the required data. This is another initiative that is being jointly funded, in this case by the central test equipment line and the Army, to provide the next-generation capability to evaluate smart submunitions.

Another one that comes to mind is the Aerial Cable Range Facility, again at White Sands. The requirement for this facility grew out of a facility that Sandia [National Laboratory] runs on the Kirtland Air Force Base reservation at a place called Coyote Canyon. At Kirtland targets are slid down the cable replicating helicopters and other targets. It was used extensively to test the Stinger missile, including flare and other countermeasures effects. You can put heaters on the target and duplicate the IR [infrared] signature of a helicopter. You can also evaluate a wide variety of flares and see how effective they are as countermeasures. Much of the development of the Stinger reprogrammable microprocessor sensor was accomplished using that facility. There are a couple of problems with the Coyote Canyon facility. One is, depending on the trajectory of the missile fired at it, there is an area that is used for grazing sheep behind it. So there are some limitations on firing directions and trajectories. In addition, the facility can't handle high-speed targets, and you don't have the variety of clutter backgrounds. As a consequence of the limitations on the system, we undertook to develop an improved system about four years ago. By the way, we were out at Coyote Canyon one weekend; in one day you could fire several shots under highly controlled conditions. It was remarkable, the number of tests that you could fire relatively inexpensively under a controlled environment. There is just no comparison to shooting a droned aircraft. The fidelity isn't 100 percent, but it's adequate for most types of development shots. As a consequence, there is a facility that's now being built at White Sands that provides a much more robust capability, both in terms of the speeds at which the targets can go down the cable and in terms of a wide variety of clutter backgrounds. That facility cost is around twenty-five million dollars, which can be amortized in the first year or two of testing.

H: That's not much.

A: The White Sands system is a three-mile-long Kevlar cable suspended across a canyon between two mountain peaks. It's the longest unsupported cable span in the world. The cable can hold up to 20,000 pounds. Captive vehicles can either be rocket propelled or gravity accelerated. This facility is a remarkable engineering achievement. Because this area is used quite a bit for low-level flying, the cable does not remain in the air all the time. When it's not being used, it's lowered.

H: It rests across the desert floor?

A: Yes, so they can raise and lower the cable, which was an additional complication in construction, but it retains the flexibility to use that airspace for a variety of other things. That airspace, by the way, has been used over the years for sensor testing. As an example, North Oscura Peak is a mesa or a plateau with a pronounced escarpment on the southwest face. Ground-based radar sensors have been mounted on the top of the plateau and target aircraft would fly up the valley. I was out there one time when they had New Mexico Air National Guard A-7s flying up the valley and using terrain masking. That area is used for those kinds of tests from time to time, so you want to minimize obstructions.

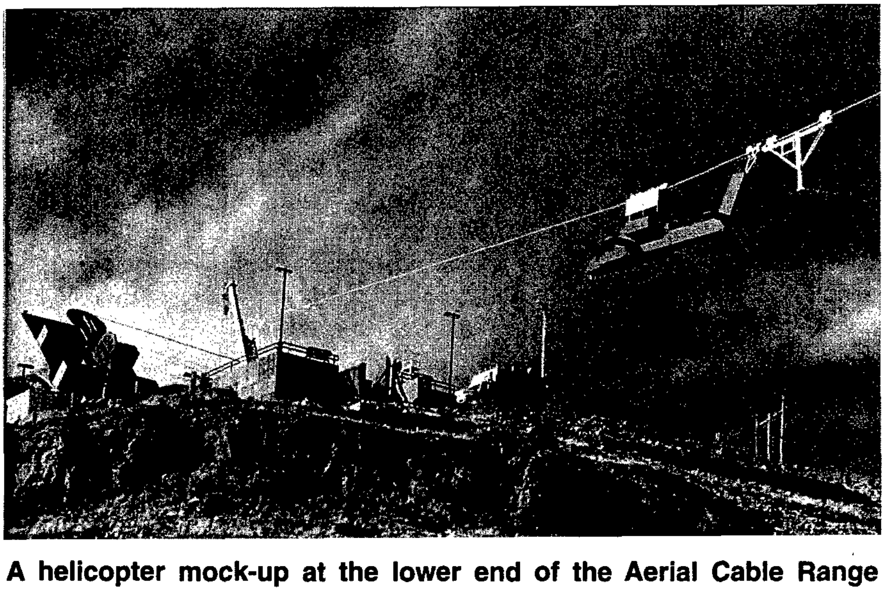
H: Is there any deterioration of the cable when you are firing?

A: You can hit the cable, so they've already bought two cables. Occasionally, the cable will require repair or replacement.

As a footnote to this facility, I think it illustrates a relatively innovative alternative application to providing what I would call full-up testing of either a platform or a weapon. In this case, full-up testing of a weapon can be accomplished against surrogate platforms that are relatively inexpensive but still have the requisite fidelity to obtain the information you need to develop the sensor. Innovations like this are particularly important in a fiscally constrained budget environment. In addition, during the Stinger test, there was no way that you could accomplish the numbers of tests that were required to refine and fine-tune that sensor, using real helicopter targets.

H: The cost would eat them alive.

A: Yes and the quality of the data from a target that is destroyed is often less. You lose some if not all diagnostic ability. The point is that you need to fine-tune software-intensive systems. To do that you



A helicopter mock-up at the lower end of the Aerial Cable Range Facility at White Sands Missile Range.

would need a relatively large database and you need to test rapidly at a variety of conditions to avoid stretching out the development.

H: What time period was this Stinger work taking place in?

A: In the early 1990s.

N: This test approach has a lot more flexibility than is immediately apparent. I've used the cable out at Sandia to hoist a [Mark 82] bomb up into position and then pull it down to impact beside a tank to test fusing. So it can be used in a variety of different ways.

A: That's right, Doug. You hit on a good point. The Air Force used the facility for testing the sensor-fused weapon [SFW—an anti-armor system that dispenses numerous self-guided projectiles over a wide area]. In fact some of the tank hulks used as targets were still there when I was there. They would slide the sensor-fused weapon submunition down the cable, and then at the appropriate point release it and let it do its trick against the target. These facilities will continue to be a powerful tool.

H: So you had the sled hooked up to a cable which basically was yanking the bomb down?

N: Right. So you ended up with a simulation of a vertical sled track, if you will.

A: Let me expand on that point and comment on fuse testing. It has always been a technical challenge, and it's becoming an even greater challenge. There is a new fuse facility being built out at China Lake. The acronym is MESA, measurement of electronic simulation something. It's under construction now and resembles a hangar in appearance. The purpose of this facility is to test proximity fuses in an air-to-air environment, doing tests similar to that talked about. You can evaluate proximity fuses against suspended targets from various aspects. This is becoming a more challenging technical problem as we go to low observable platforms, getting a prox fuse to activate at the appropriate point. The probability of kill is obviously a very strong function of the distance from the target that the warhead detonates. The limitations on our instrumentation are such that, for shots against drones, we may not know exactly where the proximity fuse initiates relative to the target. You obtain proximity fuse detonation from telemetry data on the fuse itself. You then need to time correlate between that data and the optical data relative to the platform. You can't refine the fuse from that kind of data, so you need these facilities.

H: To get a rear quadrant definition and a front quadrant definition?

A: Exactly, all of those. This facility is really unique. The capital investment is coming from the Navy; it's another illustration of a one-of-a-kind facility that we need as technology moves along. But we can no longer afford to have multiple facilities, so one service is going to develop it and it must be shared by all the services. By the way, the DOD IG questioned the need for the MESA facility. We got the issue resolved just in time to avoid construction delays.

Test and Evaluation Consolidation and the Reliance Program

H: You were involved with many of the range matters we've been talking about in connection with a major interservice effort known as the T&E Reliance Program. To put these all in kind of a single context here, we've a number of basic questions on the T&E Reliance Program. How did it get started? Who were the major players? What were the options that people were looking at? What was the role and function here of OSD? What was the role of the services, and what ideas did they have? Who were the major competitors in the field here? In your view of this and the progress to date and future prospects, how did you see your role changing and evolving as chair of the defense T&E steering group?

A: Before I address the reliance question, let me take a minute to give you some background on the origins of the defense T&E steering group. In the late 1980s the services did not have a clearly identified focal point for test and test resource matters. Jack Krings (the first Director of Operational Test and Evaluation) and I urged the services to identify a single point of contact to work with us. The Army identified Walt Hollis (Deputy Under Secretary of the Army for Operations Research) as their test focal point. Walt had a long and impressive background in operations research and test. The Air Force established the position of Director, Air Force Test and Evaluation which was filled by Lt. Gen. Howard Leaf, USAF, Retired. Because of his outstanding credentials, he was recognized throughout the Air Force from the outset as the individual responsible for test matters. The Navy's designated focal point for test matters was the Director, T&E and Technology Requirements; initially Vice Admiral Guy Reynolds and later Rear Admiral Bill Houley. These individuals were the service principals on the Defense Test and Evaluation Steering Group (DTESG), which I chaired. Maj. Gen. Mal O'Neill represented the Strategic Defense Initiative Organization (SDIO) office, and Don Linger was the Defense Nuclear Agency point of contact. The DTESG was the forum used to address and resolve a myriad of test issues, including reliance. We reviewed and approved the recommendations from the various reliance study panels. We also reviewed the services' annual test investment strategy and approved the OSD-funded CTEIP investments. DTESG meetings were frequent, always lively, and sometimes contentious. It was a pleasure to work with the caliber of people who were on the DTESG.

With that as a background, I will address the genesis of reliance. First, there was a consolidation activity which preceded reliance. The previous administration's Defense Management Review (DMR). How did it get started? There was an effort that really preceded T&E reliance. The previous administration's DMR initiative was an effort to improve the efficiency of the Defense Department. We started with a DMR study of potential consolidations of T&E facilities. That study preceded what came to be known as T&E reliance, but it really fed into it in many areas. The T&E reliance effort was initiated by a later specific defense management review item that called for it. The buzzword "reliance" came out of the science and technology area; there was a science and technology reliance effort separate from the T&E reliance initiative. It started with the DMR as well.

The major players were, of course, the services and the OSD developmental test staff. The operational test staff didn't play in it to any significant degree because the operational test organizations in the services don't run any test facilities, with the exception that the Army OT community runs [Fort] Hunter Liggett [in California]. The Defense Nuclear Agency played a significant role. It was designated as the lead for nuclear effects testing.

It turns out there are some forty-one nuclear effects facilities that made the dollar threshold cut. Of those forty-one, just to give you an example of the positive impact of reliance, a decision was made to shut, over time, seventeen of those forty-one because of duplication and downsizing. One of those is the Trestle facility out at Albuquerque.

B: That will make a lot of fire wood. [The trestle was a large raised platform at Kirtland AFB, made without metal, used primarily to test aircraft for electromagnetic pulse (EMP) effects as would be generated by nuclear detonations.]

A: What were the options? If we look at areas for possible consolidation candidates across the services—this is intuitively obvious, but it illustrates the point—there is no argument about who should test tanks or who should test submarines and surface ships. The only area of significant duplication has to do with aviation, aircraft, the weapons used on airplanes, and the weapons that shoot at airplanes. That's the area of overlap, and as a consequence, there is duplication among the services in test infrastructure and facilities. The repair depots are dealing with the same issues—excess capacity for Air Force and Navy fixed-wing aircraft. There is also some overlap in the laboratory infrastructure.

The largest single capital investment in the test infrastructure over the last several years and continuing into the future years is in the aircraft electronic combat testing facilities. That's where the major generic investments are being made in indoor facilities as well as outdoor simulators. The most contentious issue relating to aircraft that always shows up, of course, is high-performance aircraft flight testing. The issue of whether it should be consolidated in the southwest somewhere; it is not an argument of Pax River versus Edwards per se.

H: Blue skies versus gray skies, open areas versus urbanized.

A: If you consider the airspace and ground impact footprints, VFR [visual flight rules] weather requirements for developmental flight testing, low-level flight testing over varied terrain, and urban encroachment, you have to conclude that the Edwards/China Lake complex is the better alternative. It must be done in a way that preserve Navy equities. There was an article last year in the *Washington Post* which discussed the growth of the greater Washington-Baltimore metropolitan complex. The southeastern boundary of that complex is the Pax River Naval Air Station. There is population encroachment and there are also technology issues, one of which has to do with the kinds of airplanes we're building, both the Air Force and the Navy. Next-generation Navy platforms are fundamentally attack aircraft for land targets. That will require weapons separation testing, and you are going to fly over varied topography. That has to be done out west somewhere.

H: In fact, it's funny because whenever you test in an urban environment—and this is why the Navy left Anacostia in 1943 and went to Pax River—you always have the danger of accidents in a heavily populated area. Now that's not the kind of thing that's not going to happen in the southwest. You're not going to have a guy on China Lake or up on R-2508 going off and trashing a local community.

B: Maybe some old wild burro might get hit out in the desert.

A: That's the high-performance aircraft issue. The Air Force and NASA both moved high-performance testing there years ago. If the Navy's high-performance aircraft testing moves, a test pilot school move would logically follow. I think the Air Force and the Navy agree today that it could easily consolidate most elements of the test pilot training; they just don't agree on where it needs to be done. I believe that the move is inevitable, but it has to be done in a way that preserves the Navy's equity. The way you do that is you build the needed infrastructure and base the Navy air at China Lake.

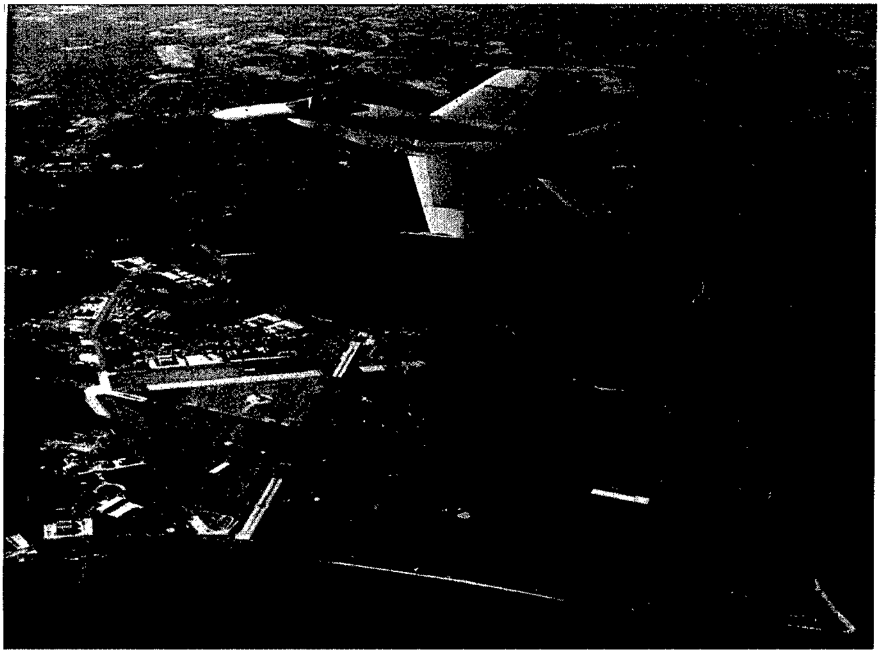
That doesn't mean closing Pax River by any stretch of the imagination. It still would accomplish the ASW [antisubmarine warfare] and the carrier suitability testing, and Pax River is evolving into a premier aircraft test support laboratory with the ACETEF [Air Combat Environment T&E Facility] and EMI [electromagnetic interference] facilities. The bulk of the weapons work and threats—the Echo Range—are out there at China Lake anyhow. Many of the activities relating to today's weapons designs are already there.

Another area is the electronic warfare range test support area. The issue there is one of cost. With the open-air threats, we can't afford to continue to upgrade numerous sites because we are talking fifty to a hundred million dollars for a simulator. There's going to be consolidation of open-air EW facilities. In the arena of the related indoor facilities, we cannot afford to have unnecessary duplication.

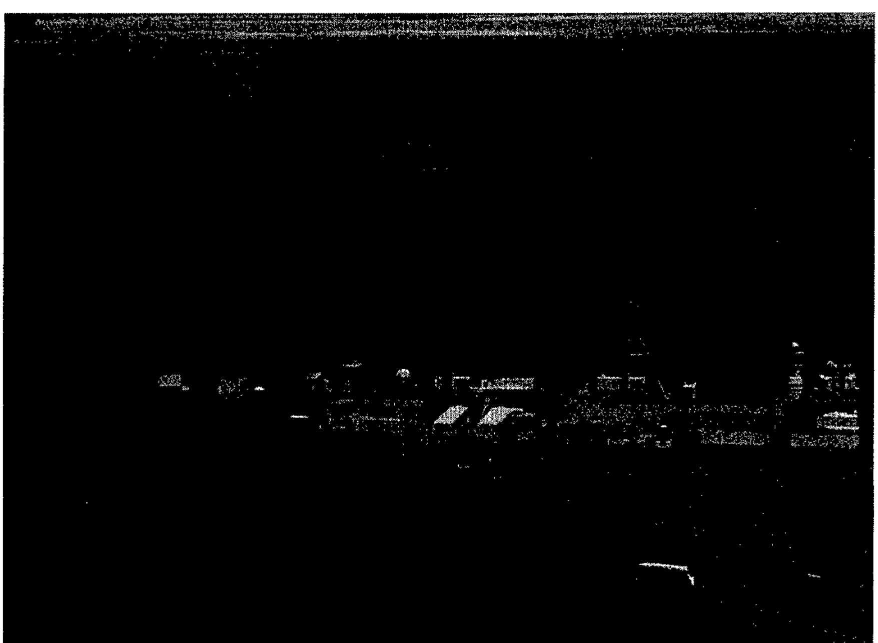
In the weapons test area, the White Sands complex has the best physical facilities and instrumentation for surface-to-air testing. Air-to-air testing is done at a variety of places: Eglin, White Sands, China Lake, and Point Mugu. I don't see much consolidation there because what goes on at each of those places is somewhat unique.* There is high capital investment in facilities and there's little to be gained by consolidation. In the air-to-air area I don't envision any big changes in the near term. By the way, White Sands has the best endgame scoring instrumentation for air-to-air testing as well. China Lake has state-of-the-art fuse test facilities. For air-to-ground testing, again, the same three locations. Because of smart submunitions, the best instrumentation for endgame scoring has evolved at White Sands brought about by the BAT submunition. It is comprised of upgrades to the multiple object tracking radars and upgrades to optical equipment. There is also a suite of instrumentation operated by DNA at the Stallion Range to test hardened targets. Testing conventional munitions against hardened targets has become more important since tactical nukes are no longer an option. There are facilities to do that both at Eglin and out at White Sands. I don't envision any significant consolidation [of these].

So the major area of overlap is related to aviation.

*Among the base realignment and closure actions of 1995 was the transfer of most of the electromagnetic test environment (EMTE) threat simulator system from Eglin to Nellis.



McDonnell Douglas F/A–18 Hornets over Patuxent River Naval Air Station, Maryland *(top)*, and China Lake Naval Air Station's Electronic Range (C-Site One) in California *(bottom)*.



Another issue is the amount of infrastructure either dedicated to T&E or used to support T&E. The T&E reliance effort consisted of a review of the following areas: threat simulators, airborne instrumentation, climatic test facilities, air breathing engines, test support aircraft, land vehicles, chem-bio effects, fixed-wing aircraft, surface-to-air, air-to-surface, and surface-to-surface missiles, nuclear weapons effects, targets, electronic warfare, eclectic guns, and supersonic sled tracks. In each one of these areas, there now exists a memorandum of agreement which basically defines an investment strategy and a lead service. All capital investments of a million dollars or over in new facilities in a given reliance test area have to be coordinated. The objective is to cut down on future investments and unnecessary duplication, so the impact will be greatest in the future. The impact on consolidating existing facilities has, with some notable exceptions, been very limited. The consolidation study provided the information that led to the closure, as an example, of the Naval Air Propulsion Test Center at Trenton [New Jersey].

In summary reliance was not designed to close major test facilities; reliance was designed to set up a process by which future capital investments in test facilities would be managed. That process is now in existence. Its effectiveness will depend on the people implementing the processes. To go back to the EC test process analogy, there was a process defined, but part of the problem was that the process was not followed. In retrospect I'm a bit disappointed by what was accomplished as a result of the reliance effort. I think the lead service concept is a step in the right direction, but that remains to be seen, as I said earlier, as to how effective it will be. In retrospect you can't accomplish much by getting people and expect them to go back home with the message that "I made a recommendation that we close our facility and consolidate."

H: That's a very difficult thing.

A: You can accomplish very little from the field with a bottom-up process when trying to consolidate. We set up a lead mechanism; it's a small step. Again, in retrospect, a tremendous amount of effort was put into reliance by a lot of people. The cost effectiveness of the project probably wasn't very high.

Turning to related issues, the T&E infrastructure activities fall into three main categories. There's the test and evaluation itself, which includes planning, test conduct, analysis, and reporting. These functions must, in my view, remain with the parent service acquiring the system. There must be an in-house government cadre within each service that knows how to execute and evaluate developmental and operational tests. You can't have that done by an outside agency.

Another major function of all of the test facilities is test support, which includes operating and maintaining the ranges, running the computers, data processing, and the real-time mission control activities. That function needs government oversight, but can and, in my view, should be contracted out to a large extent. The test support function is a mixed bag across the major ranges. On one end of the spectrum is an almost total contract operation, Vandenberg [AFB, California] and Tullahoma [site of Arnold AFB, Tennessee] being two examples. The other end of the spectrum would be almost a total in-house government civil service/military operation. In my mind, most of the test support functions should be contracted out for maximum efficiency and cost-effectiveness.

The third function that test facilities currently perform is investment management. This includes the functions associated with planning, budgeting, and implementing infrastructure upgrades. If you're going to control duplication and enhance interoperability, it's essential to have centralized management of the major infrastructure upgrades. That's one issue that I'm convinced is in need of a major change and is naturally being resisted by the service field activities. I've recommended that a joint program office be formed that would do the following: (1) function as the acquisition office for instrumentation, targets, and threat simulators; (2) manage RDT&E support platforms investment (that is, aircraft and ships); (3) manage installed system and hardware-in-the-loop development and procurement; (4) manage depot maintenance of common range assets; (5) develop and maintain a range interlinking master plan, and (6) provide instrumentation support to training.

A final function is to develop a modeling and simulation architecture to support T&E. I believe the joint program office should support the development of this modeling and simulation architecture.*

Interagency Testing

H: Which brings us up to an even broader if less-structured topic, and that is interagency testing. We've touched on this also in previous interviews, and that is the relationship that we've seen DOD have with other federal agencies, like NASA or the Department of Energy, Federal Aviation Administration, Department of Commerce, and CIA [Central Intelligence Agency].

*See also appendix I.

A: Let's start with NASA because that's the easy one and one with which I have the most familiarity, and I think it has been the most productive. There's always been a close working relationship in aircraft testing and the technology areas between NASA and the Air Force. If you look at the pictures here on your wall—the X-1, the X-2, and the X-3—the old NACA-Air Force relationship was superb.

H: A partnership?

A: Yes. It was very, very close and continues to this day with the AFTI-16 [advanced fighter technology integration demonstrator adapted from an F-16 airframe]. The AFTI-111 [F-111 technology demonstrator], which evaluated the mission adaptive wings, was not as productive as the F-16 AFTI. There were a number of those programs, joint ventures to evaluate aviation technologies. There has been a very close working relationship both on the design of the experiment and on the support infrastructure between the Air Force and NASA-Dryden Center [at Edwards AFB]. It has worked very well [and is expanding]. I'm not conversant with the level of cooperation in space-related R&D.

There hasn't been a comparable relationship with the FAA. Very few issues have been worked on a joint basis. One that comes to mind is in the area of hydroplaning during wet runway landings. The FAA had done quite a bit of work but not on a high-performance aircraft before, so we worked the joint program with the FAA. We also did one—when I say "we," I'm talking about the Air Force. Edwards supported a program on sonic boom research, providing the platforms, and we generally worked the flight profiles with NASA and the FAA. As I recall, NASA was the lead. There have been some commercial joint certification programs with the FAA, but there was never the relationship between the FAA and the DOD test community that exists with NASA.

N: One thought on interagency testing. One of the things that we have in DOD and we take it largely for granted is the intellectual property of the test and evaluation process. I recall someone giving a speech on that, and he basically said within the FAA they don't have a test culture. They don't have a process, and it seems like in many cases we could help transfer some of that.

H: That's a very good point.

B: That's sort of what happened with the NEXRAD [next-generation weather radar], where the Departments of Commerce and Transportation came to DOD to ask them to lead the operational tests of the new Doppler radar that we now see the imagery from on television every night.*

A: I'll follow up on that. The FAA has a very limited involvement in commercial aircraft testing, unlike the military services, which do a comprehensive and complete evaluation. In the commercial world, all of the developmental testing is done by a contractor. The FAA is responsible for the certification program; most of that is executed by the contractor with FAA oversight. It's fundamentally an oversight process, in contrast with the combined/joint test process typical of DOD.

H: I was struck by this going over some National Transportation Safety Board reports on some aircraft accidents. The whole thrust of the reports in some cases implied that the certification process had been too casual in a couple of particular cases. You realized just reading it how, in the commercial aviation world, we rely upon the contractors in that certification process. In a way, it's like the military services in the 1940s when they would leave so much in the hands of the contractor until the very last minute. It's not encouraging.

Test Management Oversight

H: A problem that we've seen increasingly since the 1950s, particularly as you've been dealing with programs that have major national significance or visibility, is the burden of test management oversight. More and more we've seen organizations that traditionally were not actively part of the test community positioning themselves to play an increasingly prominent role in testing, for example, congressionally mandated test objectives for the C–17 to meet. What is your view on dealing with that particular issue?

A: The basic answer is too much micromanagement. Examples include congressional hooks on the C–17, the B–2, and the ASPJ and other high-interest programs. There were twenty-some fiscal year 1993 congressional hooks in the B–2 program that had to be addressed to get funds released for fiscal year 1994. Most of those hooks have to do with flight-test results. The Secretary of Defense had to certify, in excruciating detail, that all of these hooks have been met before the funds could be released. I think that level of oversight is totally inappropriate.

*Throughout the 1980s, AFOTEC led a tri-agency operational test team that included members of the Department of Commerce's National Weather Service as well as the Department of Transportation's FAA.

There was a congressionally imposed requirement on the AMRAAM [AIM-120 advanced medium-range air-to-air missile]; I don't recall the exact words, but in effect the Director of Operational Testing had to certify that the AMRAAM met all suitability and effectiveness criteria. Most systems don't meet all the suitability and effectiveness criteria. The whole review process within the Pentagon is to make judgment calls about the adequacy of a system. What will happen over time with this congressional micromanagement is—and I've already seen the beginnings of this in some of the next-generation platforms—people will tend to water down design criteria to make sure that each goal can be met in its entirety. When you do that, you're not pushing technology. There is far too much micromanagement on the part of Congress—and some elements of OSD. The services certainly say that there is too much micromanagement on the part of OSD. There are too many agencies who are involved in the oversight process right now; we need to cut back. There are too many overseers and too few doers.

The role of OSD should be to assess a few macro-level measures and to make judgments about those measures. The whole issue of making technical judgments is discarded when you have inflexible legislative requirements that lawyers must interpret. Technical development is about setting a very high standard and trying to push technology. After you get some insight as to what is cost effective, senior decision makers collectively reach a decision about what constitutes acceptable performance. It's the classic "knee of the curve" cost-effectiveness judgment call. That was true ten years ago, and I think it's going to be even more important in today's environment where there is a relatively ill-defined threat and many different scenarios. What we are really trying to do is to keep pushing technology. People need the latitude to make judgment calls on how much is enough. That's what the entire review process is about. You need to keep setting very high standards during the initial requirements definition process. Decision makers have to continue to revisit those requirements as the system evolves. When technical wickets are translated into legislative requirements (that is, laws), you don't have any flexibility whatsoever.

H: Very quickly it becomes a judicial matter.

A: It does, absolutely.

B: I've got a couple of natural follow-up questions while you are talking about Congress. You earlier mentioned the need for integrating testing between the contractors and the government, but you also alluded to legislation that Congress passed in 1986 that doesn't allow the operational testers to have anything to do with the system contractors. What kind of progress and attempts has OSD been trying to make to get relief from that legislation over the last several years? Do you think there is any hope?

A: The so-called Section 800 panel effort looked at acquisition reform. One small part of that effort is called a streamline test statute. Elements of the proposed statute include: (a) allowing more flexibility in the use of contractor-acquired test data to satisfy operational test requirements and (b) permitting the use of contractor support where appropriate during testing. Your guess is as good as mine on the chances for approval of the proposed changes. We've gone too far in the direction of excluding the contractor from the process. The contractor has to be involved and must have access to the data. We've talked about integration laboratories and models and simulations. You either develop two completely independent sets of tools, which isn't cost effective, or you develop one and allow both the government and the contractor to use it. I think the operational test legislation was ill conceived. We've moved even further in the direction of building walls between the user and the developer. I would like to see that trend reversed; but in my view we won't get back to where we were in the late seventies in the foreseeable future.*

Electronic Combat Testing

H: Reflecting advances in technology, one of our greatest challenges has been electronic combat [EC] testing. This, it seems to me in the mid-eighties, first became highlighted by the problems we had with the B–1B on defensive avionics. This led General Larry Welch, who was then the Chief of Staff of the Air Force (CSAF), to say that the EC testing process was broken and that we needed a major effort to make it more systematic and predictable—more like flight testing. What do you see from your vantage point as the major trends in the evolution of EC testing? For example, recent progress, lessons learned that were incorporated in the systems that went into Desert Storm, lessons perhaps out of Desert Storm itself, and prospects for the future.

A: We now have a very well-defined EC test process. The big challenge is to put enough discipline in the system to ensure that the

*On 22 March 1994 Coleen Preston, Deputy Under Secretary of Defense for Acquisition Reform, appeared before a subcommittee of the Senate Governmental Affairs Committee to seek legislative relief from the artificial segregation of developmental and operational test data, the prohibition on using contractor personnel to support both OT&E and DT&E, and requirements for wasteful live fire tests. See *Defense Issues* 9, no. 26, 1994, 1–4. As of early 1996, however, no action had yet resulted.

process that has been defined is really used. That process consists of building brassboards and breadboards [readily modified prototype equipment], using indoor facilities to assist in the development and evaluation of the system at the subsystem level and then, finally, at the system level before you actually put the system in an aircraft for flight-test purposes.

You alluded to the problems with the ALQ-161 on the B-1, and there were a number of problems.* There was no integrated total system simulation available. The ALQ-161 was developed by the AIL division of Eaton in Long Island. They had no ground-based simulation facility that was used in conjunction with the offensive avionics suite, which was developed by Boeing in Seattle. That was one dimension of the problem. Another dimension was that some of the existing simulation capabilities weren't used in the development, notably the REDCAP [real-time electromagnetic digitally controlled analyzer-processor] facility at Buffalo, New York. There wasn't a disciplined use of the available tools to develop a system and troubleshoot it before it got in flight test.

The EC test process that the Air Force has subsequently defined uses indoor facilities in conjunction with flight test in a more systematic fashion. The process has been well defined and the tools are available. It remains to be seen as to the extent to which they will be used.

Let me talk a little bit about the ASPJ [airborne self-protection jammer].**

One of the many problems with the ASPJ is that EC systems don't enjoy the high priority that platforms enjoy, so their development is subject to perturbations in the budget process. As a consequence, they become stretched out. The impact of technology changes on EC systems is much more dramatic than they are in any of the other areas. One of the problems that the ASPJ had was the measures of effectiveness used to evaluate performance. The system had to demonstrate that it was significantly better than its predecessor Navy system, the ALQ-126. For some of the scenarios

*Critical deficiencies in the B-1B's ECM system did not become apparent until AFOTEC published a preliminary operational readiness assessment in August 1986, just one month before SAC declared initial operational capability as called for in the schedule announced when the Reagan administration made the production decision in 1981.

**Although the Navy had to cancel the ASPJ program in 1990 after a spate of criticism and bad publicity concerning its test program, the vulnerability of unprotected fighters over hostile territory (demonstrated by the shootdown of Capt. Scott O'Grady's F-16 by Bosnian Serbs in 1995) led the rebirth of the ASPJ program.

flown, it was physically impossible to show the degree of improvement required by the operational test criterion.

A problem we have with the testing of EW systems or electronic countermeasures [ECM] systems is our inability to define clear measures of effectiveness in measuring the performance. When you are developing a platform, that's relatively straightforward, payload, weight, range, speed, acceleration, the classical measures. When you are developing a hard-kill weapon, you can define a CEP [circular error probable] or a level of lethality. It's relatively straightforward to define and usually fairly easy to measure in the development and test process. EW systems are generally designed to "buy time" for a weapon system; to get past a threat rather than destroy it. As a consequence, it is extremely difficult to define measures of merit. Systems effectiveness is far more scenario dependent than many of the other systems.

This is Doug's area of expertise, so I would like him to expand on that a little bit. By the way, there was a recent excellent concept briefing developed by [Dr.] Marion Williams [AFOTEC Technical Director and Chief Scientist] on this issue. It had to do with the process and was a very structured approach. It is a significant step in the right direction.

N: We've talked about the B-1. I did an independent technical review of the B-1 ECM system some time back, a test of the program that at that time was foreseen. The basic question, Was adequate testing projected? When I looked at it and looked at everything that they then proposed to do—and that's been three, maybe four years ago—if they had done everything that they planned to do, it seemed like the test program was adequate because they were then planning to take it into an anechoic chamber and ream the system out end to end, do all those things that the EC test process now embodies.* The EC test process is very thorough, it's very scientific, and it's very affordable probably in the long run—in the sense that if we can avoid these problems we have seen in the past, in the sum total it will have been very affordable. To simply go out and fly hundreds, literally, of test missions in the air and expect to be able to come back and find the problems in these very complex, integrated, interrelated systems is just not going to cut the mustard anymore.

H: It sounds like storming through a haystack trying to find the needle that's in there someplace; if your process is random, you're probably not going to discover it.

*Budgetary pressures led to cancellation of original plans to fully test the B-1B in the Benefield Anechoic Chamber at Edwards.

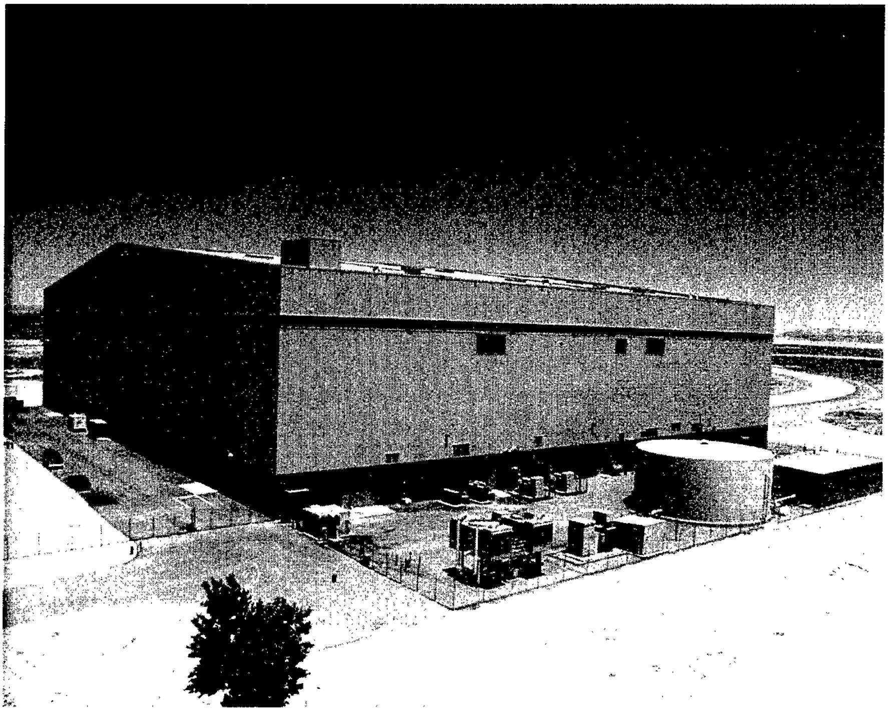
N: That's a good analogy. One of the problems that we found in this area more so than probably any other that I've been associated with is collecting too many data points in the same region. In other words, we were on a flat part of the curve and we didn't even realize it. Perhaps we didn't have the rigor in our analysis to even realize it if we had been able to, which we were able to, but we just didn't do it. It's an area where you really need to do the modeling and the simulation in advance and incorporate that into the hardware in the loop and the anechoic chamber type tests to smartly run through those test points and learn what you really need to know.

H: What that would do is actually build in a false confidence, if you will. You were accumulating all this data and not seeing any real problems, and it was all tracking fairly closely, but it is repetition of data coming within the same points of reference, if you will.

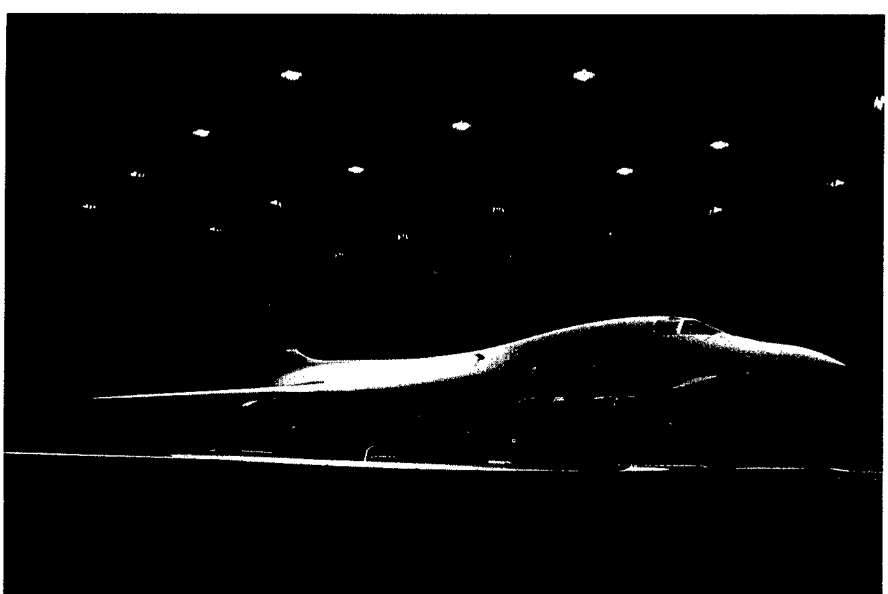
N: Right. As an example—and I'll stay away from the exact numbers—on ASPJ, it performed nominally well against certain threat simulators up until end threats. It was almost as if when we hit N plus one that was the breaking point. We didn't know that until we just happened to be able to put up a fairly dense environment during a flight test and we saw that. That's the kind of thing that you could have found easily in the lab test program.

A: To pick up on that point—the issue of threat density—there is no way we could replicate in a flight-test scenario the densities that might exist if we will still have the numbers of the threats in the former Soviet integrated air defense sector. There is certainly a continuing place for flight test, but with EC systems, we simply must use indoor tools and facilities to much greater extent than we have in the past. Part of the problem is that we look upon these facilities as test tools and they are to a degree, but they are also design tools. So we've got to get the contractor in there early on with a breadboard system. The government has to make them available to contractors for design refinements and developmental use. In some cases these facilities, because of the high capital investment required, are either run by the government or in a GOCO fashion, in the case of the REDCAP facility at Calspan-Buffalo and the AFEWES [Air Force electronic warfare evaluation simulator] facility at Lockheed-Fort Worth.* We have to make them available to the designer and must develop an attitude that these are design as well as test tools. The designer needs access to these facilities to develop the system.

*The 1995 Base Realignment and Closure Commission decided to disestablish REDCAP (relocating necessary equipment to Edwards AFB) but keep AFEWES in Fort Worth (overruling DOD's recommendation to move it to Edwards).



Benefield Anechoic Facility at Edwards AFB *(top)*, with B–1B inside *(bottom)*.



H: There was a real-world example in 1972 during Linebacker II when we started operating B–52Gs briefly against North Vietnam's air defense network. We found, largely because of the threat density question, the self-protection system that they had was completely inadequate to handle the SA–2 threat, so the B–52Gs suffered such disproportionate losses from the SA–2 that they were taken out of the Linebacker II lineup and the war was continued through its closure with Ds and Es and Fs. It was kind of a classic example of the system having gotten all the way into operational service and having indeed been in operational service for quite some time before the full magnitude of its weaknesses were known. I throw that out for what it's worth.

A: That's a good example. You just can't replicate that in a field test environment. You need to replicate the density somehow in the development process.

N: That brings up an interesting question. If it was the G-model, then that was the more advanced model and yet it had less capability.

H: That's right. That is one thing that has always struck me. The G-model—which one would expect to have the latest state of the art capable of handling a far greater degree and sophistication of threats than the earlier models—actually turned out to be quite the opposite. That's one of the great ironies, I think.

N: And on the B–52 they had the luxury of electronic warfare officers (EWO), one or more, which could use all their intellect, with just a few tools at their disposal. That's surprising, unless they had gone to automation.

H: No, they did not. They actually had an EWO in the airplane. That was very, very disconcerting, to say the least.

Software Testing and Human Factors

B: Speaking of automation, that's a good lead-in to software testing.

H: It sure is because we haven't really touched so far on the whole issue of software. Pete, you and I were talking quite a while ago about lines of code and the fact that one of the critical problems you have in software development is the people and number of lines of code that they are able to write each day. Looking at the lines of code issue as it applies to aircraft, the F–4 had no lines of code in it, the F–15A, I think, had sixty thousand lines of code, and the F–16E has something like 2.4 million, if memory serves me right.

Since they are increasing almost exponentially both in weapons systems and in C^4 [command, control, communications, and computer systems], could you comment on the evolution of software testing and some of the challenges that it has posed?

A: We were talking about the challenges of the avionics area earlier. Most aircraft subsystems today have some programmable capability and there are lots of examples. One example I like to use is fuel control. When I first got involved in performance flight testing, the fuel control was a hydromechanical device. If you wanted to change the fuel schedule, it took several months because it meant changing cams or orifices and that was something that the manufacturer had to do at its home facility. It took a significant amount of time to change the nonsteady state operating line on an engine to get better transient characteristics. Today, where there is the digital engine electronic control (DEEC), you can literally do those kinds of changes overnight. You've got a lot more flexibility, but at the same time you've got a lot more possible combinations to evaluate.

This gets back to the point you made earlier; we need some discipline in the process. You need to do as much of that evaluation as possible in a controlled environment before you actually put a software-intensive subsystem on an airplane. There are a number of dimensions. One is the reliability issue and another is regression testing. We put together a small team to look at this subject about four or five years ago. Two issues were raised by program managers and program executive officers in the services: How much regression testing is needed when a system is upgraded to reevaluate areas you had tested earlier? And how do you define a production-representative system with the software-intensive system? If you look at the OFP [operational flight program] software in an F–15, F–16, or F–18, they are making a block change a year; sometimes there is more than one.

H: Where do you freeze it?

A: Yes, where do you freeze it? The truth is it's an evolutionary process to a much greater degree than with hardware and the old hydromechanical systems, where you had rods, cables, and pulleys going to flight controls and hydraulic systems.

H: You were literally mechanically locked into that system.

A: And change was very difficult to come by, so that in itself was a powerful driver toward getting so-called configuration management. Configuration management discipline with software-intensive systems is much more important; it needs a lot more attention. That's one dimension of it on the individual subsystem. We have evolved from, say, an early F-16 to the F-16C and D and probably the F-15E where we had what were called federated systems. You had a data bus and there were some things that were common and you were sharing aerodynamic data with various subsystems, but for the most part you were still dealing with what I would call federated systems.

With the F-22, for the first time you move into—"integrated systems" is the buzzword. It drives the need for an overall architecture. A central air data computer, as an example, has the capability to process information for a flight control system, for the fire control system, the engine, the EW system, or to share this processing capability among several of the systems.

There is the architecture issue. If you will, it's distributed data processing. Doug can think of a more elegant word for it, but that's what it really amounts to. You are doing processing on a couple of relatively large computers on board, but you are distributing a lot of these functions and so the whole architecture of the design needs to be driven by building the flexibility in to accommodate all of the subsystems. And it has to be done in a modular way because the way the technology and capabilities are moving, you need to be able to go in there on a modular fashion every couple or three years with an upgrade. The design task is much more complex and much more of an integrated function.

As an example, today's flight controls—if you put enough control power and responsiveness in the surfaces, you can make the airplane respond to any reasonable control model. You do all that electronically and provide the requisite redundancy and failure mode effects analysis. There is a new crowd of engineering disciplines who have to play in the design process. The technical requirements have to be translated into a software architecture to make sure that you are implementing it in a way to provide the requisite redundancy. If you go from a federated system where you do your own failure mode effects analysis at the subsystem level and then you hook it into some bus somewhere, you've got to do a failure mode effects analysis on the entire system. If you're sharing a sensor among several functions or sharing an air data computer and you're processing data for another subsystem using that processor, say fire control, and you make an upgrade to the fire control data, what does that do to the central processor? The kinds of skills you need to develop and test a platform today have changed because we've gone to software-intensive systems.

H: That's funny in a way because this is all done in the promise of making a system more reliable, more useful to you, and it does that, but the price comes at greatly increased requirements for analytical performance in the sense of analyzing and understanding what's going on.

A: The F–22 PIO [pilot induced oscillation] is one example. On the issue of PIO and stick force gradients, that entire database is on a very specialized population of males; we don't know that it won't change with women. It may not, but the database is lacking. I think the interesting thing is that as much as we know about PIO—and I can remember in the late sixties we used to have test workshops on PIO and limit cycle testing. There was a lot of emphasis on ways to ground test high gain and variable gain control systems to try to ensure that you didn't end up having PIO. Again, surety, safety, and reliability testing before you get airborne. Despite all of the sophistication and all of those powerful analytical and laboratory tools, there were PIO problems with the shuttle, the F–22, and the Swedes had PIO problems with the Gripen [Saab JAS-39].

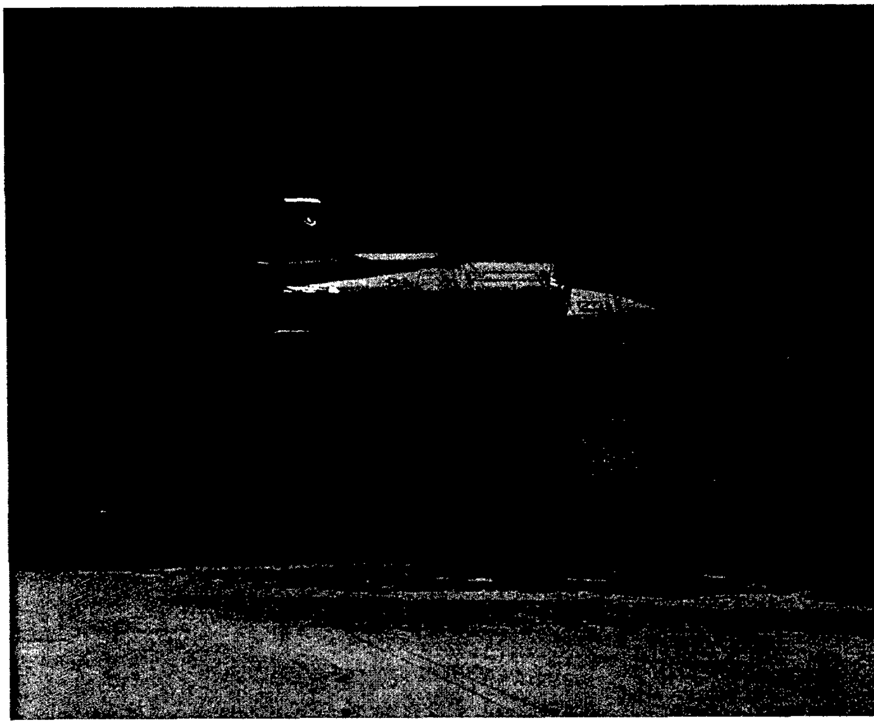
H: This is very interesting. There are some problems that you can narrow down but you never seem to get away from. Base drag is one, in the aerodynamic sense. Base drag has always been a problem in terms of evaluating it and measuring it. I hadn't thought of it until you mentioned it, but the PIO really is one of these enduring problems.

A: Regarding the human interface, this gets back to another issue with software-intensive cockpit display technology and information synthesis. We need to take advantage of the tremendous computational capability to decrease the cockpit workload, the so-called pilot's associate programs. Today's technology might be able to eliminate the need for a second person in strike aircraft, with tremendous savings. That's just one of the technical challenges for the JAST development.

H: It's not just the person in the cockpit and the weight of the person and all that, but it's the whole pipeline that produces that person and that supports them.

A: This issue of human response is an area where our database is really somewhat limited. As an example, the problem of G-induced loss of consciousness. I'm getting off on a tangent, but although that problem can be solved technically, the impediment is a cultural issue. It's similar to the cultural issue ten or fifteen years ago when aviators were against installing angle-of-attack limiting systems.

H: That's very interesting. It sounds like it's a little bit off on a tangent, but actually in many ways it's one of those so-called cultural issues that have a major impact on the technology.



YF–22 advanced technology fighter.

N: That's right. I think the fact that you have so much technology now compounds the problem. If you get into a philosophical design question, you could surely design a system that would require no pilot and that would be fully automatic, using all the sensors and data-processing algorithms, but you do require keeping a man in the cockpit, so then you end up with a balance question. How much automatic is enough? And then, given that I have certain automatic controls, what is the cut off point that I allow manual override? On the one hand, you're able to go very automatic, but you are self-limiting just by the culture that you always have to preserve the ability to override. How do you balance that out?

H: There are probably two extremes, and then there's the whole range in the middle. The two extremes are to man the airplane completely, and then on the other extreme a completely automated system, the remotely piloted fighter or whatever one cares to call it. Then in-between you have a whole range of options, many of which would be fighters of increasing complexity operating autonomous systems, namely sophisticated air-to-air weapons. It would seem to me that the further over you get toward an automated missile system, the requirement for the platform to be a highly agile maneuvering fighter itself declines.

To use the example of a current fighter that has a very long-range air-to-air weapon, presuming an F-14 was going after an antishipping missile fired with its Phoenix missile, the F-14 does not need to exploit its complete maneuver envelope to position itself to fire that weapon. The weapon basically does it for you because of the geometry of the intercept and the geometry of the positioning of the two systems.

On this software testing, I think we've really gone through it as we intended. One question we had talked about was proper baselining, and you discussed when to freeze the system and then the dilemma that correcting one glitch oftentimes causes another.

A: It's the old issue of regression testing. Before we close out, another issue I mentioned is that the program manager has to deal with what constitutes a production-representative system? That sometimes becomes an issue and the antiquated way we do operational testing. It became an issue on the AMRAAM, with some of the so-called production-representative shots. Ideally, you would like to do things during the latter stages of developmental testing on a production-representative system. During development on software-intensive systems, there are frequent software changes for a wide variety of reasons. It is a system optimization process. How you define production representative with software-intensive systems becomes an issue. Because of the ease with which you can change things, in the old days—going back to my fuel control analogy that I started with—you might have 200 hours of data or 100 hours of data with one fuel control configuration on an earlier generation hydromechanical system. Today you may be tweaking and fine-tuning the control on every flight, so you have a relatively small database on a given configuration.

That whole issue of the size of database and how one goes about defining production representative is relatively arbitrary in the sense that somebody says after the fact that we made a change and we're going to freeze the configuration. That defines for that block of time what constitutes a production representative. You don't have a large base of developmental data on that configuration prior to the time you make that decision. Our notion of how we go about things has to change to accommodate today's reality, the technologies we are dealing with today. Unfortunately, bureaucratic change is a couple orders of magnitude behind technological advance, but I guess that's always been the case.

H: That's always the case, I'm afraid. And I'm not saying that with the notion that one has to learn to live to accept that. That's a battle one has to continually wage, but the bureaucracy, as you know, is a lot more difficult to move than the technology.

Post-Cold-War Implications

A: There's a related issue that I touched upon earlier and I want to amplify it a bit, as well as the stratification and artificial separation of development from operational testing. The types of systems that have been developed over the last several years, and will be more so in the future, do not lend themselves to the highly segmented, protracted test programs that had once been the norm for major weapon system acquisition.

First, one of the key milestone decisions related to test results has to do with going beyond low-rate production. The concept of going beyond low-rate production for the majority of the systems we are dealing with today is not meaningful.

H: You will never have a high-rate F-22.

A: That's right. The whole process will be geared to some minimum level of production that sustains an industrial base and still has some measure of efficiency. That concept of going beyond low rate is certainly by far the exception rather than the rule. That's one point.

The second point is illustrated by two programs that we're familiar with, the F-15 and the F-16, both of which first flew, at least in prototype form, back in the early seventies. There has been continuous testing going on those two aircraft since they first flew. There has been a continuous process of what is now referred to as a multistage improvement program. Within each one of the major block upgrades, there may be several subblocks, so it's not just block sixty or seventy. There are combinations and permutations in an ongoing process. Within each one of those there are OFP software upgrades that go on almost annually as well.

The whole concept of going through a protracted sequential test program on a platform, and then the program ends, is no longer appropriate. It's a continuous process today, and we've got to find ways of doing that process more efficiently, not only from the test execution point but from a developmental point of view as well. The acquisition review process is geared to a model that I don't want to say no longer exists, but it is the exception rather than the rule. It will probably be applicable to something like a C-17, although even there will be software upgrades that come along, but certainly not for weapon systems where you are constantly upgrading attack or strike weapons or sensors.

H: You might find that model appropriate for a smaller article that you would be using in large numbers, like, for example, an interservice munition, but not for a major vehicle itself.

A: For the platform itself, it is inappropriate; we need to restructure and rethink the process. I mentioned, and that's what keyed me on this, that bureaucracy always moves slower than technology.

H: Previously we were talking about the fact that the low-altitude environment is now pretty much off limits to the attacker because of the threats that you encounter there.

A: A man-portable Stinger, yes.

H: Eventually, the defenders are going to have the capability to go to medium altitude. What this is going to do is force attackers (at least nonstealthy ones) to start operating further back with precision standoff weaponry.

A: The notion of close air support in other than a permissive environment, at least for a [fixed-wing] manned platform, is changing. Close support will be accomplished with unmanned precision guided weapons launched from standoff platforms.

H: And the definition of close air support may be a helicopter gunship over a ridge line, firing on something down here, using a lot of terrain masking. The idea of it being an F-16 that's going to come in—you're going to just lose an airplane.

A: I saw this in the early nineties when we went to White Sands and watched them shoot man-portable Stingers. You've got a $100,000 weapon that can hold a $50-million or a $100-million airplane at risk—a high-leverage weapon.

H: The low-cost threat forcing a very high cost in accomplishing defense. Somehow you think about a $50-million or $100-million airplane and it's flying around dispensing flares, and that's the best you can come up with to try to defeat this? Then you think, what happens with a very high contrast seeker?

A: And speed is not the answer.

H: And increasing the agility isn't.

A: The fast mover is more at risk than the helicopter because the helicopter can really use terrain masking. [At White Sands] they hit the F-100 drones with impunity.

H: It's absolutely scary.

A: It's going to change the nature of air warfare.

H: And, of course, that changes the nature of test. I think the difficulty has not been in testing to see if something works. It has been in relating development testing to the operational world. The scenario is critical.

A: Earlier we talked about the importance of an iterative process to better define and hone requirements and then use test measures of effectiveness that are really related to requirements. We are no longer dealing with a large, well-defined monolithic threat. In the past, everybody had a Fulda Gap scenario* in their models and their simulations. Today, we need a lot more flexibility to fine-tune requirements. In constructing a wide variety of potential regional scenarios, there is also the issue of surgical strike—and not only the platforms and weapons but the off-board and onboard sensors to provide targeting and command and control related information. It's a much more complex test scenario, and testing is only one dimension of it.

It also makes a much more complicated developmental program. How do you evaluate a spectrum of potential threats and then carve out from that spectrum a system that has flexibility to meet them? There is a need to totally rethink the process from one end of the spectrum—weapons of mass destruction and to the other end—a limited group of terrorists.

H: A classic example is Rwanda-Burundi. One guy with a surface-to-air missile changed the entire equation in that part of central Africa.**

A: We've talked about the test scenarios and the models and the simulations you used to hone requirements and test scenarios. There is a need to use these tools as new situations arise throughout the world. The Defense Department has begun to do this with the regional scenarios. It's a new way of approaching acquisition. We are not trying to keep up technologically with a very formidable, well-defined opponent. What we're trying to do is bring technology

*The expected invasion route of Warsaw Pact forces into the southern part of West Germany.

**After a plane carrying the presidents of Rwanda and Burundi was shot down near Kigali, the capital of Rwanda, on 6 April 1994, militant Rwandan Hutus began a massacre of the country's minority Tutsi population.

to bear in a wide variety of situations, most of which would be limited in duration and of low or moderate intensity.

H: That's really kind of sad if you look at something like the B–2, and you realize the capabilities it has, but in all honesty the expectation of the B–2 actually going to war is pretty minimal. As much as I admire what was achieved, it may be our last manned bomber—our last big bomber.

B: Like much of what we've discussed, something from another era.

Glossary

ACC	Air Combat Command
ACETEF	Air Combat Environment Test and Evaluation Facility
ACTD	advanced concept technology demonstrator
AEDC	Arnold Engineering Development Center
AETC	Air Education and Training Command
AFB	Air Force Base
AFEWES	Air Force electronic warfare evaluation simulator
AFFTC	Air Force Flight Test Center
AFFTIS	Air Force Flight Test Information System
AFI	Air Force Instruction
AFLC	Air Force Logistics Command
AFMC	Air Force Materiel Command
AFOTEC	Air Force Operational Test and Evaluation Center
AFR	Air Force Regulation
AFSARC	Air Force Systems Acquisition Review Council
AFSC	Air Force Systems Command
AFTEC	Air Force Test and Evaluation Center
AFTI	advanced fighter technology integration
AGM	air-to-ground missile
AMRAAM	advanced medium-range air-to-air missile
AOA	angle of attack
APGC	Air Proving Ground Command
AQ	Acquisition
ARDC	Air Research and Development Command
ARPA	Advanced Research Projects Agency
ASD	Aeronautical Systems Division
ASPJ	airborne self-protection jammer
ASTOVL	advanced short takeoff and vertical landing
ASW	antisubmarine warfare

ATF	advanced tactical fighter
ATS	automated test system
AWACS	airborne warning and control system
BAT	brilliant antitank (submunition)
C^4	command, control, communications, and computers
CAIS	common airborne instrumentation system
CEP	circular error probable
CFD	computational fluid dynamics
CIA	Central Intelligence Agency
COG	center of gravity
COR	Continental Operating Range
CSAF	Chief of Staff, Air Force
CTEIP	central test and evaluation investment program
DAB	Defense Acquisition Board
DARPA	Defense Advanced Research Projects Agency
DCS	Deputy Chief of Staff
DEEC	digital engine electronic control
DEM/VAL	demonstration and validation
DIA	Defense Intelligence Agency
DMR	Defense Management Review
DNA	Defense Nuclear Agency
DOD	Department of Defense
DOT&E	Director of Operational Test and Evaluation
DSARC	Defense Systems Acquisition Review Council
DT	development test
DT&E	development test and evaluation
DTESG	Defense Test and Evaluation Steering Group
EC	electronic combat
ECM	electronic countermeasures
ECTC	electronic combat test capability

EMI	electromagnetic interference
EMP	electromagnetic pulse
EMTE	electromagnetic test environment
EWO	electronic warfare officer
FAA	Federal Aviation Administration
GD	General Dynamics
GOCO	government-owned, contractor-operated
GPS	global positioning system
HAMOTS	high accuracy multiple object tracking system
HIMAT	highly maneuverable advanced technology
IADS	integrated air defense system
ICBM	intercontinental ballistic missile
IFAST	integration facility for avionics systems test
IFDAPS	Integrated Flight Data Acquisition and Processing System
IG	inspector general
IOT&E	initial operational test and evaluation
IRBM	intermediate range ballistic missile
ITTS	instrumentation, targets, and threat simulators
IWSM	integrated weapon systems management
JADO	joint air defense operations
JAST	joint advanced strike technology
JCS	Joint Chiefs of Staff
JEZ	joint engagement zone
JPO	Joint Program Office
JSTARS	joint surveillance target attack radar system
MAJCOM	major command
MRTFB	master range and test facility base
NACA	National Advisory Committee for Aeronautics

NASA	National Aeronautics and Space Administration
NATO	North Atlantic Treaty Organization
NAVAIR	Naval Air Systems Command
NEXRAD	next-generation weather radar
OER	officer evaluation report
OFP	operational flight program
OSD	Office of the Secretary of Defense
OT	operational testing
OTA	Operational Test Agency
OT&E	operational test and evaluation
PCM	precision countermeasures
PIO	pilot induced oscillation
QA	quality assurance
RAMS	RATSCAT advanced measurement system
RATSCAT	radar target scatter
R&D	research and development
RDT&E	research, development, test, and evaluation
REDCAP	real-time electromagnetic digitally controlled analyzer-processor
SAC	Strategic Air Command
SAE	service acquisition executive
SAF	Secretary of the Air Force
SDIO	Strategic Defense Initiative Organization
SFW	sensor-fused weapon
SPO	system program office
STOL	short takeoff and landing
TAC	Tactical Air Command
TFX	tactical fighter experimental
TSSAM	tri-service stand-off attack missile

TVA	Tennessee Valley Authority
UTTR	Utah Test and Training Range
VFR	visual flight rules
WPAFB	Wright-Patterson Air Force Base
WSPO	Weapon System Project Office

Appendices

Selected Papers by Charles E. "Pete" Adolph

Appendix A

"Minimizing Flight Test Time and Cost in the United States Air Force"

Presented to the
North Atlantic Treaty Organization (NATO)
Advisory Group for Aerospace Research
and Development (AGARD),
Conference Proceedings No. 424 (June 1988)

MINIMIZING DEVELOPMENT FLIGHT TEST TIME AND COST IN THE U.S. AIR FORCE

Charles E. Adolph
Technical Director
Air Force Flight Test Center
Edwards AFB, California 93523-5000

SUMMARY

Flight testing has undergone some major changes in the past 20 years. The largest single technical change, the need to evaluate software-intensive systems, resulted from advances in computer technology. Test management concepts have changed as well. These changes were driven in part by technology and in part by a need for the Air Force to become more involved early in the test process. Today's avionics systems present both a quantum leap in capability and a quantum jump in test requirements. With workload growing both in magnitude and complexity, the challenge is to meet the increased demands cost-effectively and safely. For software-intensive systems, a ground-based simulation dedicated to the support of the flight test program is essential. This paper summarizes today's methods of operation which are geared to minimizing development time and costs. The focus of this process is on productivity--doing the right testing in the proper sequence in the most efficient and safest possible manner.

INTRODUCTION

The United States Air Force Flight Test Center (AFFTC) conducts aircraft development flight tests. Nearly every new United States Air Force airplane in the past 40 years was tested at Edwards, as were NASA's high speed flight research vehicles. The Edwards Flight Test Range is used to support these flight test programs. The Center also operates the Utah Test and Training Range, the Air Force's largest overland range where remotely-piloted research and test vehicles, plus air and surface-launched missiles, are tested. Edwards Air Force Base today is the hub of a tri-service test complex which encompasses several inland and overwater ranges throughout the southwestern United States.

The last 30 years have seen dramatic changes in aircraft technology and in the tools we use in flight testing. Aircraft flight envelopes, however, have not changed significantly in the past three decades. Fighter aircraft were approaching Mach 2 and 50,000 feet 30 years ago. Today, with a few notable exceptions, we are still dealing with a 50,000 ft/Mach 2 envelope. However, there have been significant improvements in flying qualities, and in subsonic thrust and lift-limited envelopes. In the late 50's, the 1960's and the early 70's, there were numerous new aerodynamic designs. In the early 70's alone, first flights were made on the following aircraft: the F-15, F-16, F-17, A-9, A-10, YC-14, YC-15, and B-1. The 70's were the era of competitive fly offs as well (A-9 vs A-10, F-16 vs F-17, YC-14 vs YC-15, AGM-86 vs AGM-109). Until the T-46 made its first flight in 1985, the most recent first flight on a totally new Air Force aircraft was the B-1 which flew in 1974. For the past ten years, flight testing at the Air Force Flight Test Center has been confined to derivatives of existing aircraft, primarily in the area of avionics upgrades.

The most significant change, by far, in the last 30 years, is the burgeoning use of software-intensive systems. In the past five years in particular, the world of on-board computer technology has been moving in maximum afterburner.

Air Force developmental flight test management concepts have undergone a similar evolutionary process. Technology has been the catalyst for many of these changes. The flight test and evaluation programs conducted on U.S. Air Force aircraft have gravitated over the past two decades from what were largely independent, sequential test programs conducted by the contractor, Air Force development and operational test agencies, to programs conducted on a concurrent basis from a single test location, Edwards Air Force Base.

The increase in software testing has resulted in a quantum jump in workload. The best single measure of workload at the Air Force Flight Test Center is test flying hours. The test workload has increased dramatically in recent years. In the past 6 years, test flying hours tripled. In fiscal year (FY) 1980 about 1870 test hours were flown from Edwards. In FY 1986, the number was 6000. The largest contributors to the workload are the F-15 and F-16 fighters and B-1 bomber testing. The test flying hour estimate for FY 1990 is 11,600. This is firm, on-the-books workload. There has always been a significant amount of additional test work that is not identified years in advance. 1985 and 1986 were Edwards' most successful years in terms of flight safety. More test hours than ever were flown without a single Class A mishap. Class A mishaps are accidents involving the loss of a test aircraft, a fatality, or more than $500,000 worth of damage.

TEST PROGRAM MANAGEMENT

The cost and complexity of today's test aircraft and ground support equipment, test range data acquisition equipment, data processing systems and test support ("chase") aircraft, were major considerations in the consolidation of test activities. Another ingredient, the need for increased visibility by the customer during the development process, combined with the prohibitive expense of duplicative testing, led to what is referred to in the Department of Defense (DOD) as combined testing.

Air Force implementation of this DOD policy emphasizes consolidation of test events wherever practical. Data from each test event are made available to all appropriate agencies using Air Force Systems Command facilities and capabilities, including instrumentation, data processing and analysis systems, to the maximum extent practical.

The management concept used for today's Air Force flight test programs is referred to as a Combined Test Force. It is a sophisticated application of matrix management principles. The typical Combined Test Force is composed of participants from the development organization (contractors), buyer, using commands and supporting commands. The contractor contingent may involve a prime contractor and subcontractors or as in the case of the B-1, a group of associate contractors (airframe, engine, offensive and defensive avionics) with the government as the integrating contractor. A comprehensive treatment of Combined Test Force operations is contained in reference 1. Salient points are summarized below.

In the Combined Test Force approach, test activities are combined to the maximum practical extent.

a. Participants are housed in common facilities.

b. The test aircraft are instrumented to meet the needs of development testing (contractor and Air Force) and operational testing.

c. A single test plan is developed which integrates and prioritizes test requirements.

d. Test range requirements are identified.

e. Combined aircrews are used for most missions in multiple place aircraft. In single seat aircraft, either contractor or Air Force development pilots fly envelope expansion missions.

f. Aircraft are located at a single site and maintained by an integrated maintenance team.

g. All test data are included in a common data base which is available to all team members.

h. Analysis and reporting of the test results is accomplished by each organization independently.

The most significant advantage to combined testing is the opportunity for an early and continuous look at the product by both the developmental and operational military communities. There is no substitute for hands-on experience. Early involvement by users provides an opportunity to influence the design where appropriate to improve the mission capability of the aircraft or system. Early participation in the test effort by military pilots, engineers and maintenance personnel helps identify problems before the production cycle is too far along.

Another advantage of combined testing is the reduced time and cost. Combined testing virtually eliminates the duplication which existed in the past when each tester used his own facilities and completed his own test and evaluation with little or no input from the other testers. If all team members participate in the planning effort, agree with the test approach and instrumentation and data collection methods, there is no reason to duplicate tests.

Consolidating all test aircraft at a single location has distinct cost saving benefits. More flexibility in the use of the aircraft is possible if they are instrumented correctly. When one test series is delayed, another test can be scheduled on the same aircraft. As a result, fewer instrumented aircraft can accomplish the same amount of testing than is possible if the aircraft are at more than one test location. Less support equipment is required, which is particularly significant when ground support equipment is scarce.

For military test programs the location should be a government facility in most instances because hangar, office and laboratory space will support successive programs with minimum investment after the original outlay. Making capital investments in facilities that can be used again is more cost effective than paying for contractor facilities that may only be used for a single program.

Facilities and equipment such as instrumented ranges, mission control rooms, data reduction facilities and weapon delivery ranges will also support several programs at the same time. Range facilities necessary for evaluation of fighter and bomber aircraft such as air-to-air and air-to-ground weapon delivery ranges, low-level and supersonic routes, electronic combat ranges, and adequate restricted airspace are only available at government facilities. The disadvantages of shared use of a test facility is that data reduction equipment, telemetry, and ranges must be shared. This can create scheduling conflicts, but these problems are manageable.

Conducting combined tests at a government facility also offers the potential advantage of oversight by an experienced flight test management team. The Flight Test Center applies the expertise gained from managing a variety of weapon system test efforts to improve test effectiveness and safety.

SAFETY - TEST RISK REDUCTION

There are two fundamental objectives on any test program--to conduct tests efficiently and safely. The balance of this paper addresses test efficiency. It is worthwhile to focus briefly on safety as well. The Air Force development test safety record has improved dramatically over the years. More test hours were flown in 1985 and 1986 at Edwards than in any previous year, without the loss of a test aircraft. Figure 1 summarizes the fighter aircraft record over the years. It is worthwhile to explore the reasons for the improvement in safety.

There are basically two reasons: technology and management procedures. Telemetry gives test personnel the ability to monitor critical parameters in real time. But monitoring isn't enough. The system must be designed to minimize recognition time, to identify the proper corrective action, and to initiate the action. Recognition time is minimized by prominently displaying limit exceedances of critical parameters. The proper corrective actions must be defined in advance with the test conductor given the responsibility for notifying the pilot immediately. The above describes the real time element. Of more importance is the up-front planning process.

Several years ago, after several accidents and near-accidents, a decision was made at the Flight Test Center to establish a separate safety organization. The objective was to create a small organization with some degree of independence from the test managers. The organization was and is staffed by experienced pilots and engineers who are on rotational assignments. They have current experience and have a guaranteed "return ticket" to their parent functional organization. Civilian engineers are given a temporary promotion. The combination of a rotational assignment and temporary promotion attracts highly qualified individuals.

Every test program undergoes a rigorous safety review by project personnel and senior supervisors which is chaired by people from the safety organization. The review system brings to bear all expertise, government and contractor. The review process ensures that critical conditions are approached incrementally, in small steps. The safety track record is significantly improved in comparison with the past because of a concerted effort to consider the entire system. With today's complex aircraft, there is the potential for interaction among subsystems. A systems approach is taken during the safety review by including people from a variety of test disciplines in the review process. As an example, propulsion and flying qualities experts are included in the review of gun firing tests. Secondly, people who have been involved in tests of a given type (e.g., flutter, high angle of attack) on a wide variety of aircraft are a part of the review process.

Tests are categorized as low risk, medium risk, or hazardous based on the severity of potential hazards and probability of occurrence. Examples of tests which have demonstrated higher than normal risk include first flights, flight envelope expansion, flutter tests, high angle of attack testing, rejected takeoffs, and tests with explosive warheads. Minimizing procedures to prevent a mishap from occurring or to reduce the consequences of a mishap are developed for each test hazard. All hazardous tests are thoroughly reviewed by the senior staff and Flight Test Center Commander prior to accomplishment.

TODAY'S DEVELOPMENT AND TEST CHALLENGE - AVIONIC SYSTEMS INTEGRATION

The emphasis of this paper thus far has been on test managment issues. The focus will now shift to the number one technical challenge facing the development and test community - avionic systems integration.

As was noted earlier, there has been a marked decrease in totally new acquisition programs in the last decade (figure 2). Weapon systems are becoming increasingly capable, complex, and costly, so it isn't surprising that there are fewer starts, and within each program, fewer units are bought each year. Upgrading fielded weapons systems is becoming the dominant means of force modernization (reference 2). This has been apparent to the development test community as the workload emphasis has shifted from air vehicle envelope expansion and airworthiness to subsystem upgrades.

We must get "out in front" of the subsystem development process, i.e., where possible, systems should be developed before they are incorporated during production. The problem, as noted in reference 2, is that full-scale development of electronic subsystems is often begun well after the full-scale development of the platform in which they will be incorporated. The recommended approach is to decouple the development of the aircraft from that of the critical subsystems, so as to allow the necessary "head start." This concept is being used by the Air Force, where possible, to develop subsystems in an earlier version of the airframe in which it will be incorporated (e.g., the F-15, F-16). The F-16 Advanced Fighter Technology program has also been used to evaluate an impressive array of the technologies for potential application to future weapons systems.

Cost-Effective Testing of Software-Intensive Systems

The software-intensive avionics of today present both a quantum leap in capability and a quantum jump in test requirements. These systems have made the test task more demanding and complex. A high percentage of the test work we accomplished on aircraft such as the F-15, F-16, and B-1 is tied to software. The challenge for the designer, tester and operator is to maximize the benefit from the flexibility afforded by software-intensive systems, while at the same time being cost-effective. Our experience over the past several years has shown that we tend to grossly underestimate the amount of flight testing necessary to fully develop software-intensive systems. The number of flights required to develop and evaluate a software-intensive system is often 2 to 4 times higher than the initial estimates.

Old-time flight testers will usually support the notion that flight test is the ultimate proof that a system does indeed perform as designed. But with software-intensive systems, substantial savings can be realized by troubleshooting and resolving avionics problems on the ground rather than in flight. The Integration Facility for Avionic Systems Testing at Edwards AFB has validated the concept.

Edwards AFB Avionics Test Facility

The avionics simulation capability at the AFFTC resides in the Integration Facility for Avionic Systems Testing (IFAST). The IFAST is an avionics test facility housed in a three-story building containing four shielded test bays and a time-shared central computer complex. The facility is designed to meet the biggest challenge in avionics DT&E: how to satisfy increasingly complex test requirements more efficiently. The facility is being used to support the major test programs at Edwards including the B-1, F-15, and F-16. The following is a list of a few of the applications of this facility in support of these programs:

- Avionics component functional integration and checkout
- Mission software preparation
- Mission profile simulation prior to flight
- Airborne test instrumentation checkout
- System maturation before flight testing
- Crew test procedure preparation and verification
- Test mission sensitivity studies and subsystem simulations
- Technical order validation and verification

The facility is currently operating at capacity. Construction of an addition will begin later this year.

Historically, the AFFTC has maintained separate facilities for flight control system/flying qualities simulation. Because of today's highly integrated systems, these avionics and flight control development and test activities have tended to blend together. The AFFTC plans a gradual merger of these simulation facilities.

Ground Facility Cost Savings

To evaluate the cost savings potential of a ground test facility, an assessment must be made of the software problems that will remain undetected until the test phase. Studies and actual experience have shown that there are a reasonably predictable number of faults that are likely to occur in a system with a given software capacity. The issue of importance and interest for us as testers is the percentage of faults that will remain undetected until the subsystem is ready to undergo testing. A rigorous procedure for estimating the number of software errors is contained in reference 3.

Using a ground-based avionics test system which uses sophisticated simulation techniques to exercise the flight software and then analyze the results, we may test this software more efficiently and greatly reduce the flight schedule. Based on data from a number of programs, we would expect to discover 80 percent of the software errors on the ground.

Costs associated with typical current generation programs are depicted in figure 3. The cost savings associated with decreasing the number of flights by identifying problems on the ground is shown in figure 4. Calculations are made in figure 4 based on an assumed savings of 20 aircraft months (see reference 3).

The cost savings and schedule compression are dramatic. In our illustration, the cost savings factor is about 20 to 1. This is a very conservative estimate, and is based on our experience with today's fighter aircraft. Other estimates are as high as 100 to 1 (reference 4). The bottom line is that the test managers must insure that adequate funds and personnel resources are invested in the simulation in time to have it up and operating to support the test program.

It is not surprising that the problems associated with the developmental flight testing of software-intensive military systems are common to the development of current generation commercial aircraft. References 5 and 6 document the Boeing Commercial Airplane Company's experience during the development of the 757 and 767. The references also address the issue of growth in memory requirements for commercial applications. The average growth from contract award to certification was a factor in excess of 2.0. Several of the summary statements from the Boeing experience are repeated below because they are directly applicable to our recent experience.[5]

> "Hardware should provide adequate reserve capacity for inevitable growth in software."

> "Changes are a way of life in digital avionics--must be allowed for in program schedules and plans from outset."

> "Simulators/simulations are absolutely essential to development of avionics equipment--for early definition of requirements, testing of design concepts and validation of final designs. Simulations must be started early, maintained particularly early in flight test, and be as representative of airplane/engine dynamics as practically possible at all times."

The conclusion is obvious: for software-intensive systems a ground-based simulation dedicated to the support of the flight test program is essential. The cost-savings potential mandates that the simulation planning be given a level of attention comparable to the flight test planning. The use of the simulation must be planned for early on and integrated into the flight test program, data reduction and spares support. The simulation must be as representative of the airplane as is practical.

The payoffs from the use of a simulation facility to support avionics testing are summarized below:

- Three-fourths of software problems are resolvable on the ground at a fraction of the cost of flight.
- A reduction in test flying hours which translates into a reduction of test costs and an acceleration of test schedules.
- The costly and inefficient fly-fix-fly approach is minimized.
- Ground testing is more efficient because the experiment is controlled, repetition of test conditions is rapid, simple.
- Flight test time is used more effectively by isolating/keying on risk areas and smarter profile planning.
- Flight test safety for digital control systems is enhanced.

CONCLUSION

In conclusion, these are exciting times at Edwards Air Force Base. Test workload has grown dramatically in the past six years and is projected to increase in the future. Combined testing has proven to be a viable test management concept. It eliminates duplication, reduces test time and cost, and provides for earlier military participation. Aircraft flight test emphasis has shifted over the past ten years from airworthiness/aerodynamics testing to avionics subsystem test and integration. Advances in weapons system technology have had other significant impacts on the test process. Today's test aircraft require the use of a ground-based simulation which is dedicated to the support of the flight test program. Flight test challenges of the future include the need for continued improvement in avionic systems test efficiency, and increased emphasis on the use of simulators and other ground test facilities to supplement flight testing.

REFERENCES

1. Kennedy, Peter D.; Combined Test Force Operations at the Air Force Flight Test Center, AIAA Paper 86-9739, AIAA/SFTE Third Flight Testing Conference, April 2-4, 1986.

2. Rich, Michael and Dews, Edmund; Improving the Military Acquisition Process, Lessons From Rand Research, Rand Publication R-3373-AF/RC, February 1986.

3. Adolph, Charles and Montgomery, Phillip; Cost-Effective Testing of Software Software-Intensive Systems, SFTE Sixteenth Annual Symposium, July 29 - August 2, 1985.

4. Boehm, Barry W.; Software Engineering - As It Is, September 1979, TRW Report 55-79-03.

5. McElroy, John J.; The Engineering Flight Simulator: A Resource Essential to Flight Management System Development and Certification, IEEE Paper, 1983, CH 1839-0/83/0000-0020.

6. Sutcliffe, Peter L.; Lessons Learned in the Development of the 757/767 Flight Management System Equipment, IEEE Paper, 1983, CH 1839-0/83-0000-0107.

FIGHTER/ATTACK TEST AIRCRAFT

AIRCRAFT TYPE	TEST AIRCRAFT LOST	FIRST FLIGHT	TEST FLIGHTS	TEST HOURS
F-15	0	JUL 72	9,400*	13,250*
F-16	0	FEB 74	13,900*	17,800*
A-10	1	MAY 72	2,310**	3,480**
A-7	0			
F-4	1			
F-111	4	DEC 64		
1953 - 59***			INITIAL TEST PERIOD	
YF, F-100A	6	MAY 53	5/53 - 12/55	
F-101A	2	SEP 54	9/54 - LATE 56	
YF-, F-102A	2	OCT 53	10/53 - 5/55	
XF-, YF-, F-104A	11	FEB 54	2/54 - MID 58	
YF-, F-105B	1	OCT 55	10/55 - 3/60	
F-106A	2	DEC 56	12/56 - 6/59	

*AS OF JAN 87 **THROUGH 1977 ***FROM AFFTC HISTORIES, TEST-RELATED, SOME DID NOT OCCUR AT EDWARDS

FIGURE 1

NEW U.S. AIR FORCE FIGHTERS, 1940s-1990s[a,b]

DECADE	FIGHTERS DEVELOPED
1940s	P-47, P-51, P-59, P-61, F-80, F-84, F-86, F-89, F-94
1950s	F-100, F-101, F-102, F-104, F-105, F-106
1960s	F-4, F-111
1970s	F-15, F-16
1980s } 1990s }	ATF (DEVELOPMENT PLANNED)

[a]THIS FIGURE EXCLUDES FIGHTERS THAT ENTERED FULL-SCALE DEVELOPMENT BUT THAT WERE NOT PROCURED FOR INVENTORY; IT THEREFORE UNDERSTATES THE NUMBER OF NEW STARTS FUNDED IN THE 1940s AND 1950s.

[b]EXTRACTED FROM TABLE 2, REFERENCE 2.

FIGURE 2

TYPICAL TEST PROGRAM COST DATA

AIRCRAFT	TEST FLIGHT FREQUENCY (FLIGHTS/MONTH)	TEST AIRCRAFT OPERATIONS COST ($1000/HOUR)	FLIGHT TEST FIXED SUPPORT COST ($1000 MONTH)	PROGRAMMABLE CAPACITY (1000 WORDS)
CURRENT GENERATION FIGHTER	10	15	$1,500	300 ~ 700
CURRENT GENERATION BOMBER	5	50	$5,000	600 - 800
NEXT GENERATION FIGHTER	10	30	$2,500	1,000

FIGURE 3

COST SAVINGS - TYPICAL FIGHTER

COST OF FLIGHTS:	
(250 FLIGHTS) (1.25 HRS/FLT) ($15,000/HR)	= $ 4,687,500
LESS COST OF SIMULATION:	
(250 FLIGHTS) (1 HR/FLT) ($4,000/HR)	= $ 1,000,000
NET SAVINGS FROM REDUCED FLIGHTS:	= $ 3,687,500
(20 AIRCRAFT MONTHS) ($1,500,000/MONTH)	= $30,000,000
TOTAL SAVINGS: $3,687,500 + $30,000,000	= $33,687,500

FIGURE 4

Appendix B

"The Role of Test and Evaluation
in the Defense Acquisition Process"

Presented to the
Subcommittee on Procurement and Military Nuclear Systems,
Committee on Armed Services,
United States House of Representatives,
8 March 1989

Mr. Chairman and Members of the Committee:

I am Charles Adolph, the Deputy Director of Defense Research and Engineering for Test and Evaluation. I appreciate the opportunity to appear before you today to discuss the role of development test and evaluation in acquisition decision-making.

In preparation for today's hearing, the committee proposed five questions dealing with various aspects of acquisition decision making. At this time, I will provide a brief response to several of those questions. Dr. Duncan and Mr. Krings have already provided answers to some of the questions.

First, let me address the overall scope of Development Test and Evaluation (DT&E).

DT&E is an on-going process involved throughout most acquisition programs. DT&E helps to determine the readiness of technologies for application, to resolve design questions, and to measure progress of the design in overcoming development challenges. DT&E is of primary importance in conducting testing that determines if the development system meets the technical requirements of the contractual development specification. But, the scope of DT&E is much broader than technical specifications; fundamentally, DT&E, like OT&E, is concerned with the acquisition of quality systems. Therefore, we evaluate the results of our technical testing, not only against technical specifications, but also to provide information from which early insights regarding potential operational effectiveness and suitability may be drawn. Further, we share the results of development testing with Mr. Krings and his staff so that he can independently evaluate the results of technical testing in forming his early operational assessments.

My office establishes the overall policies for the conduct of DT&E and reviews the implementing policies and procedures issued by the Services. In addition, we oversee the DT&E on major as well as designated acquisition programs.

The principal oversight mechanisms we employ are:

* The Test and Evaluation Master Plan (TEMP)
* Field visits by myself and my staff to view selected testing
* System test reports from the development test and evaluation agencies within the Services
* and formal and informal briefings from the development test and evaluation agencies and the Service developer.

We participate in the preparation and conduct of each program Milestone review by the Defense Acquisition Board (DAB) structure. Before each milestone review, we hold a specific Test and Evaluation review, generally several days to a few weeks prior to the formal DAB Committee review.

For a production decision review, we ensure that virtually all DT&E on the basic system will have been completed; that any major deficiencies will have been identified, corrected, and retested. If the review is prior to a low-rate production decision, it is acceptable to have some DT&E incomplete. But for a full-rate production review, DT&E on the basic system should have been completed successfully.

At the Test and Evaluation briefing, we specifically address, with the developing Service, any areas where unresolved T&E concerns remain. For each Milestone, including the production decision, I report to the Defense Acquisition Board structure on the scope of the DT&E that has been accomplished and on the principal findings from that testing. Finally, as needed, the Director of Defense Research and Engineering represents any unresolved DT&E concerns at the formal Defense Acquisition Board review.

2. One of your questions was what is done if/when milestones are not met?

At each program milestone, e.g., entry into Full-Scale Development, approval for Full-Rate Production, etc., the developing agency is required to detail a variety of critical cost, schedule, and performance thresholds tied to the program milestones. Test thresholds are specified in greater detail in the Test and Evaluation Master Plan (TEMP), which is updated to ensure currency and completeness prior to each Milestone review.

The Test and Evaluation Master Plan details specific development objectives and performance thresholds (criteria) that must be achieved prior to entry into dedicated operational testing as well as the critical technical thresholds against which test progress will be measured at each future milestone.

The fact that a program threshold is exceeded, or a required threshold is not met at a program milestone, does not, in itself, necessarily warrant program cancellation or restructure. It does require a careful, senior level examination of the significance of the shortfall to the operational effectiveness, suitability, cost and schedule of the system. No arbitrary rigid criteria can, or should, be substituted for carefully reasoned judgments. I ensure that the scope of development testing and the results from that testing are fully considered as part of the DAB process.

3. Another question was what is being done to reduce risk to a minimum?

From the point of view of test and evaluation, the best tool we have for reducing risk is early, detailed test planning, execution, and reporting.

Thorough development testing requires early planning attention. The development tester has to understand the operational requirements, the threat, and the operational environments in which the resulting system will be used. The tester must understand the relationship between the required operational characteristics and the technical characteristics that are critical to achieving the operational performance. Finally, he must understand where the technical hurdles lie, for this provides focus for the testing. After this thorough understanding is gained, the development tester must then identify the resources he needs to conduct meaningful development testing. He must insure that the developing system and its interfaces to other elements are tested throughout its performance envelope, across the operational environments, and in the presence of appropriate surrogate threats. The final result is that thorough development testing takes resources -- funding, time, and qualified people. These will only be present when needed if thorough early planning is accomplished.

The key tool for ensuring this early attention to T&E planning is the Test and Evaluation Master Plan (TEMP). The TEMP is written by the developing Service. It must be approved jointly by Mr. Krings and myself before any testing can begin.

Clearly, one of the primary purposes of any testing is to determine if unexpected results are produced. Consequently, there is always some risk associated with any test involving high technology systems. Early planning, allocation of sufficient resources, and a structured component, subsystem, full-up system test approach are key to reducing development test risks to a minimum.

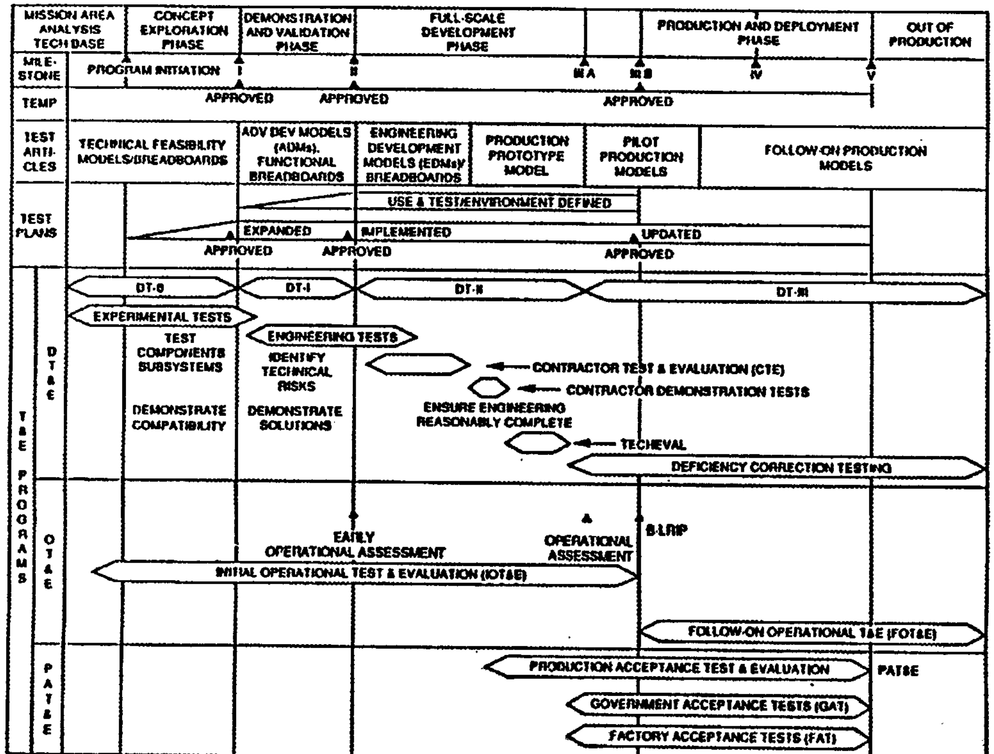
4. You also asked when/how can concurrency be justified?

It should be noted that some concurrency, with its attendant risks is the norm in most development programs. Some types of concurrency are necessary attributes in a well-structured program. For example, a program should not proceed serially from a development phase that focuses only on design, to a production phase that addresses producibility and cost, to a support phase that suffers the life-cycle consequences of the earlier phases. Instead, we expect a form of concurrency throughout each development phase -- a concurrency that integrates the concerns and considerations of future phases into the activities of the on-going phase.

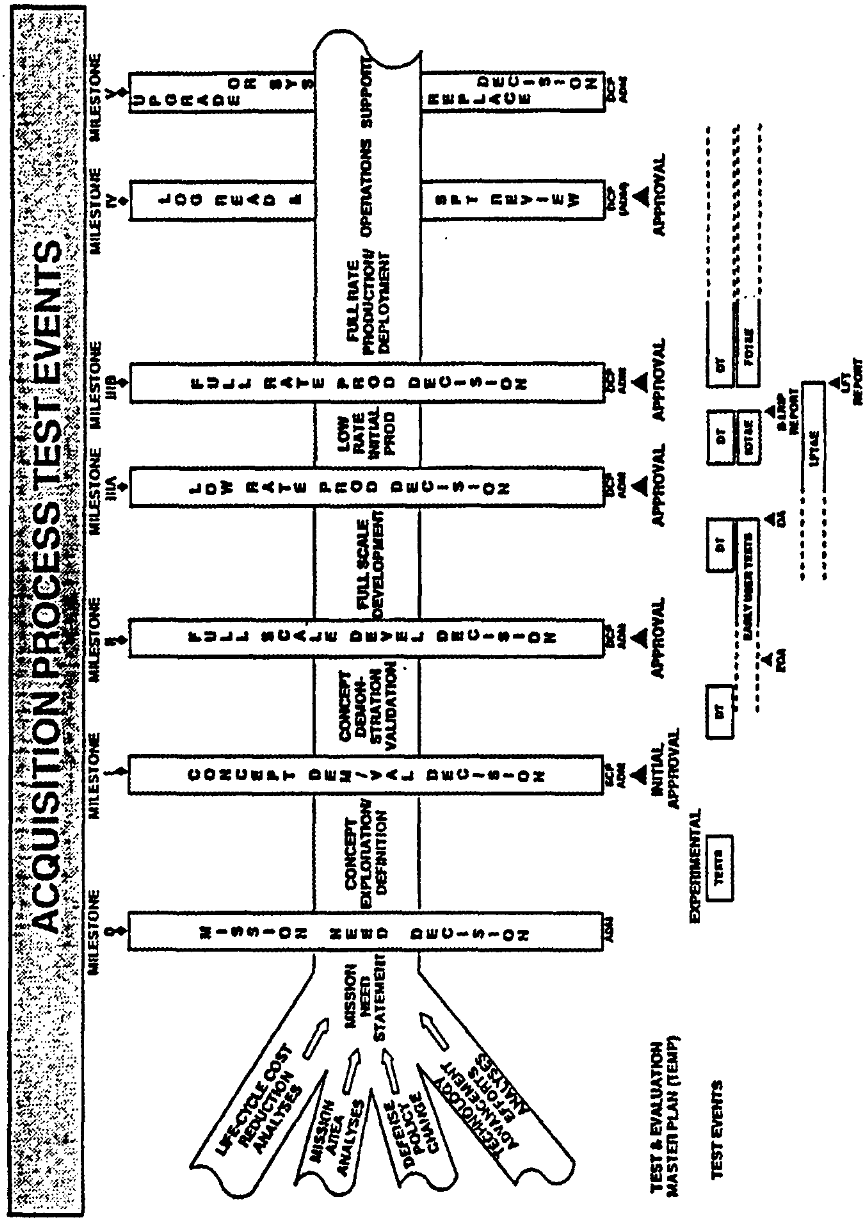
The type of concurrency that is of concern is that of entry into high rate production before activities normally associated with Full-Scale Development or Test and Evaluation on the basic system have been completed. The decision to proceed with a concurrent program of this type is an acquisition decision. Test considerations are only one aspect that must be considered in such a decision.

The primary factor that would justify approving and proceeding with a highly concurrent program is the urgency of the military requirement and the technological risks involved. However, I believe that the full scope of the issues that are involved are best and properly addressed from the perspective of the Under Secretary for Acquisition, rather than limiting it to a test perspective.

LIFE CYCLE EVENTS



9-8318


ACQUISITION PROCESS TEST EVENTS
MILESTONE 0
MILESTONE I
MILESTONE II
MILESTONE IIIA
MILESTONE IIIB
MILESTONE IV
MILESTONE V
MISSION NEED DECISION
CONCEPT DEM/VAL DECISION
FULL SCALE DEVEL DECISION
LOW RATE PROD DECISION
FULL RATE PROD DECISION
LOG READ & SPT REVIEW
UPGRADE OR SYS REPLACE DECISION
MISSION NEED STATEMENT
CONCEPT EXPLORATION/ DEFINITION
CONCEPT DEMON-STRATION VALIDATION
FULL SCALE DEVELOPMENT
LOW RATE INITIAL PROD
FULL RATE PRODUCTION/ DEPLOYMENT
OPERATIONS SUPPORT
LIFE-CYCLE COST REDUCTION ANALYSES
MISSION AREA ANALYSES
DEFENSE POLICY CHANGE
TECHNOLOGY ADVANCEMENT EFFORTS ANALYSES
INITIAL APPROVAL
APPROVAL
TEST & EVALUATION MASTER PLAN (TEMP)
TEST EVENTS
EXPERIMENTAL TESTS
DT
EARLY USER TESTS
IOT&E
FOT&E
LFT&E
LFT REPORT

Appendix C

"Statement on Defense Test and Evaluation"

Presented to the Subcommittee on
Research and Development of the Committee on Armed Services,
U. S. House of Representatives,
23 April 1991

Mr. Chairman,

I am Pete Adolph, the Deputy Director of Defense Research and Engineering (Test and Evaluation). I appreciate the opportunity to participate in this Defense Test and Evaluation Panel.

Test and evaluation supports the acquisition process by providing the capability to support weapon system and technology developments, to verify the achievement of performance thresholds, and to determine operational effectiveness and suitability. The T&E infrastructure must be responsive to the aggregate needs of the following communities: science and technology, research and development, weapons acquisition, operational test, and product improvement.

System Test and Evaluation (T&E) is a large activity involving almost 60,000 people, including contractor personnel at our T&E facilities. Our budgets aggregate to about six billion dollars per year when the costs reimbursed by our customers are considered. The funding is included in many program elements approved by this committee each year. Our facilities have a replacement cost of about 25 billion dollars. The land devoted to T&E activities encompasses about 14 million acres, which is over 50 percent of total DoD land area. The major T&E facilities are scattered throughout the United States and in the Atlantic and Pacific Ocean areas. This large enterprise exists primarily to support all phases of acquisition for U.S. weapon systems, although we also provide T&E services for other Executive Departments and Agencies such as NASA, DOE, and the Coast Guard. In addition, we perform a small amount of T&E for U.S. allies. In total, the DoD T&E establishment supports several thousand T&E projects each year.

My job is to provide developmental test oversight for the Department of Defense. My responsibilities include looking to the out-years to determine future requirements and ensuring that we have adequate facilities and equipment to satisfy future T&E needs. Our staff works with all three Services to ensure that the highest priority needs of the Department are addressed first and that duplication among the Services is avoided. Oversight extends to all major DoD T&E facilities, except nuclear weapons test facilities, and includes operation, maintenance, improvement, and modernization. It also includes targets and threat simulators for T&E, the Live Fire Test Program, and several auxiliary programs dealing with joint test and foreign comparative testing.

The T&E infrastructure must be capable of meeting current and future T&E challenges by enabling the DoD T&E community to assess the complex and evolutionary technologies being engineered into today's weapon systems. Our goals to accomplish this are summarized below.

- Responsive management of DoD-wide T&E Capability Base
- A secure, safe test environment
- Consistency and commonality of test methodology
- Improved test efficiency and effectiveness
- Transportable instrumentation
- Interoperability and commonality of instrumentation, targets, and threat systems
- Improved threat scenario definition
- An aggressive T&E technology development program
- Environmental clean-up and monitoring compliance
- Minimize impact from encroachment
- Foreign cooperation in T&E

I will discuss some of the above objectives in more detail after I outline the T&E oversight mechanism we now have in place. It's the Defense Test and Evaluation Steering Group (DTESG), a standing body that addresses corporate Defense T&E resource matters in response to Defense Management Review Decision (DMRD) 922. A similar steering group coordinates Defense laboratory programs. I am a member of that group as well so as to avoid unnecessary duplication of capabilities between the T&E and laboratory communities.

The DTESG also addresses T&E requirements that result from science and technology programs. Members include the Service senior T&E executives, who are here today, as well as members and advisors from other appropriate OSD staff, Defense Agencies, and the SDI program. We review the major issues related to oversight of the T&E establishment. Over the past three years this group and its predecessor organization have been involved in all major decisions effecting T&E resources and policy. One of the key activities of this group has been the approval of priorities for the Central T&E Investment Program, which is managed at the OSD level, but executed by the services.

A major DMRD effort is our initiative to prevent unnecessary duplication of T&E facilities. It designates a lead service or develops interservice cooperative agreements. This initiative has been dubbed the "T&E Reliance Process." (A similar process is underway as a part of the laboratory consolidation effort.) There are four broad T&E reliance categories: service-unique, lead-service, cooperative, and competitive. Numerous reliance areas have been identified for study. The list includes land vehicles, air breathing engines, chemical and biological effects testing, air-to-air, air-to-ground, and ground-to-air weapons testing, electronic guns, fixed-wing aircraft, and electronic warfare testing. The reliance process will also address support capabilities and facilities, such as ground and aerial targets, anechoic chambers, radar cross-section facilities, and T&E support aircraft.

By the end of 1991, joint OSD/Service teams will develop reliance arrangements for each of these areas. The Service Joint Commanders Group (Test and Evaluation) will make recommendations to the DTESG. Upon approval by the Director of Defense Research and Engineering, the Services will execute the agreed-to plan in each major area to implement new test facility and capability support arrangements. Budget adjustments and amendments resulting from the reliance studies will be folded into the FY 94/95 submits as appropriate. We plan to include proposed program-specific investments by program managers in our oversight process as well as generic improvement and modernization projects funded by the Services and OSD.

As you may recall, from 1985-87 we conducted an extensive analysis of our ability to support test and evaluation. What we saw caused us some concern. From 1980 to 1985, RDT&E funds for Advanced Development and Engineering Development, a category that drives T&E's near to mid-term requirements, had grown 90 percent in real terms while T&E funding had remained relatively flat. Our facilities had an average age of 35 years. Our investment rate in those facilities was about one-fourth that of private industry. In addition, Congress had become critical of T&E for our weapon systems. A large number of new or more complex technologies were projected to be incorporated into our weapon systems such as stealth, kinetic energy and directed energy, smart munitions, space systems, and complex electronic warfare systems. These new technologies needed to be reviewed to determine if we were posturing our facilities to adequately test them. Security upgrades were needed to protect technical data and satisfy special access program requirements.

Environmental monitoring and clean-up requirements were increasing and potential encroachments on land, sea, and air demanded increased vigilance. Finally, test facility maintenance and repair was falling behind in the areas of electrical power, building repairs, water distribution systems, road repairs, and other infrastructure upgrades.

As a result of the FY85-87 review, we modified our management approach. We established a joint OSD/Service review group to set priorities across the Services. The Joint Logistics Commanders were included in our planning process. We called our T&E and weapons development personnel together to identify future T&E requirements. We identified more than 15 billion dollars in new requirements over the next 15 years. These requirements were prioritized, and we built a program to support about one-half of these new investments. Much of this was incorporated into central OSD lines to fund high priority projects that would benefit more than one Service or projects that were too costly for a single Service to fund within its available resources.

Unfortunately, we were not able to persuade the Congress to approve these programs. These investments have been reduced to about 50% of our original plan. Congress reduced overall T&E funding about 16% in each of the past two years. We have experienced even more severe indirect cuts as a consequence of the yearly Congressional reductions, DoD makes negative adjustments to the out-year tail that goes along with a Congressional reduction in a current year. Over the past two years, the reduction to the FY90 to FY93 T&E program has amounted to $2 billion (about 16%) to T&E while the total RDT&E account - excluding SDI and T&E - was reduced only 1%.

The impact of these reductions will result in further aging of our facilities and more costly testing because we are not using the latest test support technology. Reductions also mean delays in the time when we can better test new weapon system technologies. This does not necessarily mean that we will not be able to test our weapon systems adequately, but it does mean we will not be able to test as effectively or efficiently. We solicit your support to fully fund our FY-92 budget request to reverse the trend set over the past few years.

The DoD infrastructure reductions planned for the next several years will not result in a significant decrease in T&E workload. Reductions in quantities of new weapon systems do not materially alter the amount of testing required. Even cancellations of new systems will not significantly decrease our T&E workload, because subsystem upgrades will be made to existing systems to prolong their utility. Upgrades fall into two basic categories, the first is one which avoids obsolescence through a technology upgrade. The second is a service-life extension program (SLEP), which may or may not include a technology upgrade. An example of an ongoing technology upgrade is the F-16C/D Multi-Staged Improvement Program (MSIP). Systematic upgrades are made to weaponry, communications, navigation, and defensive sensors in a staged approach. The program significantly changes the characteristics of the F-16 operational flight program software, as well as adding several hardware enhancements. By adding capability in stages, it allows potential improvements to be rigorously tested in development without delaying other capabilities that have successfully met the requirements.

Historically, weapon systems have evolved over time, and advances in technology may lead to Pre-Planned Product Improvements (P3I). These improvements vary from minor software efficiency changes to a major enhancement of the weapon system enhancements. Some of the new technical concepts, as well as the more technically challenging improvements will require prototyping prior to fielding.

Another challenge we face is how to accommodate environmental clean-up and monitoring requirements, including clean-up of contaminated sites, disposal of hazardous waste, clean air and water, endangered species, noise, wetlands, etc. Over the past 8 years, the number of pages of federal legislation in these areas has increased from about 1,000 to approximately 10,000 pages. State and local regulations have seen similar increases. Differences in regulations among the states limit our ability to develop common approaches for all our facilities, which are spread throughout the United States. Since T&E facilities occupy over half of DoD land, we are impacted heavily by the new emphasis on the environment. We own about half of the land but have a higher percentage of the environmental challenges because of our space-launch facilities and chemical/biological test facilities. All the exotic fuels and other materials show up first at test facilities. We strongly support initiatives in the environmental area and we are making every possible effort to program and budget for these requirements. We are taking a positive approach to addressing the myriad of environmental issues. Solutions will require intensive effort, patience, and an increase in funding. Unfortunately, Congress tends to level-fund our infrastructure accounts, which are a primary source of environmental compliance funds. We need the full cooperation of the Congress to meet the environmental challenge.

Let me move on to the subject of Test and Evaluation process improvements. Under Dr. Duncan's leadership, and with the cooperation of the Services and the OSD acquisition community, we have made significant improvements in the T&E process. The T&E process improvement initiative has focused on five major areas; (1) improving the stability of the test program planning, (2) improving the resolution of issues between the OSD T&E offices and the Services, (3) improved test facility and resource planning, (4) improving test methodology and finally, (5) improving communications between the OSD T&E offices and the Services. Dr. Duncan has addressed several of these. I will expand on two areas; threat definition which is a major factor in program stability and resource planning, and test methodology improvements.

One of the issues raised by several Program Executive Officers (PEO's) and Program Managers (PM's) was that the test community sometimes made a unilateral interpretation of the threat to be used for testing. New procedures have been implemented that provide for more systematic definition of the threat and interpretation of scenarios for test purposes. In the early stages of test plan development, a meeting will be held to define the threat for test purposes. A Defense Intelligence Agency (DIA) threat assessment will be the basis for threat scenario development. The meeting will address targets, countermeasures, scenarios, doctrine and tactics employment, and the degree to which models and simulations will be used. A significant new threat will be addressed in a program review called by either the service or OSD. Resource requirements for testing major new threats will also be addressed as necessary at Defense Acquisition Board program reviews. We are also dealing with the issue of evolving threats in our program reviews.

The OSD T&E offices and the DIA co-sponsored a conference in April to address how scenarios are used in system acquisition. The objective was to improve the connectivity in threat definition among acquisition documents (e.g., Mission Need Statement), System Threat Assessment Report (STAR), Cost and Operational Effectiveness Analysis (COEA), Test and Evaluation Master Test Plan, and Test Plans through the use of consistent scenario definitions. The conference was a definite step forward in an effort to reach closure in this area.

Now let me move on to the area of test methodology improvements. One of the major issues we will continue to address in the future is improvement in the efficiency and effectiveness of the test process. A high priority area is electronic warfare (EW) testing. An OSD/Service EW test capability study was recently initiated. The study will define the future EW T&E capabilities required to support the acquisition of major weapon systems within realistic fiscal constraints. Live fire and software testing were two areas recommended by program managers and PEO's for methodology and/or policy reviews. The Army's assessment methodology for combat vehicle vulnerability was reviewed in 1989 by the National Research Council. A similar review of aircraft live fire test methodology is in the planning stages, with completion anticipated in 1991. Two issues relating to software policy were identified. The first is a definition of what constitutes production representative software for operational test purposes. The second is what role the Service and OSD T&E communities should play in the frequent software upgrades that many systems undergo, sometimes annually. These two issues have been addressed by a Service/OSD team, and Dr. Duncan and I recently signed out draft software test guidelines for Service review.

The GAO recently assessed our efforts (GAO report B-241231, Sept 28, 1990) "Improvements are being made in the Department of Defense's Test Planning." The following are the Results in Brief from that report:

> DoD has focused a great deal of attention on improving test planning processes and has made several improvements in Test and Evaluation Master Plan (TEMP) guidance. More specifically, our review indicated the following:
>
> - The military services have each instituted similar TEMP and test planning processes in accordance with Office of the Secretary of Defense (OSD) guidance.
> - DoD has adequately identified the key weaknesses and strengths of the TEMP and test planning processes.
> - DoD actions address historical weaknesses and, if properly implemented, should significantly improve the timeliness and quality of TEMPs and the test planning processes.

In conclusion, the Department of Defense is well into the detailed planning to reduce the cost of operating T&E activities by consolidating future investments and areas of responsibility. Efforts are also under way to improve the efficiency of the test process. These combined initiatives are designed to posture the T&E community to better meet the challenges of the 1990's. We need the support of Congress to ensure that the required test facility upgrades are made in a timely manner.

Appendix D

"Statement on DOD Laboratories"

Presented to the Subcommittee on Defense
Industry and Technology of the Committee on Armed Services,
U.S. Senate,
21 May 1991

**DEPARTMENT OF DEFENSE
STATEMENT BY
CHARLES E. ADOLPH
TO THE
SUBCOMMITTEE ON DEFENSE INDUSTRY AND TECHNOLOGY
OF THE
COMMITTEE ON ARMED SERVICES
UNITED STATES SENATE
MAY 21, 1991**

Mr. Chairman, members of the committee, ladies and gentlemen. I am Pete Adolph. I have been designated by the Secretary of Defense to perform the duties of the Director, Defense Research and Engineering (DDR&E), until a Director is appointed. With me is Ray Siewert, Acting Deputy Director, Defense Research and Engineering, for Research and Advanced Technology. Also testifying this morning are George Singley, Deputy Assistant Secretary of the Army for Research and Technology; Genie McBurnett, Principal Deputy Assistant Secretary of the Navy for Research, Development, and Acquisition; and Brigadier General Pat Condon, Deputy Assistant Secretary of the Air Force for Management Policy and Program Integration. He is here today in his capacity as Acting Director of Science and Technology for the Air Force. I'm pleased to be here with my colleagues to testify on our current initiatives to strengthen the technology base, to streamline the management of the Department of Defense (DoD) laboratories, and to reduce the cost of conducting business.

Much of the work we are doing falls under the aegis of the Defense Management Review (DMR), which requires us to seek ways in which the Department of Defense can perform its vital national security mission more efficiently. We are making steady progress. I am pleased by the high level of coordination and cooperation between the Office of the Secretary of Defense (OSD) and the Services in working these initiatives. We are here together today to report on our progress and continue our dialogue with you.

As evidenced by the success of Desert Storm, the Department has been correct in basing its strategy for national defense on developing superior technology. We will continue our commitment to provide our soldiers, sailors, airmen, and marines with superior technology as our national security policy adapts to the dramatic changes taking place in the world. Our strategy for defense technology includes fostering those conditions essential to maintaining our qualitative superiority in deterrence and warfighting capabilities. As a result, we have a heightened need for a coherent, integrated, stable, and strong research, development, test, and evaluation (RDT&E) program; one that links technological leadership to military needs. A critically important ingredient is the DoD in-house laboratory system.

In response to the requirement for a more corporate, unified approach to management of RDT&E, the DDR&E has established two steering groups, one for test and evaluation (T&E), which I chair, and one for science and technology (S&T), which Ray Siewert chairs. These two steering groups will be able to advise the DDR&E in a more cohesive manner on resource and policy matters.

The laboratory system falls under the purview of the S&T Steering Group, which will advise the DDR&E on integrating Service and Defense Agency technology objectives, infrastructure problems, and investment strategies. I mention this steering group this morning because, on behalf of the DDR&E, it will oversee our major laboratory management initiatives, including laboratory restructuring, the Laboratory Demonstration Program, and Project Reliance.

I want to take a moment to describe the Tri-Service Reliance effort because it exemplifies the teamwork I talked about earlier. Although Reliance focuses on both S&T and T&E, I will concentrate my remarks on the S&T portion.

Reliance is a major DMR effort being conducted by the Services and OSD to enhance the effectiveness and productivity of our facility investments and in-house work, eliminate unwarranted duplication and unnecessary overlap, connect related technology efforts, and refocus investments to strengthen and modernize our facilities -- all the while preserving the Services' unique, mission-essential capabilities. The intent is to build on the strengths inherent in each Service's system for the benefit of everyone and to foster inter-departmental reliance when that makes sense. We have thoroughly reviewed our laboratories and identified various programs, first by assigning them to 28 technology areas and then to several operational categories: programs unique to one Service, coordinated programs, joint efforts, collocation, consolidation, and competition. Having done this, we are now reorganizing away from coordination -- the weakest and least productive alternative -- and toward the other options. The DDR&E reviews all Reliance strategy and implementation plans and makes recommendations to the Under Secretary of Defense for Acquisition. I will not elaborate further on Reliance since the Service representatives who spearheaded this tremendous and productive effort are here with me and will address specifics in their presentations.

The Defense Technology and Test Action Plan, which presents the specific actions and milestones necessary for the RDT&E community to effect the cost reductions and increases in efficiency directed by DMR 922, will be signed out soon by the Deputy Secretary of Defense. It includes actions and milestones for investment planning, for workload measures, and for implementation of Project Reliance decisions.

Consolidation, conversion, and realignment of the Defense laboratory system have been the subject of intense study in the Department and the Services since the summer of 1989. Stimulated by the recommendations of the 1988 Report of the Secretary's Commission on Base Realignment and Closure, the Secretary's Defense Management Review, and other reports such as those from the Office of Technology Assessment, the Department conducted expansive and systematic studies of the near- and long-term savings that we could achieve through consolidations and closure of Defense laboratories. The Deputy Secretary of Defense established the DMR Laboratory Consolidation Working Group in the fall of 1989 to develop recommendations. This OSD/Tri-Service working group studied management alternatives such as government-owned/contractor-operated facilities (GOCO's), employee-owned contractor operations, government corporations, and hybrids of various options in addition to closures, consolidations, and realignments. These studies were in the final stages when, in Senate Report 101-384 on the National Defense Authorization Act for Fiscal Year 1991, Congress directed the Secretary to establish a Commission on Laboratory Consolidation and Conversion with the charter to review the health and effectiveness of defense laboratories, using DMR studies as a starting point, and to make recommendations on the means to improve the operation of these laboratories.

This Federal Advisory Commission has initiated its review of ways to improve the operations of the laboratories.

In selecting the Commission members, the Secretary chose three DoD laboratory directors who had played an active role in the DMR studies, the Director of DARPA, executives from the Department of Energy and the National Institute of Standards and Technology, the vice presidents for research at Bell Labs and IBM, the director of Los Alamos Laboratory, and other equally qualified scientists and engineers. These individuals have brought an outstanding breadth of expertise and experience to review the Services' laboratory consolidation plans and other means of improving the operations of the Defense laboratories.

The Commission's first two meetings consisted primarily of information gathering and policy discussions. This is an important process. They have been briefed in detail by the Services on their consolidation plans and on Project Reliance. The base closure office in OSD has briefed the Commission on the Secretary's proposal to the Base Closure and Realignment Commission. They were briefed on the Laboratory Demonstration Program, and they have interviewed staff from the House Armed Services Committee and the Congressional Research Service to make sure they are accurately responding to the intent of Congress. Additionally, the Commission has invited a staff member of the Senate Armed Services Committee to meet with them later this week. The Commission members have a well-defined meeting schedule, are addressing their duties systematically, and fully expect to meet the mandated 30 September, 1991, reporting date to the Secretary of Defense.

Let me turn now to the Laboratory Demonstration Program. The Department of Defense laboratory system has an enviable record of technological accomplishments. Nevertheless, for too long bureaucratic impediments have hampered the Defense laboratories' ability to compete in the labor market for high-caliber scientists and engineers. The effectiveness of their research efforts has been reduced because of inadequate facilities and equipment and other management inefficiencies.

I want to briefly review some of what has transpired in the last two years. Through Senate Report 101-81 on the National Defense Authorization Act for Fiscal Year 1990 and 1991, Congress endorsed recommendations of the 1987 Defense Science Board study and directed the Department to implement a Laboratory Demonstration Program.

On November 20, 1989, Deputy Secretary of Defense Donald J. Atwood initiated this program. In his authorizing letter he stated, "Recent studies conducted by the Defense Science Board and validated by the DDR&E Interagency Task Force have shown that the productivity and effectiveness of DoD laboratories can be significantly improved by implementing specific changes in procedures involving personnel management, research related contracting, facilities refurbishment, and management authority of technical directors."

The Department has taken a five-level approach for 1) implementing changes at the laboratory level; 2) implementing changes that are within the authority of the Services and Defense Agencies; 3) implementing changes that are within the authority of DoD; 4) negotiating with other government agencies (e.g., the Office of Personnel Management) to secure increased authority within existing law; and 5) when no other recourse is available, seeking legislative changes.

Participation in the DMR Laboratory Demonstration Program was granted to those Defense laboratories nominated by the Services and the Defense Nuclear Agency and approved by the Under Secretary of Defense for Acquisition. Collectively, and prior to recently proposed restructuring plans, the Demonstration Laboratories employed approximately a quarter of DoD scientists and engineers (S&E's), including approximately 21 percent of Army S&E's, 31 percent of Navy S&E's, and 12 percent of Air Force S&E's. These laboratories represent a broad spectrum of the mission areas, science and technology strengths, laboratory sizes, workforce profiles, R&D programs, and financial and management arrangements that characterize the DoD laboratory system as a whole. As a group, they represent an effective cross-section of essential in-house laboratory capability.

I would like to highlight some of our principal areas of concern and briefly illustrate some of our accomplishments.

Our personnel management initiatives seek acquisition of direct hiring authority for all positions; acquisition, development, and retention of S&T positions; implementation of special salary and pay-banding schedules; delegation of increased personnel authority to base commanders; recruitment and retention bonuses; and non-career appointments for retired distinguished senior experts. Some accomplishments to date include: The Office of Personnel Management (OPM) delegated increased local hiring authority to DoD in June 1990. We have been able to hire 59 new PhD's since then. OPM has also authorized 27 additional science and technology positions; and automated classification systems are operating at selected installations. These and other contemplated actions will help the Department develop a personnel climate that will restore our competitiveness in recruiting and retaining the most talented S&E's.

Contracting and procurement initiatives have been put in place to reduce complex and time-consuming procurement methods in contracting for R&D supplies and services; to increase the small purchase threshold to accommodate unique, one-of-a-kind R&D

purchases; and to increase the small purchase competition threshold to significantly reduce procurement acquisition lead times. A seemingly innocuous measure, the delegated use of VISA credit cards to S&E's for small purchases, reduced the average lead time from 100 days to 6 days while ensuring adequate management control. We believe these measures, together with other proposed measures, will give us streamlined procurement procedures that will eliminate excessive and burdensome management controls of the past.

Major facilities modernization initiatives will increase the effectiveness of resources targeted at improving the laboratory infrastructure. We want to give laboratory directors the ability to use scarce resources efficiently and establish a process to create and sustain modern facilities suitable for the conduct of state-of-the-art research and development.

Management authority initiatives seek delegation of increased authority to laboratory directors to control essential support activities in a manner similar to that practiced under the separate-profit-center concept of the private sector; longer tenure for laboratory directors; fewer reporting levels; increased contracting authority for laboratory directors covering services, products, and facilities; and discretionary R&D funding for laboratory directors for enhanced in-house independent research and exploratory development with after-the-fact annual performance reviews. Accomplishments to date are noteworthy: The Navy has delegated contracting authority for an unlimited amount to its laboratory directors; the Army's Laboratory Command has provided its laboratory directors with increased direct control over critical administrative support services; Air Force laboratories have cut the average acquisition lead time in half for some types of equipment and services; and the Defense Nuclear Agency has provided increased management authority to its laboratory director in contracting and technical publications. These measures are designed to vest Defense laboratory directors with levels of authority commensurate with their responsibilities.

A draft legislative proposal, the DoD Laboratory Revitalization Demonstration Act of 1991, has been prepared and is undergoing formal coordination within the Department. This will request legislative authority to demonstrate new and innovative approaches to managing and operating DoD laboratories. The ideas that improve laboratory operations will be exported to other DoD laboratories that are not a part of the demonstration.

As a final note on laboratory demonstration, I would like to address the criteria we are using to evaluate various reforms. In his letter of November 20, 1989, the Deputy Secretary tasked the Services and Defense Nuclear Agency to identify those parameters that are especially important to their particular operations and missions that could be used as "measures of success." The selected indicators were reviewed and refined by the OSD Laboratory Demonstration Program Working Group and will be used to track the progress of the Services and Defense Nuclear Agency. However, I need to caution that using such indicators as "measures" may not always reflect the actual health or improvement in any one laboratory or in the aggregate of demonstration laboratories. The current measures are only a start; we will refine them as needed.

In summary, we have made a determined effort to strengthen our management of the Department's laboratory system through an integrated, corporate approach that provides policy direction and continuity from the DDR&E and the S&T steering group. The DDR&E is also ensuring that the many initiatives we have undertaken are coordinated toward a common purpose. I anticipate that we will see an increasing amount of inter-Service collaboration, both in our efforts to organize and operate the laboratory system more efficiently and in the day-to-day conduct of laboratory work. Because of DMR 922 and the many in-house efforts it has initiated, the development of proposals for base realignment and closure, the guidance we expect from the Advisory Commission, and the many and varied successes and proposals of the Laboratory Demonstration Program, we are making real progress in addressing the complex and abiding problems that hamper the laboratory system today. The result will be a more efficient and effective laboratory system, one that allows this nation to maintain and strengthen the technological edge that has been -- and will remain -- a vital part of our national security strategy.

Appendix E

"Statement on Domestic Technology Transfer"

Presented to the Subcommittee on Oversight and Investigations
of the Committee on Energy and Commerce,
U.S. House of Representatives,
25 July 1991

DEPARTMENT OF DEFENSE
STATEMENT ON DOMESTIC TECHNOLOGY TRANSFER
TO THE
SUBCOMMITTEE ON OVERSIGHT AND INVESTIGATIONS
OF THE
COMMITTEE ON ENERGY AND COMMERCE
UNITED STATES HOUSE OF REPRESENTATIVES

JULY 25, 1991

Good morning. I am pleased to be able to appear before your Subcommittee in order to describe the role of research and development (R&D) within the Department of Defense (DoD) as it relates to domestic technology transfer.

The Department conducts an aggressive R&D program to maintain U.S. defense capabilities at the leading-edge of technology so that our scientists, engineers, and the manufacturing community can design and produce military weapon systems at the highest level of efficiency and effectiveness. A key element of U.S. defense strategy is to conduct R&D for the purpose of maintaining weapons superiority. This longstanding strategy pays off as we have recently seen.

The DoD R&D program is carried out by industry, universities, National Laboratories, Service laboratories, and other organizations such as the National Institutes of Standards and Technology. The private sector performs about 75 percent of DoD's R&D. We routinely organize teams within industry,universities and other laboratories including our own and other Government activities.

Before explaining the overall basis of DoD's domestic technology transfer activities it is important to establish some facts about the Department's R&D funding. In FY 1991, we fund basic research at about $1 billion, and exploratory development (applied research) at about $2.4 billion. The development portion of the budget is $33.2 billion and includes design, engineering, testing and evaluation of weapon systems and equipment.

The DoD RDT&E program emphasis is on development to accomplish our principal purpose -- new and effective technology in military hardware. DoD is unique among federal agencies in the emphasis on systems development. This approach reflects the Administration policy to limit direct Federal support for development to specific Defense areas where there is a clear Federal Government mission. In other areas, product development is appropriately supported by the private sector.

DoD plays a modest role in Federally supported research, providing about 4 percent of the total Federal support for basic research and about 17 percent of Federal funding for applied research. DoD is dominant in Federal support for development, accounting for about 90 percent of the total Federal development expenditure.

Much of the Department's R&D is so-called dual use technology, applicable to both Defense and commercial requirements. Basic scientific knowledge (as opposed to DoD application to weapons) is often not classified. Therefore most of our technology is available to the private sector.

While some companies receiving DoD R&D support concentrate their sales in the defense sector, others are active in both defense and commercial arenas, e.g., General Electric (GE), General Motors (GM), Ford, Boeing, Texas Instruments (TI), Westinghouse, Xerox, and others. Dual use technology developed for DoD by industry inevitably affects both their defense and commercial products. The use of the private sector to perform government is a most effective means for rapid and efficient domestic technology transfer.

The Department's policies in implementing the domestic technology transfer program are pertinent. The program is based upon the Stevenson-Wydler Technology Innovation Act of 1980 and the President's Executive Order 12591 on "Facilitating Access to Science and Technology" of April 10, 1987. We are currently working with other departments, through the Federal Acquisition Regulatory Council and in other forums, to improve our implementation of the Executive Order concerning Rights to Technical Data and Recoupment of Non-recurring Costs. These efforts will further enhance domestic technology transfer. The

Federal Technology Transfer Act of 1986 incorporated major amendments. The Department's current policy directive implements the provisions of those statutes and the Executive Order.

The Federal Technology Transfer Act grants direct statutory authority to the Departments of Army, Navy, and Air Force to execute the programs. Consistent with that authority and Departmental guidance, the Military Departments define the technology transfer program within their organizational framework. The principle elements of the DoD program are:

1. A full-time equivalent position for an Office of Research and Technology Application (ORTA) at each laboratory with at least 200 scientific, engineering, and technical personnel.

2. An active patenting and licensing program. In FY 1990, DoD laboratories received 574 patent awards and submitted an additional 790 patent applications. The DoD patenting and patent licensing facilitates the transfer of technology by allowing us to convey the commercial rights to the private sector through license arrangements. Since contractors conduct most of the DoD R&D program, the full innovative impact of DoD-funded R&D is much greater than the above figures. In this regard, our procurement contracts permit contractors to patent, license, and commercially market technology they initially

develop at government expense subject to reservation by the Government of a nonexclusive, irrevocable, paid-up license for Government use of the invention.

3. Recognization of individual inventors. Inventors have both a strong professional and personal interest in seeing their inventions used in a broad and expeditious manner. In addition they share in the royalties (20%) for commercially licensed inventions. Also, the Services provide monetary and professional excellence awards for significant contributions to domestic technology transfer.

4. Documentation of R&D results. We collect, announce, and release for public purchase, DoD technical reports through the National Technical Information Service (NTIS) at the Department of Commerce. In FY 1990, DoD provided NTIS over 15,000 unrestricted technical reports.

5. Access to DoD laboratory expertise. We have developed a database to refer inquiries from the domestic sector to appropriate experts in DoD laboratories. This interaction increases the potential for commercializing technologies.

6. Listing of DoD patents. We include our patents and licensing contacts in the annual catalog of Federal patents issued by the NTIS.

7. Promoting local laboratory contacts. DoD laboratories directly interact with economic development groups at regional, state, and local levels.

8. Engaging in cooperative research and development agreements (CRDA's). Within DoD, 152 CRDA's are in place and an additional 70 are in negotiation at this time. They encompass a range of technical areas and include state and local governments, the private sector and university participants. The DoD has no significant problems in implementing the provisions embodied in the Federal statutes or Executive Order 12591 in this area.

An important element in understanding the technology transfer process is the realization that there is no single technique that will ensure success. Creating an awareness of the opportunities to commercialize the fruits of Government R&D is a multi-faceted marketing effort.

DoD uses a variety of techniques to facilitate its half of this marketplace equation. Our ORTA's strive to make the non-Federal sector aware of both the capabilities of their laboratories and their technical products (R&D results). Also the Federal Laboratory Consortium is an excellent vehicle for interaction at the local, regional and national level.

In addition, the Department: provides input to databases and directories; hosts conferences, demonstrations, open houses and exhibits; promotes publication in technical literature, newsletters, and the media; and does target mailings to potential commercial interests.

Any successful effort to push Government technology to the commercial marketplace requires industry to shape technology into commercial products by applying their design, manufacturing, distribution, and sales expertise.

As indicated earlier, contractors perform the majority of DoD's R&D. Technology advances developed to meet DoD mission needs routinely result in corollary commercial activities. Past spin-offs to the commercial sector from high risk dual use research supported by DoD have played a major role in U.S. economic strength. In the computer industry, DoD supported efforts have led to computer time sharing, computer graphics used in computer aided design and biotechnology, and computer aided instruction.

DoD has supported materials science in areas such as ceramics and ceramic composites, microelectronics materials, metal alloys and rapid solidification, and advanced polymers. Modern materials coming out of DoD research are found in diesel and gasoline powered automobiles, sporting goods, medical prosthetic implants, and many other products of U.S. companies.

Commercial application of Defense R&D is an ongoing process. Recent research conducted by the Navy in synthetic red blood cells shows great promise. The goal of assuring a supply of disease free blood cells is of great significance, both from a military and civilian perspective.

In addition to the general approaches to technology transfer discussed previously, we would like to mention some specific DoD outreach efforts. One is the Navy's Domestic Technology Transfer Fact Sheet. This monthly publication describes recent technical innovations and explores their commercial application potential in a newsletter format. With over 8,000 (mostly private sector) subscribers for this monthly bulletin, it provides a medium to disseminate technologies with commercial potential.

Through its civil works authority, the Corps of Engineers sought and received Congressional approval to engage in cost shared cooperative R&D for construction purposes. As the Federal Government's executive agent, the Corps is one of the largest purchasers of general purpose commercial type construction in the U.S. As concluded in a study by the National Research Council of the National Academies of Science and Engineering in late 1986, the U.S. construction industry is in the bottom tier (0.39 percent of revenues) of U.S. R&D investment for mature industries.

This Army effort identified as the "Construction Productivity Advancement Research" program provides for up to 50 percent cost sharing of projects. In the three years since its inception, the Corps has funded 37 proposals with a committment of $9.6 million of Federal funds and $19.3 million in non-Federal funds. The Corps has received proposals from a mix of large and small businesses, universities and colleges, state and local governments, and trade associations. They cover a wide range of applications including hardware, machinery, robotics, materials, and software.

The most significant problem we have experienced to date is being addressed in legislation introduced as H.R. 191, "Technology Transfer Improvements Act of 1991". This would permit the Federal Government to copyright computer software created or developed by the Government as part of a CRDA. Software is one of our national competitive strengths. The ability to copyright and license software will give the private sector time needed to commercialize our software and develop the required user documentation. Without this protection of their investment, the private sector elects not to use our valuable asset.

The Department of Defense recognizes that national security is closely tied to economic strength. We believe that investment in DoD R&D has provided an excellent return for the nation. The DoD domestic technology transfer program continues to grow by all measures. We will continue to nurture our efforts in this area which is vital to our national security.

Appendix F

"Report to the Secretary of Defense, Federal Advisory Commission on Consolidation and Conversion of Defense Research and Development Laboratories"

Executive Summary
September 1991

CHARLES E. ADOLPH, CHAIRMAN

FEDERAL ADVISORY COMMISSION ON CONSOLIDATION AND CONVERSION OF DEFENSE RESEARCH AND DEVELOPMENT LABORATORIES

COMMISSIONERS
DR. SOLOMON J. BUCHSBAUM
MR. ROBERT M. HILLYER
DR. O'DEAN P. JUDD
DR. JAMES C. McGRODDY
DR. JOHN M. PALMS
DR. FRANK D. VERDERAME

COMMISSIONERS
DR. JAMES F. DECKER
DR. JOHN W. LYONS
DR. WILLIAM C. McCORKLE
MR. EARLE L. MESSERE
COL. RICHARD R. PAUL
DR. VICTOR H. REIS

September 30, 1991

Honorable Richard Cheney
Secretary of Defense
The Pentagon
Washington, DC 20301

Dear Mr. Secretary,

We are pleased to submit the report of the Federal Advisory Commission on Consolidation and Conversion of Defense Research and Development Laboratories. This report contains the Commission's findings and recommendations on the feasibility and desirability of various means to improve the operation and effectiveness of the DoD laboratories. To prepare this report, the Commission met for nine sessions, received extensive briefings from the Services, and reviewed previous Defense laboratory studies and other written documentation. Among the individuals who presented testimony were members of Congress and officials representing the Department of Defense, the Office of Technology Assessment, and the General Accounting Office. In addition, the Commission conducted on-site reviews at one laboratory of each of the Military Departments.

The Defense laboratories provide the technical expertise to enable the Services to be smart buyers and users of new and improved weapons systems and support capabilities. There is significant room for improvement in the operation of the laboratories, and our report contains specific recommendations to move in that direction. The Services' planned restructuring and realignment, coupled with interservice cooperation through Project Reliance and vigorous implementation of the DoD Laboratory Demonstration Program, afford a unique opportunity to improve DoD laboratory effectiveness.

Respectfully,

Charles E. Adolph
Chairman

James F. Decker	Richard R. Paul	O'Dean P. Judd
John W. Lyons	Victor H. Reis	James C. McGroddy
William C. McCorkle	Solomon J. Buchsbaum	John M. Palms
Earle L. Messere	Robert M. Hillyer	Frank D. Verderame

FEDERAL ADVISORY COMMISSION ON CONSOLIDATION AND CONVERSION OF DEFENSE RESEARCH AND DEVELOPMENT LABORATORIES

EXECUTIVE SUMMARY

Public Law 101-510 established the Federal Advisory Commission on Consolidation and Conversion of Defense Research and Development Laboratories to study the Department of Defense (DoD) laboratory system and provide recommendations to the Secretary of Defense on the feasibility and desirability of various means to improve the operation of the DoD laboratories. Among the means the Commission was directed to study were: (1) conversion of some or all of the DoD laboratories to Government-Owned, Contractor-Operated laboratories, (2) mission and/or function modification of some or all of the laboratories, and (3) consolidation or closure of some or all of the laboratories.

The DoD operates a large and complex laboratory system. The DoD laboratories (42 Army, 20 Navy, and 4 Air Force) spend approximately $6.5 billion annually and employ nearly 60,000 people, of whom over 26,000 are scientists and engineers. The DoD laboratory system has evolved over the past 150 years. Each Service's system is different and is a product of its historical origins, culture, and method of systems acquisition. Several laboratories are embedded in larger organizations. A significant number of the laboratories are relatively small and geographically isolated.

In undertaking its task, the Commission started with the fundamental issues concerning the laboratories: Why does the DoD have in-house laboratories? What are their primary functions? What is their current level of effectiveness? What are the attributes of an effective laboratory? How best can these attributes be achieved within the current environment? Is conversion to Government-Owned, Contractor-Operated necessary and/or feasible?

With a consensus on these fundamental issues, the Commission focused on the efficacy of the Services' laboratory reorganization plans and other opportunities for improving the productivity and effectiveness of DoD laboratories. Additionally, the Commission sponsored an independent assessment of the methodology and data that the Services used in evaluating the costs and savings associated with implementing their laboratory reorganization plans.

PRINCIPAL FINDINGS AND RECOMMENDATIONS

FINDINGS

1. The mission of the Defense laboratories is to provide the technical expertise to enable the Services to be smart buyers and users of new and improved weapons systems and support capabilities.

2. The functions provided by the DoD laboratories are an essential part of the acquisition process. Dedicated organizations free from commercial pressure are required to provide these functions.

3. The Services operate laboratories that span the range from those with broad research, development, and engineering responsibilities to those focused on science and technology. The Army and Navy operate both types of laboratories, while the Air Force operates the latter type. The laboratory types within each Service are a function of that Service's weapons systems acquisition structure. There is no need to force the Service laboratory systems into a single model.

4. While the Services are making progress, there is the need to improve the effectiveness of the DoD laboratories.

5. The following attributes are essential to achieving high quality and effectiveness:

- o Clear and substantive mission
- o Critical mass of assigned work
- o A highly competent and dedicated work force
- o Inspired, empowered, highly qualified leadership
- o State-of-the-art facilities and equipment
- o Effective two-way relationship with customers
- o Strong foundation in research
- o Management authority and flexibility
- o Strong linkage to universities, industry, and other Government laboratories.

6. Restructuring the in-house laboratory system is not only essential to achieve cost reductions, it also should be used as a major opportunity to improve effectiveness.

7. In general, the Services' cost and savings estimates associated with their laboratory reorganization plans are in

accordance with established procedures for base closures and are reasonable.

8. The restructuring of the laboratories, as proposed by the Services, could result in work-force turbulence, loss of key technical personnel, and disruption of critical research and development activities, therefore requiring special attention.

9. Strong advocacy on behalf of the laboratories at Service headquarters and in the Office of the Secretary of Defense is needed to ensure the effectiveness of the laboratories.

10. The effectiveness of the DoD laboratories suffers from regulatory and policy impediments to the authority and flexibility of the individual laboratory directors.

11. DoD-wide commitment to laboratory management excellence, high-level advocacy, and removal of obstacles to management authority and flexibility will provide an environment for greatly improving the productivity and effectiveness of the laboratories.

12. The Laboratory Demonstration Program and the recently enacted Federal Employees Pay Comparability Act contain many of the provisions needed to enhance organic management flexibility.

13. Conversion of some or all of the laboratories to Government-Owned, Contractor-Operated organizations could improve their effectiveness. However, fixing the problem organically is preferable to such a conversion.

14. The recently initiated interservice Project Reliance offers considerable potential for strengthening the effectiveness, productivity, and cohesiveness of DoD science and technology.

15. Many of the observations in this report have been made numerous times in the past, but have not been acted upon. The planned laboratory restructuring and realignment effort affords a unique opportunity to achieve significant improvements.

RECOMMENDATIONS

To materially improve the Defense laboratory system, the Commission recommends the following:

1. The proposed Army and Navy laboratory consolidations and realignments should begin in January 1992. The Army should delay implementation of the microelectronics function at Adelphi, Maryland, and construction of the facility to house the function until the completion of the study in recommendation 7. The Air Force should continue implementation of its laboratory consolidation plan. All service plans should be implemented so as to minimize disruption during the transition to a new structure.

2. The Secretary of Defense should direct the Services to implement all the provisions of the Laboratory Demonstration Program without delay, extend the program to all DoD laboratories, and seek legislative action required to complete the Laboratory Demonstration Program initiatives, including the personnel-related actions.

3. The Secretary of Defense should instruct the Services to delegate the authorities provided under the Federal Employees Pay Comparability Act immediately to the individual laboratory directors.

4. The Secretary of Defense should direct each Service to establish a high-level advocate who will report to the Service Assistant Secretary level and who will be accountable for the effectiveness of its laboratories.

5. The Services should strengthen the selection process for laboratory directors, emphasizing technology and technology-management qualifications. These positions should be for a minimum term of 4 years.

6. Each laboratory should establish an advisory committee of outside experts to review periodically the status of the laboratory and its work, and make recommendations to the director.

7. An independently appointed review group should assess the advantages and disadvantages of a single microelectronics research facility for all three Services. If a single facility is a viable solution, consideration should be given to a Government-Owned, Contractor-Operated laboratory.

8. The Services should continue to implement Project Reliance and the Director, Defense Research and Engineering should review the implementation of Reliance agreements periodically to ensure that there is no unwarranted duplication and that optimum resource utilization is achieved.

9. The Director, Defense Research and Engineering should ensure through periodic reviews that the recommendations contained in this report are being implemented. In addition, the Director should review the status of the individual Service laboratory consolidations and realignments at least semi-annually to ensure that they are being accomplished to maximize effectiveness and minimize disruption to personnel and ongoing technical programs.

SERVICE-SPECIFIC FINDINGS AND RECOMMENDATIONS

ARMY - FINDINGS

1. The Army's proposed laboratory consolidation and realignment should result in a more effective laboratory structure: eight streamlined Research, Development, and Engineering Centers within the commodity commands and the Combat Material Research Laboratory. The Commission supports this proposed consolidation.

2. Strong leadership at the Combat Materiel Research Laboratory is crucial to that laboratory's successful startup and long-term success.

3. The Combat Materiel Research Laboratory and the Research, Development, and Engineering Center technology base activities must interact with and support each other to achieve maximum effectiveness. High-level leadership must oversee and manage an active cooperative effort.

4. High-level leadership must institute active measures to maintain the connectivity between the Combat Materiel Research Laboratory and the Army user community.

5. The effectiveness of the laboratories can be improved by significantly increasing their connectivity to the acquisition process.

6. An underpinning research program within each Research, Development, and Engineering Center is important to its success.

7. The large capital investment planned for a new Army microelectronics research facility at the Combat Materiel Research Laboratory may not be warranted.

8. The Army's plan to ensure responsiveness of the Combat Materiel Research Laboratory science and technology program to the Research, Development, and Engineering Centers through a Board of Directors is sound.

9. The Army's plan to allocate a substantial part of its 6.1 budget to in-house laboratory independent research at the Research, Development, and Engineering Centers will ensure basic research programs in support of their missions.

ARMY - RECOMMENDATIONS

The Army should:

1. Appoint a strong civilian director for the Combat Materiel Research Laboratory as soon as possible. The new director must be given extensive authority to form Combat Materiel Research Laboratory divisions for maximum effectiveness and to recruit Combat Materiel Research Laboratory division leaders. This director should be a scientist or engineer with stature as a research and development leader and administrator.

2. Hold the Assistant Secretary of the Army for Research, Development, and Acquisition responsible for appointing and rating each of the Research, Development, and Engineering Center directors and the director of the Combat Materiel Research Laboratory.

3. Use all possible incentives to minimize turbulence, loss of key personnel, and disruption of critical research and development programs. These incentives include retention bonuses, relocation services and assistance, placement services, and time flexibility.

4. Defer the capital investment for an Army microelectronics research facility at the Combat Materiel Research Laboratory pending the outcome of principal recommendation 7.

5. Implement procedures for the Combat Material Research Laboratory and the Research, Development, and Engineering Centers to evaluate and interact with each other's programs and with the user.

6. Include all Army Research, Development, and Engineering Centers and laboratories in the Laboratory Demonstration Program.

NAVY - FINDINGS

1. The Navy's proposed laboratory consolidation and realignment will result in an organizational structure that includes a range of functions from science and technology to depot support within each of four major Naval Warfare Centers (Air, Surface, Undersea, and Command and Control), each of which has one or more research and development elements embedded within it. This overall structure provides flexibility for change in the face of uncertain future budgets and problems. The Commission supports the warfare center concept with the following reservations:

a. There is risk that the research and development elements of the warfare centers will lose their identity as laboratories in the planned structure.

b. A high-level official responsible for laboratory effectiveness is not identified in the plan.

2. The Navy's planned personnel relocations (approximately 4800) present a particular challenge to minimize work-force turbulence, loss of key personnel, and disruption of critical research and development programs.

NAVY - RECOMMENDATIONS

The Navy should:

1. Modify the plan to identify the research and development element or elements within each warfare center as Navy Research, Development, and Engineering Laboratories. These activities will be DoD laboratories, as will the realigned Naval Research Laboratory and the Navy medical laboratories. Each of these laboratories should be led by a scientist or

ES-7

engineer with stature as a research and development technical manager.

2. Use all possible incentives to minimize turbulence, loss of key personnel, and disruption of critical research and development programs. These incentives include retention bonuses, relocation services and assistance, placement services, and time flexibility.

3. Include each Navy Research, Development, and Engineering Laboratory within each warfare center along with the Naval Research Laboratory and the Navy medical laboratories in the Laboratory Demonstration Program. Consistent with the Laboratory Demonstration Program, these laboratories, including the Naval Research Laboratory, should have their own organic support.

AIR FORCE - FINDINGS

The Commission finds:

The Air Force Laboratory Consolidation Plan will improve the overall effectiveness of the Air Force laboratory system. That plan, already partially implemented, provides for the following:

a. Organizational consolidation of 14 laboratories into four laboratories that align with and reside in the Air Force Systems Command's four product divisions.

b. Gradual geographical migration of the elements associated with each laboratory to that laboratory's headquarters location.

c. A Technology Executive Officer who provides integrated science and technology investment strategy guidance to the four laboratories and serves as a dedicated Air Force laboratory system advocate.

d. Strong emphasis on technology transition and support of the weapons systems acquisition process through direct reporting of laboratory commanders/directors to their product division commanders.

AIR FORCE - RECOMMENDATIONS

The Air Force should:

1. Continue implementation of the Air Force Laboratory Consolidation Plan.

2. Use all possible incentives to minimize turbulence, loss of key personnel, and disruption of critical research and development programs. These incentives include retention bonuses, relocation services and assistance, placement services, and time flexibility.

3. Continue to improve the connectivity between the laboratory structure and the acquisition elements of the product divisions.

4. Include all Air Force laboratories in the Laboratory Demonstration Program.

Appendix G

"Test and Evaluation Challenges in the Nineties"

Paper Presented to the NATO
AGARD Conference Proceedings 519
May 1992

TEST AND EVALUATION CHALLENGES IN THE NINETIES

by

Charles E. Adolph
Director, Test and Evaluation Office
of the Secretary of Defense
Room 3E1060
The Pentagon
Washington DC 20301-3110
United States

INTRODUCTION

Good morning ladies and gentlemen. It is a pleasure to be here with you today. I'm pleased to have the opportunity to speak to you.

The Flight Mechanics Panel devotes a symposium to flight testing every 3-4 years. Many of you attended the last one at Edwards Air Force Base, California in October 1988. No one could have anticipated the dramatic changes that have taken place in the world since we last met to discuss flight testing. The Warsaw Pact is gone. The Soviet Union no longer exists. Germany is reunited. The United States has been involved in two limited conflicts. We are drawing down the Defense establishment in the United States to reflect the new world situation. I want to briefly discuss the Department of Defense plans for restructuring, because they have a major impact on the weapon system acquisition and test community.

Revised Acquisition Approach

Let me start with a few excerpts from Secretary of Defense Dick Cheney's news briefing at the Pentagon on January 31 of this year when he summarized DoD restructuring and budget plans.

Figure one graphically depicts the change in funding for the DoD. The cumulative decline from 1985-1997 is 37 percent. The rate of real decline, which is often referred to in the Pentagon as the "glide slope", is 4 percent per year.

At the same news briefing, Mr. Atwood, the Deputy Secretary of Defense, noted that "We are no longer being chased by the Soviets in the development of new technology. We have weapons that have proven themselves to be the finest in the world. And we don't have to get into production on new ones just because we're chasing someone or they're chasing us."

In view of changes in the world situation and tight constraints on defense resources, the United States is reducing the size of its armed forces and revising our approach to defense acquisition. We will place greater emphasis on technology demonstration. We can now afford to take time to develop and evaluate technologies thoroughly before proceeding to production. We will decide to produce and field a system only after we minimize technical, production, and operational risks and verify the absolute need for production of a weapon system. We will also emphasize the insertion of proven technologies into existing weapon systems over producing new weapon systems, when technology insertion can meet operational needs.

In summary, the salient features of the revised approach to acquisition are:

(1) increasing research on advanced technologies;

(2) increasing development and evaluation of demonstrators;

(3) incorporating advanced technologies into existing weapon systems only under the following conditions:
- (a) we have thoroughly tested the technology and associated subsystems;
- (b) we have a substantive need for improved performance or reliability; and
- (c) the incorporation or retrofit program is cost-effective; and

(4) producing new weapon systems using advanced technology only under the following conditions:
- (a) we have thoroughly tested and proven the technology and associated subsystems;
- (b) we have significantly minimized the technical, production, and operational risks;
- (c) the production program is cost-effective; and
- (d) we have verified an absolute need for a new weapon system.

Under the revised approach, we will produce less. Rather, DoD will emphasize system upgrades, incorporating technological innovations that have been proven and developed. We will emphasize sound development and carefully managed risk reduction over accelerated development and hasty fielding. Maintaining a technological advantage will continue to be a priority; however, we can reduce the quantity of new weapon systems we produce. Dr. Vic Reis, Director of Defense Research and Engineering described the strategy for science and technology in the following way in his testimony this March before the House Appropriations Committee: "The core of this strategy has these primary elements; sustaining and applying the explosion in information technology, involving the user early and continuously, and demonstrating the technology extensively and realistically."

To better manage science and technology efforts, DoD is establishing a Defense Technology Board, which is chaired by the Director, Defense Research and Engineering. The Director, with the advice of the board, will focus science and technology efforts and establish broad policy. This will help us maintain a strong defense technology base and avoid major costs associated with premature systems development.

I am a member of the Technology Board and the DoD test and evaluation community will participate in the science and technology working groups to ensure that we are posturing our test facilities to rigorously evaluate the new technologies to be incorporated in weapon systems and subsystems.

AIRCRAFT DEVELOPMENT AND TEST CHALLENGES FOR THE 1990's

Now that you have some insight into the restructuring that is underway within our Department of Defense and our revised approach to weapon system acquisition, I want to move on to a discussion of the aircraft design and test challenges which lie ahead. The issues have become increasingly complex over the past 30 years. In the 50's, the airplane community assumed that flying faster and higher was better. In the 60's, excess thrust or energy - maneuverability came into vogue as the primary measure of merit. We also began to pay increasing attention to improving lift-limited performance and to providing more benign high-angle-of-attack characteristics. Flying high detracted from survivability in most situations; so we emphasized operating effectively in a low-level, terrain-following, high-dynamic-pressure environment.

In the 80's, signature reduction came into vogue as the primary means of improving survivability. In the U.S. we also began to pay far more attention to reducing the vulnerability of aircraft. We now take a methodical systems approach to designing in and testing for survivability.

Survivability: Improving the Odds

We all know that the probability of survival in mathematical terms is one minus the product of susceptibility times vulnerability, or one minus the probability of a hit times the probability of a kill given a hit. Simply stated, survivability focuses on the tactical and technical means to lower the risk of being killed.

Susceptibility reduction measures require a design that balances agility and signature reduction while incorporating an electronic warfare suite for warning and deception, electronic countermeasures, combined with threat suppression weapons and electronic techniques. Vulnerability reduction is now a major design consideration. Damage tolerance (redundancy) and resistance (hardening) are designed in from the outset.

What are the implications of this highly integrated systems approach to survivability design for the aircraft test community? Simply stated, it has made our task much more complex. Evaluating survivability is emerging as our most challenging task.

Any well-executed test program is a product of good up-front planning. Unfortunately, combat survivability requirements tend to be relegated to lower priority in the competition for funds and other test resources since survivability information is usually not needed to demonstrate initial system operation. Although developmental and operational testing address the components of survivability to some extent, we need a more systematic and quantifiable approach. Often, because of funding and other constraints, we can't subject some survivability features to actual tests. Hence, we must establish modeling, simulation, and analysis techniques early in our planning to supplement testing.

Vulnerability is the more easily quantified term. From live fire testing and detailed simulations, we can estimate the probability of a kill given a hit. Not everyone would agree that the estimates are valid, but at least there is agreement on the approach. We can assess damage tolerance, damage resistance, personnel protection and battlefield damage repair.

I would now like to move on to a discussion of susceptibility. The task of measuring susceptibility is far more complex than measuring vulnerability or lethality. Many factors are involved in lowering susceptibility. Signature reduction techniques are currently being emphasized in the design of most aircraft. This involves various types of signature reduction including infrared, electro-optical, acoustical, and physical size. Improvements in agility, countermeasures, threat warning, threat suppression, terrain masking, and employment tactics also lower susceptibility. I might note that low observable technologies have created a new set of challenges for the test community; the ability to measure the signature characteristics of the vehicle without altering those characteristics.

For susceptibility, numerous factors are difficult to properly quantify and vary from scenario to scenario. For example, the susceptibility of an aircraft to a specific threat system is a function of many factors, including the number of aircraft in the raid, radar and communications support jamming, defense suppression, jinking maneuvers, and so forth. Seldom can we include all these factors in a test because of the availability of assets, cost, and the complexity of discerning just what happened in the field test of a very complex scenario. Simulation is often used in conjunction with field testing to assess the above factors. However, the campaign-level simulations needed to accomplish these analyses require improvement.

Testers must define those factors that impact on susceptibility of a system--speed, altitude, terrain-following and electronic countermeasures, and compare them to user requirements. As an example, does flying 200 knots faster improve survivability more than flying 100 feet lower? The correct approach is for the user, developer, and tester to agree early in the requirements definition process to the parameters which most directly influence survivability and also to establish criteria for each of the parameters that, if met, would ensure an adequate level of survivability.

Threat Definition

Threat definition is obviously a sensitive subject, from a political as well as a security perspective. We have given considerable attention to the threat to be used for design and test purposes. One of the issues raised was that the test community sometimes makes a unilateral interpretation of the threat to be used for testing. In the early stages of design and test plan development, we need to address the basis for threat scenario development, to include targets, countermeasures, scenarios, doctrine, and tactics employment, as well as the degree to which models and simulations will be used.

We all know that the threat is not constant. It continues to evolve, and, occasionally, new technologies may be introduced which were not known when the weapon system requirements documents were defined. In addition, our knowledge of the threat improves with time. This demands a disciplined process to update threat scenarios. Both short-term and long-term projections of the anticipated threat are needed. Until recently, our major threat was clearly the former Soviet Union and the Warsaw Pact countries. Today's environment, while less threatening, is also less clear. We have less predictable regional threats. Both allies and adversaries may have a mix of so-called blue, grey, and red weapons.

Weapons System Performance Envelope

I would like to briefly discuss, on a philosophical level, testing in the context of weapon system performance envelopes and relate this to the threat scenario issue. One of the functions of developmental testing is to expand and define the system performance envelope. The T&E community is sometimes accused of having a preoccupation with testing at the fringes of the envelope. The implication is that we testers try to set up scenarios where the system is likely to fail.

My philosophy on this issue has evolved to something like this: Once the envelope is defined, operational testers should insure that the system works very well and reliably in the heart of the requirements envelope. Note that I used heart of the requirements envelope rather than the heart of the performance envelope. The heart of the requirements envelope takes into consideration the scenario that is most likely as well as the middle of the system performance envelope. Operational testers should also assess performance at the fringes of the envelope but not become preoccupied with that area. Some degradation can be expected at the fringes. By the way, testing at the fringes of the envelope is generally extremely resource-intensive and expensive. Now let me relate this discussion to threat scenarios.

At the heart of the threat envelope are the highly probable, operationally realistic threats against which we test our weapon systems. There are other, less probable, threats at the fringes of the envelope. These are the threats and technologies that could show up and against which our weapon systems should be tested, such as exotic electronic countermeasures. We must develop a process that tests our weapon systems at the fringes of the test envelope without judging the results too critically. Certainly, test and evaluation using these threats must not be viewed as a pass/fail final exam. At the same time, we must not bury our heads in the sand and ignore the threats on the fringes or hope they will not be fielded by the time our weapon systems are fielded.

We all know from past experience that our airplane platforms of the future will be in service for 30 years or more. Design modularity and increased attention to life-cycle costs are essential; subsystem upgrades and planned product improvements will be a way of life. It is also important to recognize that we actually obtain as much if not more new information when a system fails a test as we do when it succeeds. During the development process, we expect failures. That's what development is all about.

Testing Software-Intensive Systems

Let's move on now to what I consider to be another challenge facing the military test community in the next 20 years - the efficient evaluation of software-intensive systems. Integration and checkout of software-intensive systems is the one cost and schedule risk area associated with most of today's development and test programs -- whether aircraft, surface ships, submarines, missiles such as the advanced medium range air-to-air missile, and stinger reprogrammable microprocessor missile or over-the-horizon radars.

The software-intensive systems of today present both a quantum leap in capability and a quantum jump in test requirements. A high percentage of the test work accomplished on aircraft such as the F-14, F-15, F-16, F-18, B-1 and B-2 is tied to software-intensive systems. Almost two-third's of the flight testing on these aircraft is related to avionics. The challenge for the designer, evaluator, and operator is to get the highest benefit from the flexibility afforded by software-intensive systems, while at the same time being cost-effective. Experience over the past several years has shown that we tend to grossly underestimate the amount of testing necessary to fully develop software-intensive systems. As an example, the number of test flights required to develop and evaluate a software-intensive system is often two to four times higher than the initial estimate. Our estimate of test requirements is much poorer today than it was 20 years ago with hydro-mechanical systems.

Ground Simulation Facilities

We also need to pay more attention to system integration. On a programmatic level, we need to make greater use of simulators to support both the development and test process. Using simulation facilities to support aircraft avionics testing can give us tremendous payoffs. Let me cite some examples:

- We can resolve three-fourths of our software problems on the ground at a fraction of the cost of flight.
- We can reduce flying hours which translates into reduced test costs and accelerated test schedules.
- The costly and inefficient fly-fix-fly approach is minimized.
- Ground testing is more efficient because the experiment is controlled, and repetition of test conditions is rapid and simple. We can work the problem 3 shifts a day, 7 days a week, if necessary.
- We use flight test time more effectively by isolating/keying on risk areas and smarter profile planning.

Electronic Warfare Systems

I will now move on to a particularly troublesome set of software-intensive subsystems. Flight and ground testing of electronic warfare (EW) systems is becoming increasingly complex and more time-consuming. Our test facilities have evolved around developing aircraft and hard-kill weapons. We need significant improvements in our EW test capabilities. We need better indoor, hardware-in-the-loop test facilities to develop and evaluate the effectiveness of EW systems. We also need improved ground test and flight test range capabilities.

The complexity of today's weapon systems has increased the number of test modes, conditions, and parameters to the point where flight testing against simulated environments on outdoor ranges alone is insufficient. As our technology moves on into the realm of integrated avionics with situation-adaptable processing, we have no alternative but to find ways to test smarter, cheaper, and more efficiently. We are challenged to create realistic threat situations that include the density and complexity of the whole electronic order of battle.

To meet this challenge we are beginning to use full-scale hardware-in-the-loop simulation facilities. Electronic combat test facilities will ultimately allow us to insert a new aircraft into a full force-on-force combat environment without leaving the ground. One such facility is the U.S. Navy's Air Combat Evaluation Test and Evaluation Facility located at the Patuxent River Naval Air Station. This facility consists of a tactical aircraft-sized anechoic chamber, an electronic warfare integrated systems test laboratory, a closed-loop missile threat simulator, and a motion-based full-visual man-in-the-loop cockpit simulator. All of these simulations, when tied together, provide a very high degree of combat fidelity that is difficult if not impossible to duplicate in flight.

Modeling and Simulation

I have used the words simulation or simulation facility several times today. Over the past several years, we have used models and simulations increasingly to support the test and evaluation process. We can expect such use to continue---indeed, increase--in the future. There is enormous potential for the prudent use of models and simulations to both mitigate some of the shortcomings of the test and evaluation process and to reduce the risk of developing new weapon systems. In the current budget environment, the potential to help us test smarter, evaluate more completely, and bring down the cost of doing business make models and simulations essential tools for the T&E community. I am not going to discuss all the applications of modeling and simulation to our business, but I do wish to make one point with respect to them: models and simulations are T&E resources and need to be treated as such!

This means that, first of all, we must recognize that validated, standardized models and simulations are long-term capital investments in our test and evaluation capability just as laboratories and ranges are. As such, we must plan ahead carefully for their development and employment, we must ensure that they are properly operated and maintained and we must fund their support accordingly. The most difficult task is deciding where and when to use models and simulations.

In the use of models and simulations, as in all aspects of test and evaluation, credibility is the most essential characteristic. The credibility of a model or simulation cannot be considered separately from its application to a specific problem, the validity of its data inputs, and the qualifications of those executing the simulation and interpreting the results. The T&E community needs to place the same kind of high priority on the quality of its models and simulations as it does on its other tools of the trade. The models and simulations to be used in support of a specific weapon system need to be identified early on, validated, and accredited for the specific application. As an example, accreditation for use in sensitivity analyses does not necessarily imply accreditation for other purposes.

And just as we seek to harness technology to constantly improve our instrumentation, measurement, and data processing capabilities to support test and evaluation, we must also exploit emerging technologies in the area of simulation. There is dynamic growth in the computing and networking areas, software structures, module and database interfaces, and languages. We need to take advantage of all the available advances to most effectively capitalize on the powerful potential of models and simulations.

We have a major initiative within the Department of Defense to make better use of models and simulations. A Defense Modeling and Simulation Office was recently established to provide oversight to the process. Test and evaluation is one of five panels.

We must reduce the proliferation of models and simulations used to support development and testing. There is an inverse correlation between numbers of models/simulations and their credibility. The more often a model or simulation is used for valid and accredited purposes, the better its credibility. Additionally, we are either maintaining, or worse, failing to maintain far too many models at too great a price. We can accomplish this reduction in part by use of standard architectures and protocols -- thus promoting interoperability and reducing the desire to construct specialized, stand-alone models. We can also reduce the proliferation through information sharing. Models and simulations should be tools to make us more effective and efficient in our primary business, test and evaluation. These tools will do a better job for us if we also manage their use efficiently and effectively!

Increased Emphasis on Evaluation and Evaluation Tools

A major challenge and cultural change that the test community must deal with is the need to place more emphasis and thought on the evaluation process. We have talked about Test and Evaluation (T&E); the emphasis in the past has been test, in today's world it needs to be evaluation.

The test community must work much more closely with the organizations that define the weapon system requirements and associated threats. As I discussed earlier, we must develop threat and scenarios and use them to define critical and measurable technical performance parameters and operational measures of effectiveness. These must include quantifiable survivability criteria. Models and simulation tools are essential to augment field test results. Our objective is to provide more effective and efficient evaluation by using a suite of tools that include modeling, simulation, and ground test facilities for subsystem and system-level testing to augment flight testing. We perform Cost and Operational Effectiveness Analysis (COEAs) on U.S. weapon systems to assist in determining whether new weapon system or system upgrades offer sufficient military benefits to be worth the cost. Test and evaluation aids decisionmakers by verifying that systems have attained their technical performance specifications and objectives and are operationally effective and suitable for their intended use.

Our acquisition policy mandates that a linkage exist between COEAs and test and evaluation. To judge whether an alternative is worthwhile, we must first determine what it takes to make a difference. Measures of effectiveness are defined to measure operational capabilities in terms of engagement or battle outcomes. Measures of performance, such as speed, are related to the measures of effectiveness such that the effect of a change in the measure of performance can be related to a change in the measure of effectiveness. We develop measures of effectiveness to a level of specificity such that a system's effectiveness during developmental and operational testing can be assessed with the same effectiveness criteria as used in the cost and operational effectiveness analysis. This permits us to further refine the analysis to reassess cost effectiveness compared to alternatives in the event that performance, as determined during testing, indicates a significant drop in effectiveness (i.e., to or below a threshold) compared to the levels assumed in the initial analysis. A comprehensive test and evaluation program is an essential part of an operational effectiveness analysis. Test results give credence to the key assumptions and estimates that were made in earlier cost and operational effectiveness analyses.

Next Generation Aircraft

I would like to get my crystal ball out and take a look at the future. You all know that there will be fewer platforms and weapons. In the late 50's, the 1960's, and the early 70's, there were numerous new aerodynamic designs. In the early 70's alone, we made first flights on numerous aircraft. The number of totally new aircraft designs has decreased dramatically in the past 10 years, and there may be no major new tactical and strategic designs in the U.S. beyond the B-2, F-22, and AX for the next 20 years. As I mentioned earlier, "stealth" will be designed into most systems. Signature reductions may include physical size, acoustical, as well as the more common and infrared RF. Stealthy aircraft carry their weapons internally, which implies fewer weapons and weapons types. This means that there will be less integration testing per type of store and that requires more rigorous analysis.

Advanced guidance and control technologies provide for more precise maneuvering of platforms. One trend in aircraft design is ever increasing reliance on integrated automatic control. This can have a profound effect on test requirements because the way the aircraft is flown can be significantly altered through minor modifications of the flight control system. The incorporation of "carefree maneuvering" will increase the test requirements close to max load factor. Advanced guidance technologies, combined with an increasingly lethal battlefield air defense environment, have been the driving forces behind the development of precision guided stand-off weapons.

Aircraft of the future will have more highly integrated modular avionics and electronic warfare suites. We will have fewer test programs in the year 2000 but these programs will require fewer but more sophisticated test capabilities. Higher performance vehicles and weapons also drive requirements for more airspace and faster movement through the airspace, as well as larger impact footprint areas.

CONCLUSION

In conclusion, these are exciting times for the test and evaluation community. Aircraft systems and aircraft test programs have become highly integrated, increasing our technical and management challenges. The changing world situation and fiscal environment of the 1990's presents unprecedented changes and a level of turbulence that will not subside in the next year or two. The software-intensive systems of today require new and innovative technical and test management approaches. There is a need for greater use of simulators and other hardware-in-the-loop ground test facilities to accelerate the integration and checkout of software-intensive systems. Electronic warfare systems testing will place even greater demands on the use of computer models, simulations, and other indoor and ground test facilities. Finally, we must bear in mind that we in the test community are the disciplining factor in the acquisition process. It's up to us to ensure that weapons systems meet both design and operational requirements.

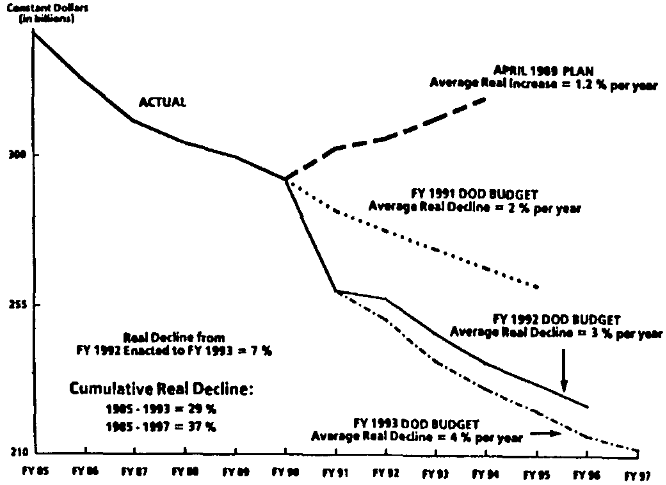


FIGURE 1

Appendix H

"Test and Evaluation Facilities, Fast-Track Study"

Office of the Secretary of Defense
4 August 1993

SECDEF DIRECTION

- OSD, ASSISTED BY THE SERVICE SECRETARIES, WILL STREAMLINE THE T&E INFRASTRUCTURE. THE OBJECTIVE IS TO ACHIEVE MANAGEMENT EFFICIENCIES BY BETTER INTEGRATING THE ACTIVITIES OF INDEPENDENT FACILITIES AND POSSIBLY **BY LINKING CONUS TEST AND EVALUATION RANGES ELECTRONICALLY WITHIN BROAD GEOGRAPHIC AREAS. CONSIDERATION SHOULD BE GIVEN TO DESIGNATING AN EXECUTIVE AGENT AS SINGLE MANAGER FOR TEST AND EVALUATION RANGES. CONSIDERATION SHOULD ALSO BE GIVEN TO DESIGNATING A LEAD SERVICE FOR TEST AND EVALUATION OF CERTAIN CLASSES OF SYSTEMS (E.G., ARMY FOR SURFACE-TO-AIR MISSILES AIR FORCE FOR AIR-TO -SURFACE MISSILES, NAVY FOR AIR-TO-AIR MISSILES).** CJCS, ASSISTED BY OSD, WILL EXAMINE THE FEASIBILITY OF ELECTRONICALLY LINKING SERVICE TRAINING RANGES IN CONUS.

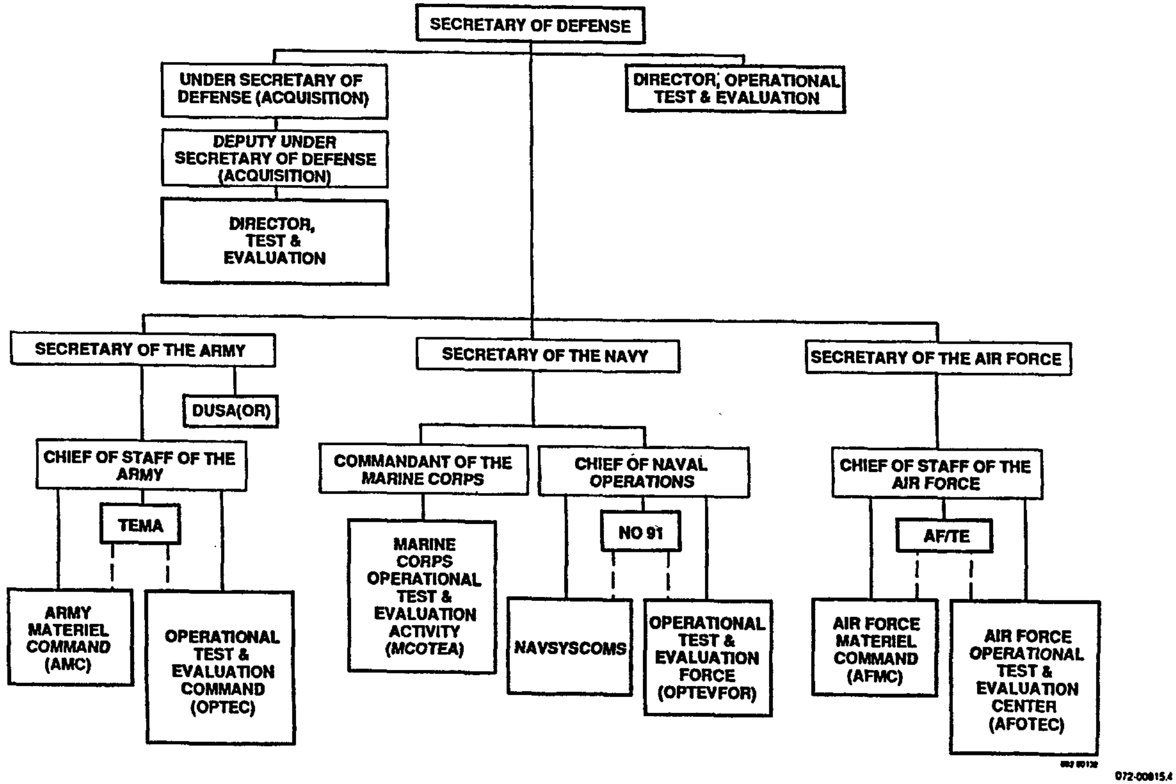
2--August 10, 1993

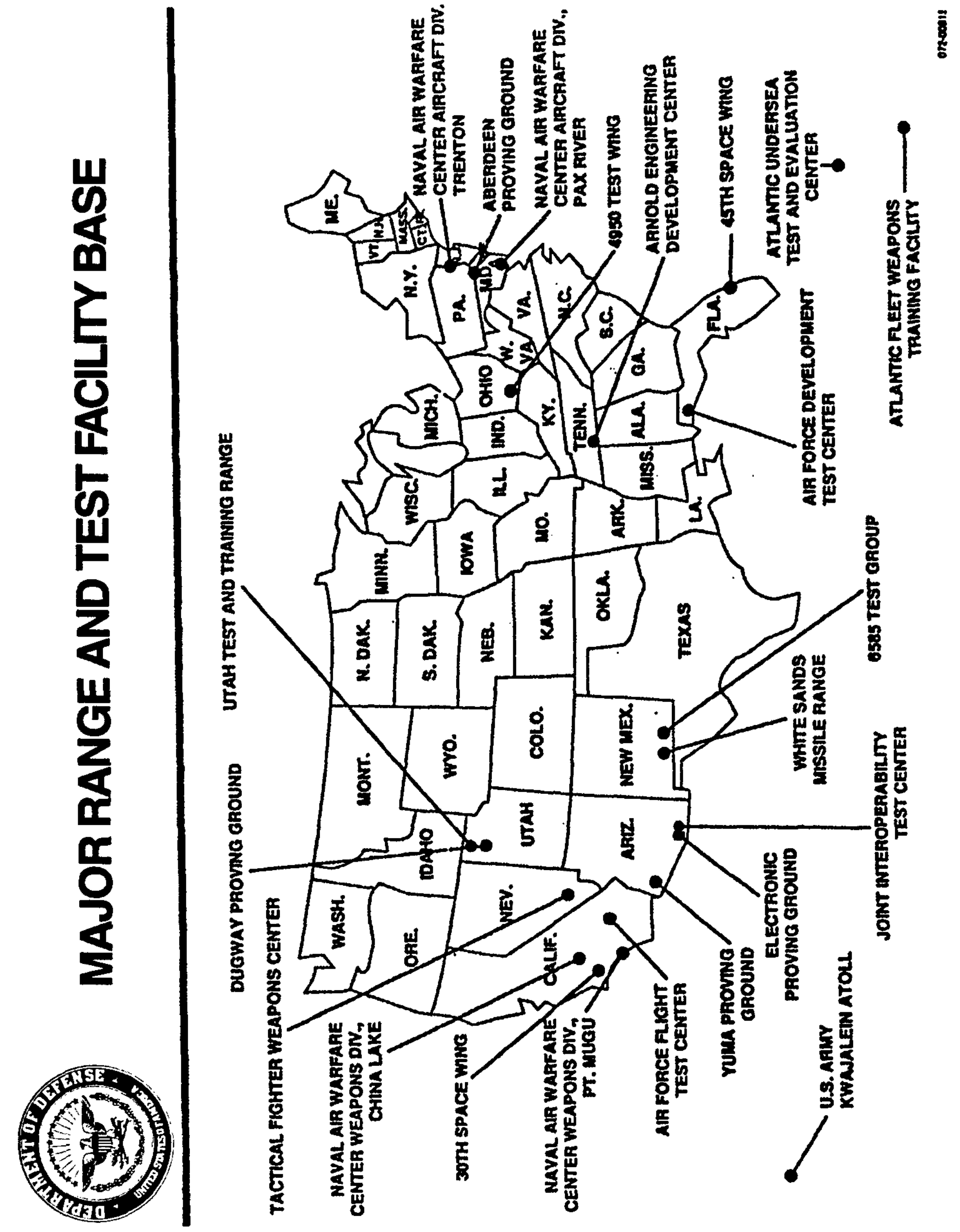
OUTLINE

- ☞ **BACKGROUND**
- CONSIDERATIONS AND RECOMMENDATIONS
  - DESIGNATION OF AN EXECUTIVE AGENT AS A SINGLE MANAGER FOR TEST & EVALUATION RANGES
  - DESIGNATION OF A LEAD SERVICE FOR TEST & EVALUATION OF CERTAIN CLASSES OF SYSTEMS
  - LINKING OF CONUS TEST & EVALUATION RANGES ELECTRONICALLY WITHIN BROAD GEOGRAPHIC AREAS
  - FEASIBILITY OF ELECTRONICALLY LINKING SERVICE TRAINING RANGES IN CONUS (CJCS)

3—August 10, 1993

T&E COMMUNITY




MAJOR RANGE AND TEST FACILITY BASE
DUGWAY PROVING GROUND
UTAH TEST AND TRAINING RANGE
TACTICAL FIGHTER WEAPONS CENTER
NAVAL AIR WARFARE CENTER WEAPONS DIV., CHINA LAKE
30TH SPACE WING
NAVAL AIR WARFARE CENTER WEAPONS DIV., PT. MUGU
AIR FORCE FLIGHT TEST CENTER
YUMA PROVING GROUND
ELECTRONIC PROVING GROUND
JOINT INTEROPERABILITY TEST CENTER
U.S. ARMY KWAJALEIN ATOLL
WHITE SANDS MISSILE RANGE
6585 TEST GROUP
AIR FORCE DEVELOPMENT TEST CENTER
ATLANTIC FLEET WEAPONS TRAINING FACILITY
ATLANTIC UNDERSEA TEST AND EVALUATION CENTER
45TH SPACE WING
ARNOLD ENGINEERING DEVELOPMENT CENTER
4950 TEST WING
NAVAL AIR WARFARE CENTER AIRCRAFT DIV., PAX RIVER
ABERDEEN PROVING GROUND
NAVAL AIR WARFARE CENTER AIRCRAFT DIV. TRENTON
WASH.
ORE.
CALIF.
NEV.
IDAHO
MONT.
WYO.
UTAH
ARIZ.
COLO.
NEW MEX.
N. DAK.
S. DAK.
NEB.
KAN.
OKLA.
TEXAS
MINN.
IOWA
MO.
ARK.
LA.
WISC.
ILL.
MISS.
MICH.
IND.
KY.
TENN.
ALA.
OHIO
W. VA.
GA.
FLA.
S.C.
N.C.
VA.
MD.
PA.
N.Y.
VT.
N.H.
MASS.
CT.
ME.

MRTFB FACILITIES

- MORE THAN T&E -- PERFORM A SPECTRUM OF FUNCTIONS
- **PURE T&E** -- EDWARDS AFB, YUMA PROVING GROUNDS, ARNOLD ENGINEERING DEVELOPMENT CENTER
- **TEST AND TRAINING**
  - UTAH TEST AND TRAINING RANGE
  - WHITE SANDS--HOLLOMAN--FORT BLISS COMPLEX
  - ATLANTIC FLEET WEAPONS TRAINING FACILITY
  - NAWC at PMTC
  - KWAJALEIN MISSILE RANGE
  - AFDTC at EGLIN AFB
  - TACTICAL FIGHTER WEAPONS CENTER
- **OPERATIONAL RANGES SUPPORTING TEST FUNCTIONS**
  - ESMC (45 th SPACE WING) at PATRICK AFB
  - WSMC (30 th SPACE WING) at VANDENBERG AFB
- **FULL SPECTRUM: LAB--TEST--SERVICE ENGINEERING**
  - NAWC at PAX RIVER and CHINA LAKE
  - ABERDEEN PROVING GROUND

6--August 10, 1993

BASIC CONCEPT

- DIFFERENTIATE T&E, TEST SUPPORT, AND INVESTMENT MANAGEMENT
  - **T&E**
    - TEST PLANNING, CONDUCT, ANALYSIS AND REPORTING
    - RELATES TO SPECIFIC ACQUISITION PROGRAMS
    - INTEGRAL PART OF DEVELOPMENT PROCESS
    - WILL CONTINUE TO BE EXECUTED BY THE SERVICE RESPONSIBLE FOR ACQUISITION
  - **TEST SUPPORT**
    - FUNCTIONS RELATED TO OPERATING TEST FACILITIES, INSTRUMENTATION AND DATA COLLECTION
  - **INVESTMENT MANAGEMENT**
    - FUNCTIONS PERFORMED TO PLAN, BUDGET AND EXECUTE INFRASTRUCTURE INVESTMENTS, DIVESTITURE OF CAPITAL ASSETS, AND INTERLINKING RANGES
    - CAN BE MANAGED SEPARATELY FROM TEST SUPPORT

7--August 10, 1993

FUNCTIONS WHICH REMAIN WITH FIELD ACTIVITIES

- BASE OPERATING SUPPORT (BOS)
- REAL PROPERTY MAINTENANCE ACTIVITIES (RPMA)
- USER REIMBURSEMENTS
- TEST AND EVALUATION PERSONNEL
- OPERATIONAL WORKLOAD MANAGEMENT AT INDIVIDUAL ACTIVITIES
- MINOR UPGRADES
- ENVIRONMENTAL COMPLIANCE

8--August 10, 1993

OUTLINE

- BACKGROUND
- **CONSIDERATIONS AND RECOMMENDATIONS**
  - DESIGNATION OF AN EXECUTIVE AGENT AS A SINGLE MANAGER FOR TEST & EVALUATION RANGES
  - ☞ **DESIGNATION OF A LEAD SERVICE FOR TEST & EVALUATION OF CERTAIN CLASSES OF SYSTEMS**
  - LINKING OF CONUS TEST & EVALUATION RANGES ELECTRONICALLY WITHIN BROAD GEOGRAPHIC AREAS
  - FEASIBILITY OF ELECTRONICALLY LINKING SERVICE TRAINING RANGES IN CONUS (CJCS)

9—August 10, 1993

LEAD SERVICE OPTIONS

- **EXAMPLES** OF ALTERNATIVES
  - ARMY: SURFACE-TO-AIR MISSILES
  - AIR FORCE: AIR-TO-GROUND MISSILES
  - NAVY: AIR-TO-AIR MISSILES
- **FUNDAMENTAL PREMISE**: SERVICE RESPONSIBLE FOR ACQUISITION MUST BE SERVICE RESPONSIBLE FOR T&E
  - TESTING IS AN INTEGRAL PART OF DESIGN SUPPORT AND DEVELOPMENTAL PROCESS
- IF ACQUISITION PROCESS IS CHANGED TO DESIGNATE LEAD SERVICES FOR ACQUISITION, T&E RESPONSIBILITY WILL FOLLOW.

10–August 10, 1993

RECOMMENDATION

- RECOMMEND AGAINST DESIGNATION OF LEAD SERVICES FOR TEST AND EVALUATION OF CERTAIN CLASSES OF SYSTEMS

11--August 10, 1993

OUTLINE

- BACKGROUND
- **CONSIDERATIONS AND RECOMMENDATIONS**
  - **DESIGNATION OF AN EXECUTIVE AGENT AS A SINGLE MANAGER FOR TEST & EVALUATION RANGES**
  - DESIGNATION OF A LEAD SERVICE FOR TEST & EVALUATION OF CERTAIN CLASSES OF SYSTEMS
  - LINKING OF CONUS TEST & EVALUATION RANGES ELECTRONICALLY WITHIN BROAD GEOGRAPHIC AREAS
  - FEASIBILITY OF ELECTRONICALLY LINKING SERVICE TRAINING RANGES IN CONUS (CJCS)

EXECUTIVE AGENT OPTIONS

- INTERSERVICE BOARD OF DIRECTORS
- JOINT COMMAND FOR TEST RANGES AND FACILITIES
- LEAD SERVICE
- OSD--MANAGEMENT ROLE, IN ADDITION TO OVERSIGHT FUNCTION
- NEW AGENCY

13--August 10, 1993

EXECUTIVE AGENT OBJECTIVES (TEST SUPPORT)

- **POOL RESOURCES AMONG SERVICES**
  - AVOID DUPLICATION
  - CONSOLIDATE TO "CRITICAL MASS"
  - CAPITALIZE ON EXISTING CAPABILITIES
  - COORDINATED DIVESTITURE/RESTRUCTURE ACTIONS
  - JOINT PRIORITIZATION - FOREIGN MATERIEL, INVESTMENTS
  - "RELIANCE" ON AVAILABLE RESOURCES BY ANY SERVICE
- **IMPLEMENT POLICY**
  - STREAMLINED CHAIN OF COMMAND
  - IMPROVED VISIBILITY/FEEDBACK
  - RESOLUTION OF INTERLINKING OPERATIONAL ISSUES
  - COORDINATED MASTER PLANNING
  - PERMANENT ANALYSIS SUPPORT (WORKLOAD, UTILIZATION)
- **OPTIMIZE RETURN ON INVESTMENT FOR T&E ACQUISITIONS**
  - JOINT PROGRAM OFFICE
  - MASTER PLAN (INVESTMENTS, CONSOLIDATIONS, DIVESTITURES)

14—August 10, 1993

JPO INVESTMENT MANAGEMENT FUNCTIONS

- FUNCTION AS ACQUISITION OFFICE FOR INSTRUMENTATION, TARGETS, and THREAT SIMULATORS (ITTS)
- RDT&E SUPPORT PLATFORMS (AIRCRAFT, SHIPS) INVESTMENT
- INSTALLED SYSTEM AND HARDWARE-IN-THE-LOOP DEVELOPMENT AND PROCUREMENT
- MODELING AND SIMULATION GENERIC ARCHITECTURE PER DMSO MASTER PLAN
- MANAGE DEPOT MAINTENANCE OF COMMON RANGE ASSETS
- MAINTAIN INTERLINKING MASTER PLAN
- PROVIDE INSTRUMENTATION SUPPORT TO TRAINING

15—August 10, 1993

PROPOSED MANAGEMENT STRUCTURE

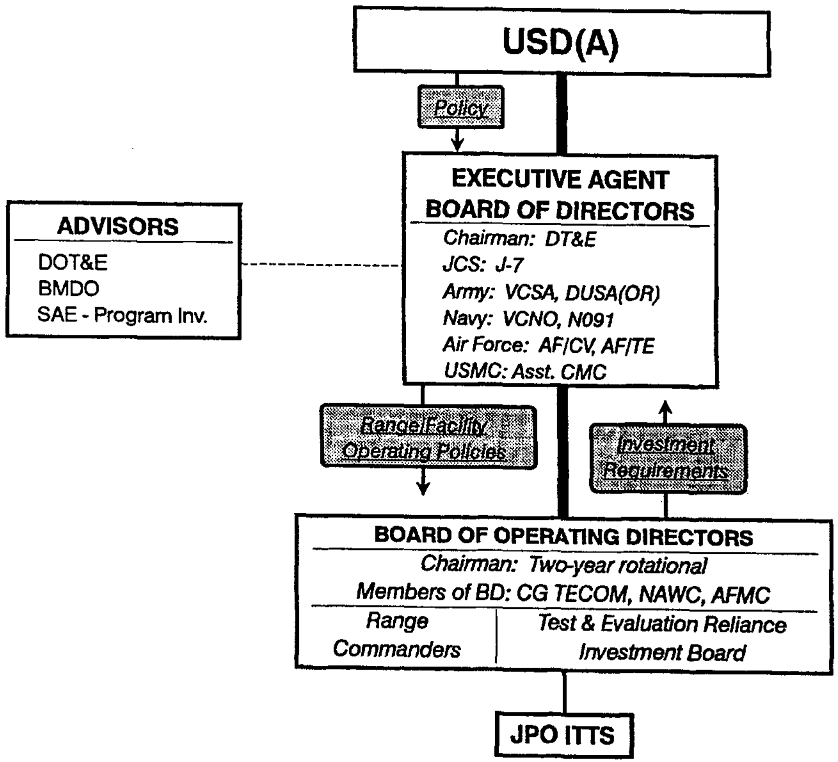
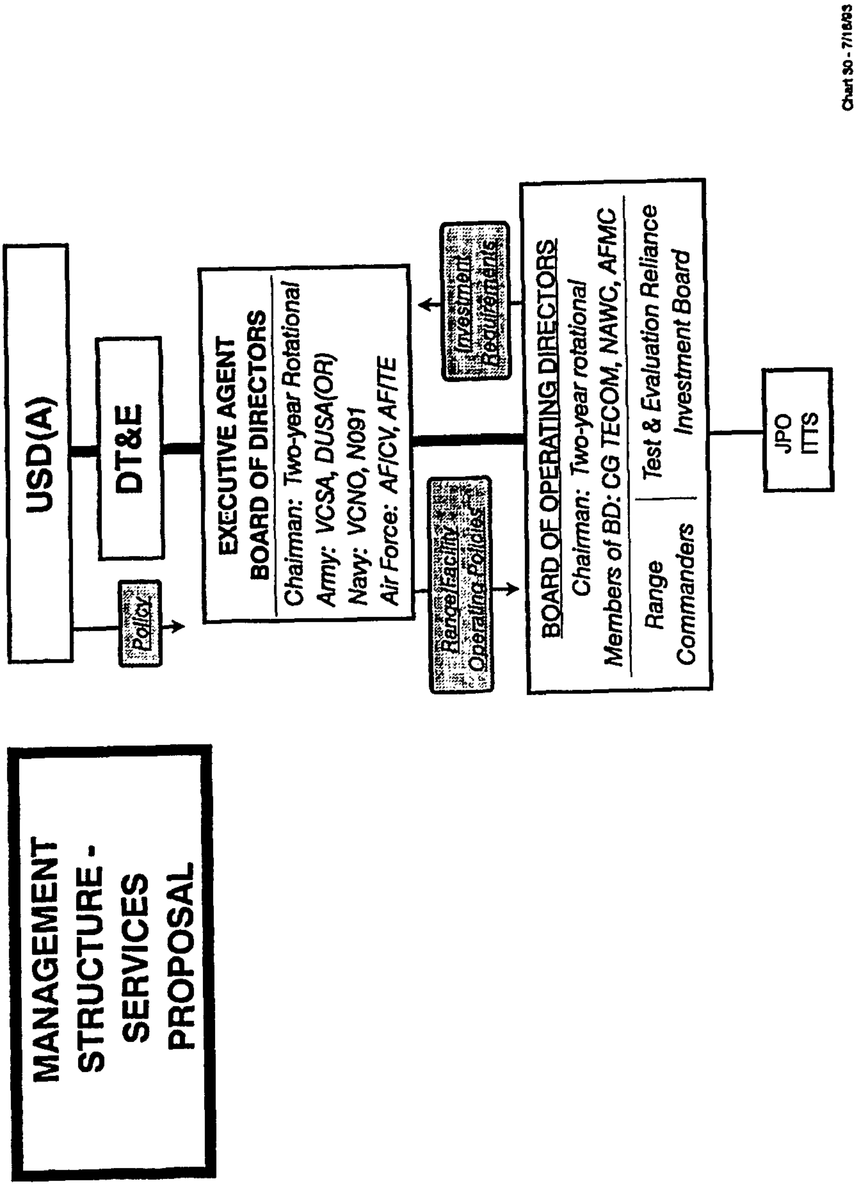


Chart 37 - 8/11/93


MANAGEMENT STRUCTURE - SERVICES PROPOSAL
USD(A)
Policy
DT&E
EXECUTIVE AGENT BOARD OF DIRECTORS
Chairman: Two-year Rotational
Army: VCSA, DUSA(OR)
Navy: VCNO, N091
Air Force: AF/CV, AF/TE
Range/Facility Operating Policies
Investment Requirements
BOARD OF OPERATING DIRECTORS
Chairman: Two-year rotational
Members of BD: CG TECOM, NAWC, AFMC
Range Commanders
Test & Evaluation Reliance Investment Board
JPO ITTS
Chart 30 - 7/16/93

BUDGET CONSIDERATIONS

- CONSOLIDATE MODERNIZATION & INVESTMENT
- JOINT PROGRAM OFFICE (JPO) EXECUTES FUNDS NECESSARY TO MANAGE INVESTMENTS
- IMPROVE VISIBILITY INTO MILCON and PROGRAMMATIC INVESTMENT
- SEEK ADDITIONAL (NON-DoD) REIMBURSABLE CUSTOMERS TO IMPROVE UTILIZATION

17–August 10, 1993

RECOMMENDATIONS

- **APPROVE EXECUTIVE AGENT CONCEPT**
  - BOARD OF DIRECTORS, RESPONSIBLE FOR:
    - TEST SUPPORT INFRASTRUCTURE REQUIREMENTS
    - INFRASTRUCTURE OPERATING POLICIES
  - DT&E AS CHAIRMAN
  - VICE CHIEFS, JCS(J-7), T&E PRINCIPALS AS MEMBERS
  - JCG(T&E) COMES UNDER LINE MANAGEMENT
- **ESTABLISH AN INSTRUMENTATION, TARGETS, and THREAT SIMULATORS (ITTS) JOINT PROGRAM OFFICE -- 1 JAN 94**
  - REPORTS TO JCG(T&E)
  - STAFFED AND FUNDED BY THE SERVICES
  - EXECUTE PROGRAMS FUNDED BY SERVICES
- **POOL MANAGEMENT OF TEST INVESTMENT BUDGETS BEGINNING FY94**

18--August 10, 1993

OUTLINE

- BACKGROUND
- **CONSIDERATIONS AND RECOMMENDATIONS**
  - DESIGNATION OF AN EXECUTIVE AGENT AS A SINGLE MANAGER FOR TEST & EVALUATION RANGES
  - DESIGNATION OF A LEAD SERVICE FOR TEST & EVALUATION OF CERTAIN CLASSES OF SYSTEMS
- ☞ **LINKING OF CONUS TEST & EVALUATION RANGES ELECTRONICALLY WITHIN BROAD GEOGRAPHIC AREAS**
- ☞ **FEASIBILITY OF ELECTRONICALLY LINKING SERVICE TRAINING RANGES IN CONUS (CJCS)**

19—August 10, 1993

TRAINING RANGES

- CONTINUING TO WORK
- WILL DEVELOP JOINT PROCEDURES

20--August 10, 1993

DEFENSE SCIENCE BOARD
1992 SUMMER STUDY

DDR&E, T&E COMMUNITIES, AND SERVICES SHOULD

- FULLY INTERNET TRAINING RANGES, TEST FACILITIES, LABORATORIES, SERVICE SCHOOLS, AND INDUSTRY –MAKE THEM **DISTRIBUTED INTERACTIVE SIMULATION** (DIS) COMPATIBLE
- INCORPORATE STANDARDS AND PROTOCOLS INTO ALL DEVELOPMENTS AND PROCUREMENTS - ENHANCE THE **ADVANCED DISTRIBUTED SIMULATION** (ADS) ENVIRONMENT AND ITS USE
- ESTABLISH AND ENFORCE STANDARDS AND PROTOCOLS TO FACILITATE INTEROPERABILITY AND REUSABILITY OF ADS TOOLS AND TECHNOLOGIES IN TRAINING AND MATERIEL DEVELOPMENT

21–August 10, 1993

TEST RANGE INTERLINKING

- **TECHNICAL FEASIBILITY DETERMINED FOR TEST AND TRAINING RANGES**
- **FUNCTIONS TO BE PERFORMED BY THE EXECUTIVE AGENT**
  - DEVELOP MASTER PLAN TO SATISFY TEST REQUIREMENTS
- **FUNCTIONS TO BE PERFORMED BY THE JOINT PROGRAM OFFICE**
  - TEST SYSTEM INVESTMENT CONTROL
  - COMMON ACQUISITIONS AND DEPOT SUPPORT

22–August 10, 1993

TEST RANGE INTERLINKING

RECOMMENDATIONS

- **SECDEF POLICY: USE DIS-BASED STANDARDS AND PROTOCOLS IN ALL INVESTMENTS AND UPGRADES**
- **INTERSERVICE BOARD OF DIRECTORS TO DEFINE REQUIREMENTS**
- **ACCOMPLISH TECHNICAL INTERLINKING**
  - ITTS JPO IS IMPLEMENTING ORGANIZATION
  - PROGRAMMED and BUDGETED IN SERVICE PROGRAM ELEMENTS
  - INTERLINK HIGHEST PAYOFF NODES FIRST AS PROTOTYPE AND TEST BED

23--August 10, 1993

Appendix I

TEST AND EVALUATION
Reflections and Recommendations

1994

TEST AND EVALUATION
REFLECTIONS AND RECOMMENDATIONS

Charles E. Adolph

Abstract

This paper reviews some changes that have occurred over the past several years in Test and Evaluation (T&E) policy and procedures. The author makes several recommendations for legislative changes and policy and procedural improvements.

I. Introduction

The last several years have been the most dynamic for the Department of Defense since WWII. The mid-late 1980s saw the end of the Reagan buildup, the demise of the Soviet Union and the Warsaw Pact, and the end of the cold war. Today's environment is dynamic, complex, and to a large degree unpredictable. The acquisition policy is in need of reform to provide more flexibility and better align it with the times. Secondly, there is a need to make test and evaluation process improvements. I will address both issues as they relate to aircraft.

II. Policy Changes

A. Legislative Changes

As part of the acquisition reform process, changes have been recommended to the existing test statutes. A "streamlined" test statute has been proposed which will, if enacted, allow more flexibility in structuring tests without compromising the integrity of the process. The proposed statute retains the general rule that operational and live fire testing must occur before going beyond low-rate production. However, the proposal gives the Secretary of Defense the authority to modify operational and live fire testing when unreasonably expensive or impractical or when it would be unnecessary because of the acquisition strategy for the system. A prohibition on system contractor involvement would be relaxed with Secretary of Defense assurance that the integrity of the test would not be compromised. The Test and Evaluation Master Plan would define the anticipated contractor involvement. There would be no restrictions on support contractors unless they are affiliated with the system contractor.

B. Improving the Developmental and Operational Test Relationship

There is widespread recognition within the T&E community that legislative and process improvements are needed. A review made during 1993 identified three of the most pressing issues that were addressed at an International Test and Evaluation Association T&E Workshop in November 1993. The three issues identified were: harmonizing developmental and operational testing; improving the correlation between requirements and test criteria; and, internetting test and training ranges and simulations. The first two issues are addressed below.

Over the last 15-20 years, developmental testing (DT) and operational testing (OT) have become increasingly stratified. This stratification is a major (but by no means the only) contributor to an increasingly protracted and inefficient weapons system developmental process. The rigid DT/OT stovepipes that have evolved are, for the most part, artificial. Barriers, real and perceived, discourage cooperation and in some instances prohibit it. The vast majority of test objectives provide developmental insights as well as operationally relevant information. The stovepipes are also counterproductive, fostering separate, duplicative testing, separate data bases, and isolated development and use of models, simulations, and data processing. The legislation that created the office of the Director, Operational Test and Evaluation in the Office of the Secretary of Defense has exacerbated the problem, as have the legislative prohibitions on contractor support. A streamlined test statute is included in the DoD acquisition reform package. The proposed statute is a step in the right direction. It will provide more flexibility in structuring tests and in the use of contractor support without compromising the integrity of the process.

The increasing politicization of the process creates additional inefficiencies; numerous legislated test "hooks," require annual restructuring of test programs. The definition of what constitutes operational testing has become broader; today's operational effectiveness criteria typically include a large number of technical characteristics.

Given where we are today, particularly in view of the current fiscal environment, it's time for the pendulum to swing in the other direction. My specific recommendations follow. We should continue to have separate DT and OT assessments, but totally integrated planning and largely integrated use of test support tools and test conduct. This includes integrated contractor, government DT and OT use of models, digital and hardware-in-the-loop simulations, and test data reduction. This can readily be done without compromising the integrity of the process, while simultaneously realizing significant efficiencies. My friends in the Air Force might assert that many elements of what I have described are already taking place within the Air Force's Combined DT/OT Test Force (CTF). I would argue that the CTF is a small step in the right direction, but doesn't go far enough. Some elements of the support infrastructure are consolidated within a typical CTF, but the government DT community is narrowly focused on what they have defined as developmental objectives. The developmental test community in all the Services pays little, if any, attention to operational test criteria. This is one of the major problems with the way we currently do business. We need to "front load" the development and test process to be more operationally oriented. The Service DT and OT communities should develop an integrated evaluation plan for each system. Testing should be almost totally integrated, except a very limited period of dedicated OT&E. In addition, the government acquisition community needs to pay more attention to evaluation, and not use such a high percentage of its declining resources on detailed test planning, and test conduct. The manufacturer is responsible for executing most elements of the development process.

DT should be expanded to include a preliminary assessment of suitability and effectiveness. The integrated product team concept should be expanded to include the developmental and operational test community.

III. Test Process Issues

A. Improving Correlation Between Requirements and Test Criteria

The correlation between system design requirements and test criteria needs to be improved. The overall objective is enhanced consistency among requirements documents, the cost and operational effectiveness analysts (COEA), the program baseline, the test and evaluation master plan (TEMP), and test-related criteria. For every program, there should be a "transformation matrix" that links these documents. Test scenarios and threat arrays must be defined in a manner that is consistent with programmatic documentation. There should be constraints on the test community's latitude to "invent" test criteria to insure consistency with user-validated requirements. The control volume of the evaluation (not necessarily the test) needs to be limited to the system undergoing testing. Where practicable, test results should be used to validate the performance assumptions underlying the COEA.

B. Modeling and Simulation

We use simulations extensively today in support of system development and test and evaluation. Simulations and models are used at the subsystem level, in standalone, engineering-level simulations as well as for mission-level assessments. What is needed in the future is not necessarily more simulation but more effective use of simulation tools to improve the development process and to supplement and focus tests and evaluation. More effective use will also facilitate reaching agreement on basic system requirements. It will also improve performance definition, which in turn leads to testable and traceable measures of effectiveness and measures of performance. Better use of simulation tools will focus investment of test resources on critical test parameters, issues, and criteria. Finally, simulations can be valuable in test planning and for test rehearsals. To accomplish these objectives we need to better plan for the use of models and simulations (M&S), to address the validation and verification issue at the outset, to mitigate the risk associated with using imperfect simulation tools, and to invest in M&S technology. An integrated approach to M&S must be developed. There is a need to build on the credibility offered by hardware in the loop, system software in the loop, and man in the loop simulations. We also need to plan for efficiency; as an example, plan to use engineering simulators as the basis for system training simulators.

The solution to simulation validation problems lies, at least in part, in standardizing architectures, methods and procedures to the extent practical, an in using test results to validate models. This has been accomplished very effectively by the air defense threat simulation community. Hardware in the loop, software in the loop, and man in the loop simulation can be used to help in the validation process, progressing from component to subsystem to system to mission-level simulations. There is obviously risk associated with excessive reliance on M&S that do not adequately represent the system or its environment. This risk can be mitigated by using test data at every level in the validation.

There is also a need to refine, maintain and make available those simulations having wide applicability; as an example, air defense threat digital simulations. In summary, improved use of simulation will improve the efficiency of the developmental process and decrease, but not eliminate subsystem and system-level testing.

C. Software-Intensive Systems

Another challenge facing the aircraft test community is the efficient evaluation of software-intensive systems. Integration and checkout of software-intensive systems are cost and schedule risk areas associated with most of today's development and test programs - aircraft, missiles, surface ships and submarines. The software-intensive systems of today present a quantum leap in capability as well as a quantum jump in test requirements. Almost two-thirds of the flight testing on today's aircraft is related to avionics. The challenge for the designer, evaluator and operator is to get the highest benefit from the flexibility afforded by these systems, while at the same time being cost-effective. Experience over the last several years shows that we tend to grossly underestimate the testing necessary to fully develop software-intensive systems. The testing required is often two to four times higher than the initial estimate. Our ability to plan a test program is much poorer today than it was twenty years ago when we were dealing primarily with hydro-mechanical systems.

One of the most vexing issues related to testing software-intensive systems is defining test criteria: measures of performance, effectiveness, and reliability. Because of the ease of making changes, the definition of what constitutes a production-representative system is far from straightforward. Regression testing requirements are difficult to define. Test methodology for software intensive subsystems needs to be given far more attention in the test planning process.

D. Weapons System Performance Envelope

I would like to briefly discuss testing in the context of the aircraft performance envelope. A function of developmental testing is to define and expand the envelope. The T&E community is sometimes accused of having a preoccupation with testing at the fringes of the envelope. There are occasional implications that we testers establish scenarios where the system is likely to fail.

My philosophy on this issue has evolved to something like this: Once the envelope is defined, operational testers should ensure that the system works very well and reliably in the heart of the requirements envelope. Note that I said the heart of the requirements envelope; not the heart of the performance envelope. We should concentrate on most likely operational scenarios. Testers should also assess performance at the fringes of the performance envelope without becoming preoccupied with that area. Some degradation should be expected. By the way, testing at the fringes of the envelope is generally extremely resource-intensive and expensive.

E. Test Program Infrastructure

There are several issues to keep in mind when planning for the test program infrastructure. The first is the number of test support sites. Historically, many aircraft test programs have been conducted from multiple locations; some programs have been conducted primarily, if not exclusively, from a single site. The pros and cons of both approaches have been debated over the years, as has the degree of integration of contractor and government test and test support personnel. The Army, Navy and Air Force all use slightly different approaches. In today's resource-constrained environment, programs cannot afford to unnecessarily fragment test expertise and the test support infrastructure. Site selection criteria include test support infrastructure, flexibility to plan and execute the test program, and weather, particularly for high performance aircraft.

With the real-time data transmission capability that exists today, a real-time data link to the design team is highly desirable, particularly in the early stages of the program. Contractors should have the maximum flexibility and autonomy to execute the development effort. The government's role in development is primarily oversight and evaluation, not test execution. Developmental and operational testing should be almost totally integrated, as was discussed earlier. The fixed cost of the test support infrastructure is a larger cost driver than the per-flight support cost. Among other things, this implies that the support infrastructure should be geared to a multi-shift, six to seven days per week operation during the early stages of development. This is the norm for commercial certification programs and for the highly successful competitive fly-offs and demonstration-validation programs (dem-vals). These programs typically clear a complete flight envelope in a year or less.

A related issue is contractor teaming, which is the norm today in aircraft development programs. The F-22 and the V-22 feature teaming arrangements and teaming is likely for the Join Advanced Strike Technology (JAST) development efforts. In Europe, co-development consortiums have been the norm for military developments (e.g., the Tornado) as well as commercial ventures (e.g., the Airbus). These multi-organizational and multinational programs present additional technical and management challenges. One challenge is balancing test efficiency with the desire for each service, contractor organization or country, to maintain currency by participating in the test process. This must be accomplished without unnecessary duplication of test locations and test support infrastructure. One of the many problems with the V-22 test program was the attempt to operate and maintain several test sites, at the expense of not having a critical mass at any one site early in the program.

F. Next Generation Aircraft

I would like to look into my crystal ball to determine what lies ahead for military aircraft development and testing. You all know that there will be fewer platforms and weapons. In the late 50s, the 60s and early 70s, there were many new aircraft designs. In the early 70s alone there were first flights on numerous new aircraft. The number of new aircraft designs has decreased dramatically in the past 15 years, and there won't be any major new tactical, strategic, or cargo aircraft reach the engineering model development stage for the next ten years at least.

Let's take a look at the direction technology is headed in aircraft and aircraft subsystem design. The number one design criterion at the system level and for every subsystem will be "affordability," which has been the overlooked "ility" for the past twenty years. Affordability needs continuing emphasis in every technology area, both the cost of acquisition and the cost of ownership. "Stealth" will be designed into most systems. Signature reduction will be incorporated into physical size as well as the more common infrared and RF signatures. Stealthy aircraft carry their weapons internally or conformally, which implies fewer weapons and weapon types. This means less integration testing per type of store. Air vehicle designs will also incorporate lift/drag improvements, more electric subsystems, and reductions in structural weight through the use of advanced composite designs and low cost composite materials. Next generation propulsion systems will have significantly improved thrust-weight ratios and thrust-fuel specifics. There is a superb government-industry propulsion technology program with well-defined goals. It is a model technology effort. Human system interface enhancements will include improved helmet displays. Pilots' associate technologies offer the potential of reducing crew size with a corresponding reduction in acquisition and life cycle costs. Finally, the data base on human interfaces is based almost exclusively on men. There is an urgent requirement to expand the data base to include women.

The greatest challenge is in the area of avionics since avionics subsystems' costs are approaching 50 percent of the total aircraft cost. Avionics developments are rapidly evolving and diverse, influenced by technology developments in electronics in many sectors, government, industry, academia and foreign manufacturers. There is often a blurred boundary between technology development and engineering applications. The process can be characterized as evolutionary development; as an example there are often annual operating flight system software upgrades. There are extensive and complex interfaces distributed on-board and off-board, including aircraft-aircraft links, ground and satellite links.

Summary

In summary, the turbulence that exists today is not likely to subside in the foreseeable future. The acquisition community, which includes test and evaluation, must adjust and adapt to these turbulent times. Test and evaluation remains the disciplining factor in the acquisition process. We must seek out ways of exerting that discipline in a more cost-effective and efficient manner.

Index