Integration of Manned Simulation and Flight Test into Operational Testing and Evaluation

M. J. Williams* and M. A. Hartford†
McDonnell Douglas Corporation, St. Louis, Missouri

A basic outline is presented for integration of a manned simulation test program with a flight-test program to accomplish the operational testing and evaluation of a fighter aircraft or a specific fighter weapon subsystem. A description of and the requirements for an operational test are given. The intrinsic advantages and disadvantages of operational flight testing and manned simulation are explained. The McDonnell Aircraft Company's manned aerial combat simulator is briefly described along with the benefits of using a simulation to support an operational test. A method of integrating flight test and manned simulation to compensate for their inherent limitations is also presented. This integration provides for a better overall test program than would be possible using either method alone. Issues addressed include the use of a manned simulator for training test participants and as a method of significantly expanding the analytical data base. A sample test matrix and scenario are presented. Validation between simulation and flight test results is discussed. Analytical issues addressed include possible measures of effectiveness to be collected and analyzed during both simulation and flight test.

Introduction

THIS paper presents a basic outline for integrating a manned simulation test program with a flight-test program to accomplish the operational testing and evaluation of a fighter aircraft or a specific fighter weapon subsystem.

Operational flight testing is the traditional yardstick by which a new aircraft or subsystem can be evaluated before large-scale production can begin. It provides an excellent indication of how such a system can and should be used in actual operations and how effective it will be. Manned aircraft simulation has reached a point where it can significantly improve an operational flight test and add significant amounts of operational data that may not be available from flight testing alone.

This paper will examine requirements and definition of military operational testing; an overview of the McDonnell Aircraft Company (MCAIR) simulation facility that has been used to support U.S. Air Force operational tests; a detailed examination of the benefits of a manned simulation in an operational flight-test program; and, finally, the integration of an operational test simulation into an overall test program to provide comprehensive pilot training before the start of flight test and to significantly expand the analytical data base of the test.

Requirements and Definition of Military Operational Testing

Operational testing is defined in the U.S. Navy's *Operational Test Director's Guide* (Norfolk, VA, 1978) as:

1) Exercising a system or equipment under conditions that simulate the expected operational combat environment as closely as possible.

2) Recording sufficient data during the exercise to document all operationally significant system or equipment characteristics.

The same source defines operational evaluation as the analysis and interpretation of data from an operational viewpoint, for the purpose of predicting the operational effectiveness and operational suitability of a system. These general definitions are suitable for any military operational test program.

There are a number of basic elements in any well-planned and well-conducted operational test. The first element is a realistic scenario or operational environment in which to conduct the test. This should include all of the friendly and adversary assets that are expected to be present on the battlefield that will directly impact the test article. It is especially important that the adversaries are present in realistic quantities and that they fight back. A scenario that might be suitable for the operational test of a new or improved air defense fighter or subsystem is illustrated in Fig. 1. Realistic tactics must be employed by both sides and should replicate those expected on the battlefield to the greatest extent possible.

A good operational test should also be designed to collect specific measures of effectiveness (MOE's) that will allow the analysts to quantify the significant operational characteristics of the test article. Some MOE's suitable for the evaluation of a

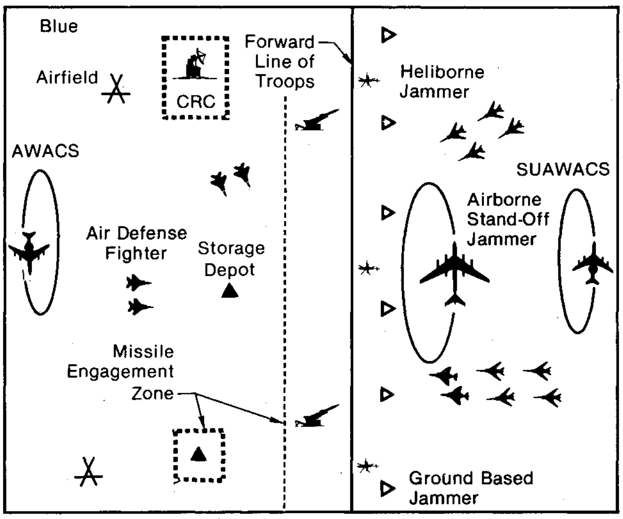

Fig. 1 Typical Central European scenario.

Presented as Paper 87-2924 at the AIAA/AHS/ASEE Aircraft Design Systems and Operations Meeting, St. Louis, MO, Sept. 14–16, 1987; received Dec. 14, 1987; revision received Oct. 6, 1988.

*Lead Engineer, McDonnell Aircraft Company.
†Engineer, McDonnell Aircraft Company. Member AIAA.

new air defense fighter or subsystem are shown in Table 1. To determine whether these operational characteristics represent actual improvements in operational performance, a point of comparison, or baseline, is required. This can be obtained by evaluating a currently fielded alternative to the test article. For example, a good baseline for a new air defense fighter might be the currently fielded F-15 or F-16.

Finally, the test should be designed to provide a quantity of data sufficient to draw meaningful conclusions regarding the benefits of the test article. In general, this will mean numerous repetitions of the same basic scenarios. At the heart of such a design might be a factorial test matrix that allows for the controlled variability of key elements of the scenario. A factorial test matrix indicates the number of runs to be made within each configuration of independent variables. An example of a matrix for the evaluation of a new fighter aircraft or subsystem with different weapon loadings, different levels of electronic warfare, and different force mixes (i.e., two F-15's vs two Red Fighters plus four Red Bombers) is shown in Table 2.

The output of an operational test usually consists of an evaluation report and a tactics guide. The evaluation report attempts to answer the questions of operational effectiveness and operational suitability of the test article. It is then directed to the decision-makers, who must determine whether the test article warrants termination, further testing, or actual procurement. The tactics guide is directed to potential users of the system and contains preliminary suggestions regarding how the new system should be used in combat.

McDonnell Aircraft Company's Flight Simulation Capability

Current state-of-the-art air-to-air combat simulators, such as MCAIR's, allow pilots to engage in realistic multi-aircraft encounters. These simulators are typically comprised of fighter cockpit domes, simplified crew stations, ground-controlled intercept (GCI) or airborne warning and control system (AWACS) controller stations, a test director station, an overview area, and a computing system. The MCAIR manned air combat facility is shown in Fig. 2. The number of each of these components to be used is determined by the most demanding scenario in the simulation.

The heart of the simulator is the computing system, which calculates the dynamics and performance of each aircraft and missile. It also determines the relative geometry of aircraft and missiles and computes weapon flyouts, communication and radar electronic warfare signals, radar detections, and radar warning receiver (RWR) detections, and models the performance of the system being tested from accurate software models.

Table 1 Possible measures of effectiveness

- Percent of bombers reaching target
- Red losses
- Blue losses
- Simultaneous firings at same target
- Number of voice transmissions

The system being tested is usually modeled in the high-fidelity fighter cockpit domes. Pilots who are testing the system are located at the center of the dome in an accurate reproduction of the fighter cockpit. Information from the computer is used to drive displays within the cockpit. Sound generators simulate engine, speed brake, and weapon launch noises. Detailed terrain surfaces, which present pilots with altitude and velocity references, images of other aircraft, and weapons in flight are projected on the dome. This visual presentation to the pilot is shown in Fig. 3.

The limited availability of these domes usually restricts the number of manned aircraft that can be simulated. Additional aircraft can be modeled using a lower-fidelity simplified crew station, consisting of a throttle, control stick, and a large CRT, where a simplified visual field-of-view, radar scope, radar warning receiver (RWR) scope, and other sensor displays are presented. This crew station is illustrated in Fig. 4. Computer information is also displayed on a simplified GCI or AWACS controller screen that allows for experienced GCI controllers to participate in the simulation. Computer-controlled, unpiloted digital aircraft can also be utilized.

An overall picture of the engagement can be presented to a test director station, where the test director can start, restart, or terminate engagements. This overall picture may also be

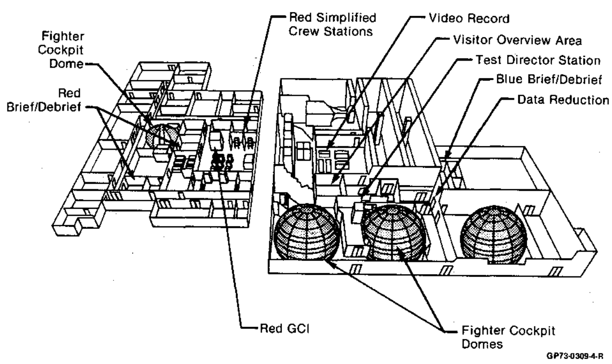

Fig. 2 MCAIR flight simulation facility.

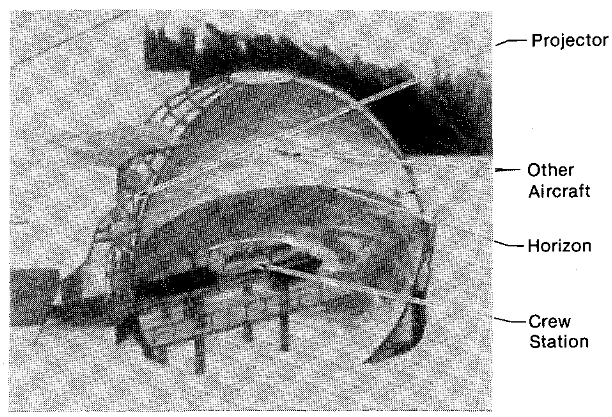

Fig. 3 Fighter cockpit dome.

Table 2 Possible test matrix

Force	With test system				Without test system			
	ECM		No ECM		ECM		No ECM	
mix	AIM-7	AMRAAM	AIM-7	AMRAAM	AIM-7	AMRAAM	AIM-7	AMRAAM
2v2+4	10	10	10	10	10	10	10	10
2v4+4	10	10	10	10	10	10	10	10
2v2+6	10	10	10	10	10	10	10	10
Total of 240 runs								

displayed in an overview room where analysts and visitors can watch the engagements unfold. Displays of the system being tested, aircraft radar displays, other sensor displays, and threat and friendly voice communications may also be presented in this overview room. The computer can also calculate MOE's and dump pertinent information onto data tapes so that analysts may compute the MOE's off-line or on another computer.

Benefits of Simulation for Operational Testing

Based on the requirements of an operational test and the capabilities of a state-of-the-art flight simulator, a manned simulator can improve procedural training and tactics development and can expand the data base. A summary of the improvement in procedural training, tactics development, and data-base expansion is shown in Fig. 5.

The simulator is especially beneficial as a procedural trainer in an operational test allowing the pilot to learn the switchology, mechanizations, limitations, and use of the system being tested. In the simulator, the pilot can concentrate on learning the system to be tested without continuous distractions that are present in a real aircraft. There is no danger of fuel exhaustion, mid-air collisions, engine fires, getting lost, hitting the ground, or loss of certain aircraft functions. Procedural training can be further enhanced by "freezing" the simulator so that aircraft can stand still in mid-air, allowing the pilot to operate the test article without the need to fly the aircraft.

In the simulator, a pilot may become familiar with the test article in a fraction of the time it would take in an aircraft. Scenarios can be brought up or restarted almost instantaneously in the simulator, whereas training in an aircraft usually requires several flight hours spent flying to and from a test range, in air traffic control, or refueling from a tanker. An additional difficulty with using an aircraft as a procedural trainer is that a failure in the test article or other critical system can waste an entire day of training. Significant time savings and learning advantages provided by the simulator result from flying engagements back to back by rotating crews through the system. In the joint tactical information distribution system (JTIDS) and the initial operational test and evaluation (IOT&E), roughly 21 runs per day were run in the simulator using three different crews, whereas only 2 or 3 were run in the flight test with one crew.

Using the simulator as a procedural trainer also provides direct interaction with experts. During procedural training, an engineer who helped design the system, a human factors engineer familiar with the system, and a pilot already familiar with the system can stand beside the training pilot and show him or her pertinent switch actuations. The experts see what the pilots are doing; therefore, they are in an excellent position to point out mistakes, discuss objectives directly with the pilot, and aid in developing the optimal operational use of the system.

A less obvious benefit of simulation is tactics development, where pilots attempt to identify and exploit any tactical advantages that the test article may give them while they are engaged in realistic test scenarios. The simulator removes any artificial limitations placed on the pilot's flying, whereas aircraft flying at a test range have limited flying areas, maximum Mach numbers, and minimum altitude restrictions. These safety, environmental, and range restriction limitations severely hamper a pilot's ability to fully explore the system's tactical implications. For example, in evaluating a new sensor that provides pilot information without the danger of detection, a pilot may want to utilize the sensor feature along with a near-sonic, extremely low-altitude flight profile to avoid being detected by the enemy. This flight profile would violate several range restrictions if tactical training occurred in an aircraft; however, if a simulator is used, a pilot's flying is only limited by his ability and the aircraft's performance.

Yet another benefit of the simulator for tactical training is the immediate feedback from new tactics. In an aircraft, pilots must fly to and from a range and then wait about an hour before debriefing. The pilots must then wait until the next mission, usually the next day, to try any new tactics. The simulator provides a novel approach to tactics development. A pilot can brief, fly a mission, and debrief within an hour. This rapid learning cycle allows pilots to investigate tactics while the success or failure of previous tactics are still fresh in his mind. Since no maintenance or inspections are required between mission in the simulator, two or three teams can train during the same day. While one team is briefing or debriefing, another team can fly, effectively doubling or tripling the number of training sorties generated. If a pilot feels that the opposing team is using unrealistic tactics, the test director may allow members of opposing teams to meet and discuss rules of en-

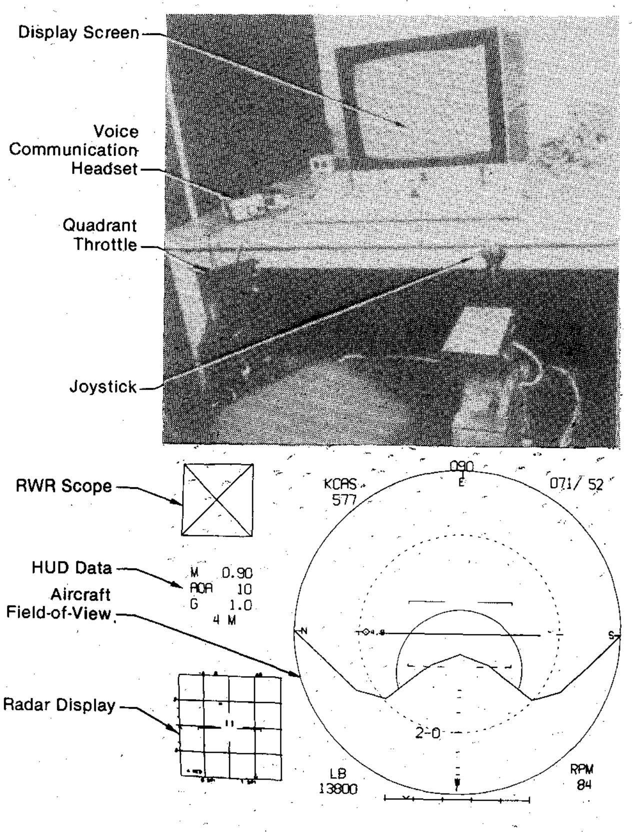

Fig. 4 Simplified crew station and display.

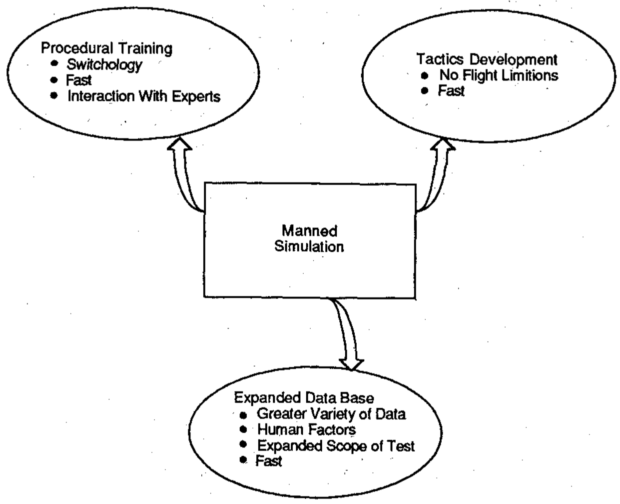

Fig. 5 Benefits of manned simulation.

gagement immediately after the incident occurs. If tactics development is done in the simulator, a pilot can generate dozens of sorties per day. In this unique dynamic learning environment, even subtle tactical implications of the test article can be reviewed in detail immediately after engagements have been flown.

Although lower stress and work load are an advantage for the procedural trainer, they are a disadvantage in the tactics development phase. This lower stress and work load can cause slight distortions in tactics employment, such as a pilot taking unnecessary risks. These distortions can be controlled to some extent by establishing appropriate rules of engagement, which are carefully monitored by the test director. Simple pseudo-punishment techiques may also be employed. For example, if a pilot is killed by colliding with the ground, he may have to sit in total darkness and listen to the other pilots until the next engagement starts.

From the analyst's point of view, the greatest benefit of the simulator to an operational test is the ability to greatly expand the data base in terms of data quantity, types of data, and fidelity through the use of an expanded scope simulation utilizing actual threat aircraft. A larger volume of data can be created in the simulator by rotating teams, which increases sortie generation by a factor of two or three. This, and the fact that aircraft do not have to be flown to and from a test range, speed the generation of test engagements significantly. In the flight test, generation of sorties may also be hindered by aircraft availability and reliability.

Another benefit of using the simulator is the almost unlimited amount of data available because the simulator computer calculates all of the aerodynamic effects, control laws, radar and infrared and other sensor inputs, weapon parameters, and relative positions and velocities of every aircraft, jammer, and surface-to-air missile many times every second. Any of these data that are of interest to analysts may be used to calculate MOE's. For a system with even a fraction of this data-gathering capability in a flight test, aircraft and the test range would have to be extensively instrumented.

Extensive human factors data that are difficult, if not impossible, to collect in an actual aircraft are available in the simulator. Switch actuation frequencies and response times can be collected via the simulator computer data. Video cameras recording pilot responses and movement can easily be installed in the simulator. Eye movement trackers and other test equipment requiring a large amount of support equipment can be installed relatively easily in the simulator. Much of this equipment cannot be used in aircraft due to space and weight limitations. The use of such equipment in the simulator can provide the analyst with expanded data on the human factors implications of the test article.

Use of the simulator in an operational test can enlarge the data base by expanding the scope of the test through utilization of models of equipment not available and tactics not allowed in a flight test. In a flight test, the interactions of surrogate equipment and limited tactics with the test article may be significantly different than actual combat conditions. A flight test also imposes an artificial environment of tactics constraints, such as disallowing head-on gun passes and not allowing an aircraft with extremely low fuel to pursue an enemy. An expanded scope simulation allows pilots to engage in combat against current and postulated threat systems utilizing both current and future friendly aircraft defense technologies. An operational test attempts to evaluate the test article against the latest or future model threat systems. Current friendly aircraft systems (i.e., bombers, radars, and missiles) and future aircraft systems (i.e., the test article) are actually used in a flight test; however, there is no way to accurately model any of the threat systems. In the simulator, models of the latest or future threat aircraft, missiles, communication networks, and electronic warfare jammers can be modeled relatively accurately. For example, the primary threat in the flight test may be a surrogate F-4 representing a MiG 29. This is certainly not a very good substitute for the latest threat aircraft, which may be several generations beyond the F-4 in almost every capability. Similarly, surrogate missiles are used in a flight test. An AIM-7F or IHAWK may not represent any of the threat missiles in service or their follow-ons.

The simulator can accurately model interactions of the test article with other future friendly systems, such as a new missile not yet operational. An additional area where the simulator can increase the quality of the operational test is in the area of electronic warfare. The simulator can predict the jamming resistance of a new radio or radar by modeling postulated threat jammers, which will almost certainly appear if the test article becomes operational. Estimates of a system's jamming resistance may not be attainable in flight test because the current friendly inventory may lack a surrogate jammer.

Simulation Design and Integration

An operational test simulation should be designed and integrated into the flight-test program to accomplish the basic objectives of procedural training, tactics development, scenario development, and expansion of the scope of the overall test. Additionally, a portion of the simulation should duplicate the conditions of operational flight testing as closely as possible. Data taken from these runs, when directly compared to flight-test results, establish the credibility of simulator results and assist in the analysis and interpretation of flight test and simulation operational test data. Four objectives and tasks require the division of the simulation effort into four distinct phases: 1) procedural training phase, 2) tactics development phase, 3) flight-test comparability phase, and 4) expanded scope phase. The rest of this paper will discuss the design of each of these phases and their integration into a test program.

Figure 6 illustrates the integration of the four simulation phases with the flight test. The first phase is the procedural training phase. A good simulation design will allow test pilots a sufficient amount of procedural training in the use of the new system or aircraft to be able to proceed to operational testing. The amount of time required will vary with pilot experience and ability as well as with the complexity of the system to be learned. In practice this can vary from only a few hours of simulator time to a more complex training scheme lasting days or even weeks. Since pilots learn at different rates, it is a good idea not to fix hard-and-fast schedules for procedural training. Rather, the test conductors should be flexible and terminate procedural training early if pilots find the system easier to learn than originally anticipated or extend it somewhat if the pilots are having learning difficulties. Training scenarios should be based on a planned learning progression and approach the actual test scenarios in complexity and scope. Other participants in the simulation such as adversary pilots and GCI controllers can also begin learning the operation of the crew stations and equipment at this time. During procedural training enough data can be taken to verify selection and computation of the measures of effectiveness. Some human factors data may also be taken that address difficulties pilots have in learning the operation of the new system. At the completion of the procedural training phase, the tactics development phase begins.

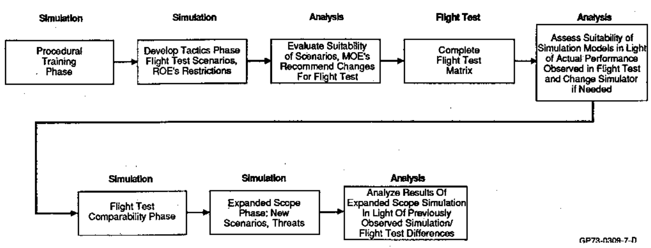

Fig. 6 Flight-test–flight-simulation integration scheme.

After the basics of any new system have been mastered, a significant amount of time is still required for the pilots to develop and understand all of the tactical implications of the system. Ideally, pilots should be given enough simulator time to "peak out" on the tactics learning curve before the start of the operational test. Tactics development is complicated, requiring much trial and error; thus, it is not practical to provide enough time for this to be fully accomplished. However, a significant amount of simulation time should be allowed before any operational testing is done to give the pilots a reasonable start at tactics development. A careful examination of pilot experience and the tactics implication of the new system should be made to determine the amount of time that should be devoted to tactics development. The scenarios used in this phase of simulation should be comparable in complexity to the planned flight-test scenarios. In fact, much is gained from using scenarios that duplicate flight-test conditions and by collecting operational type data. Simulator trials run during this phase in the baseline configuration provide for analyses of the data to determine whether there is adequate sensitivity between MOE values and the major test variables. If, for example, the percent of enemy bombers reaching target is usually zero for both the baseline and test configuration, it is a strong indication that this MOE is not varying sufficiently to show the benefits of the new system, at least with the current scenarios. A change of scenarios or a definition of new MOE's that better reflect the operational benefits of the system being tested is called for. This phase may also be used to test the suitability of pilot questionnaires and other subjective data being collected. If these require modification, it should be accomplished before the start of the flight test.

After procedural training, tactics development, and any necessary refinements to scenarios and/or MOE's, the operational flight test can begin. Because of the complexity of a flight-test program running scenarios of the intricacy shown in Fig.1, and due to the normally high demand for resources and range time, it is usually necessary to finish the complete test matrix without any significant breaks so that the flight-test resources are not tied up any longer than necessary. Since the pilots and test team will be fully occupied with the flight-test program, nothing can be done with the simulator during this time as far as testing is concerned. However, it is important that during this phase the actual performance of critical test items be monitored closely and compared to their modeled performance in the simulator. If significant performance differences exist, the test team should consider the desirability of changing certain simulator models to reflect more realistic performances. Sufficient time should be allocated in the schedule to allow such changes to be made to simulator software if the need arises.

After the completion of the flight test and after any necessary changes were made in the appropriate simulator models, the flight-test comparability phase can begin. The benefits of doing this phase may be less obvious than for the other phases of simulation, but they are nonetheless significant and important. The two basic reasons for this phase are to establish the credibility of the simulation results and to support the analysis and interpretation of those results.

This phase should consist of duplicating at least a significant part of the flight test in the simulator, including flight-test scenarios, surrogate equipment, rules of engagement, and safety and range restrictions. Both baseline and test configuration trials should be run in numbers comparable to these flown in the flight test. A side-by-side comparison of appropriate MOE's calculated in flight test and simulation is then completed. Absolute values of the MOE's, as well as the relative differences between baseline and test article, can be compared for both flight test and simulation. Small differences in MOE's are further proof of the credibility of the simulation data and the validity of the conclusions. Other differences should be rationally explained by known differences between the simulation and the flight test. If the scenarios are nearly identical, as they should be, then the differences should consist mostly of intrinsic differences between the flight simulation and the flight-test, such as pilot work load, stress, and simplifications made in the simulator's modeling of systems or environment. It may be possible to use expert opinion to define and quantify the effects of these differences with a simple multiattribute rating technique and develop transfer functions for each MOE that modify flight-simulation results to more closely match flight test results. Transfer functions that can modify simulation data taken during this phase and produce results close to those of the flight test may be used on data taken during the expanded scope phase of simulation. This should remove simulation unique factors affecting the data. In any event, the conclusions drawn from simulator data should take into account any significant differences between flight test and simulator test conditions.

The final phase of the simulation, the expanded scope phase, is the most important because it provides operational data in areas that cannot be collected in flight test. The scenarios, rules of engagements, and threats used for this phase will be most like those that will be encountered in the time frame that the test article will be operational. This phase is also the longest of the simulation phases. Enough repetitions of the basic scenarios should be made that some statistical significance can be placed on the result. The number of repetitions should be based on the expected variability of the data, as well as its type (discrete or continuous), distribution, and the preferred statistical test. The test matrix should have an equal number of runs in each bin by major factor and by pilot to prevent inadvertent biasing of the data. Once this phase has been accomplished, the final analysis of the data can be started. If desired, the appropriate transfer functions developed from the side-by-side comparison of simulator and flight-test data can be applied to the data from this phase.

The result of a careful simulator test design and integration with the flight test include more efficient procedural training; development of tactics before the test; analysis of suitability of flight-test scenarios and MOE's before their use in the flight test; observation of actual system performances in the flight test to modify simulator models before the simulation operational test; comparison of flight-test and simulator results using identical scenarios, adversaries, and rules of engagement to validate fidelity of simulation; and a simulated operational test using scenarios, advanced threats, and other modeled components not available in the flight test.

The flight test is, to most authorities, the best measure of any new aircraft or subsystem. Nothing in this paper should be taken as a recommendation to replace or reduce the scope of flight testing in favor of manned simulation. Rather, it should be seen as a recommendation for the effective design and integration of simulation programs that improve flight-test programs and provide additional sources of data to test team analysts and better information to the decision-makers.