VERISIMILITUDE TESTING: A NEW APPROACH TO THE FLIGHT TESTING OF AIR FORCE SIMULATORS

MAJOR JAMES A. RICHMOND, USAF
Aeronautical Systems Division
Wright-Patterson Air Force Base

INTRODUCTION

The purpose of this paper is to explain the mechanics and rationale of an improved approach to simulator flight testing that the Air Force has been taking with its most recent simulator procurement. The goal has been to fly and test the simulator so as to identify problems that it may have had in performance and handling qualities. Problem identification has been done in a scientific manner and "tweaking" has been avoided. The test method that has been developed and used, I call verisimilitude testing.

"What is verisimilitude and what is verisimilitude testing?" you ask. Webster's dictionary[1] defines verisimilitude as "the appearance of being true or real." It is just that - how closely the simulator appears to be true or real - that is of interest in the end result. Verisimilitude testing is an approach to testing the performance and handling qualities of a simulator that permits determining just how closely the simulator appears to be true or real.

Recognizing that heretofore simulator testing, within the Air Force at least, has been a highly subjective, iterative process commonly referred to as "tweaking," verisimilitude testing is a somewhat different approach. It is an approach that seeks to determine how closely the simulator duplicates aircraft performance and handling qualities by applying aircraft flight test techniques to the simulator.

One assumption is critical to this approach. That assumption is that the results that are obtained by using standard flight test techniques can be related directly to the pilot's perception of how closely a simulator duplicates aircraft performance. Although there are undoubtedly some shortcomings in this assumption, the assumption should be basically sound.

REQUIREMENTS

The introduction of an ordered approach to simulator testing was considered to be necessary because of the slow, iterative, and frequently nonrepeatable results of "tweaking." The Acceptance Test Procedures written under contract (frequently by the simulator contractor) were clearly not the answer. This is because the Acceptance Test Procedures are written to tell the contractor whether or not he has programmed the math model that he thinks he has. Verisimilitude testing, on the other hand, is aimed at determining whether or not the contractor has programmed a math model that reflects the aircraft being simulated. The big advantages of verisimilitude testing over "tweaking" are that it can be used to compare known aircraft values to those of the simulator, it can be accomplished quickly, and it produces repeatable results.

The program on which this approach to simulator testing has been introduced is the Undergraduate Pilot Training-Instrument Flight Simulator (UPT-IFS) program. In this program two different aircraft are being simulated, the T-37 and T-38, both of which are used to train Air Force pilots. The simulators that are being procured have six-degree-of-freedom motion systems and incorporate a terrain model board visual system. The demands of the Air Force are that these simulators possess a high degree of fidelity (or verisimilitude!) with performance characteristics not "perceptibly different from the characteristics of the real world aircraft."[2] Since the level at which differences become perceptible is not specified and is furthermore not even known, the application of verisimilitude testing seeks to make objective comparisons between the simulator's and the aircraft's performance and handling qualities. Areas where any differences exist between the aircraft and the simulator can be more directly pinpointed and significant problem areas can be made the subject of necessary corrections.

Insuring that there are "no perceptible differences" between the performance of the simulator and that of the aircraft is not an easy task. When a simulator is subjected to "tweaking," the results reflect what a small group of pilots perceive to be differences between the performance characteristics of the simulator and that of the aircraft. There are a number of problems with this approach. Three of these problems bear closer scrutiny:

a. First, differences that are perceived may not exist as perceived. For example, a pilot might feel that control forces

[1] Webster's New World Dictionary. Cleveland and New York: The World Publishing Company, 1958.

[2] Aeronautical Systems Division. ASD Exhibit ENCT 73-1, Wright-Patterson AFB, Ohio, 5 Nov 73.

are too high to produce a given roll rate. What bothers the pilot is his perception of control force, but the problem may not be one of incorrect control forces at all. It could be that stick deflection is too little for a given force or that aerodynamic aileron control power is too little. *In other words,* the pilot can identify what surfaces in his sensory perception as a problem, but will likely be unable to identify the source of such problems.

b. Second, a pilot might identify a problem that is really not a problem at all, or he may exaggerate the extent of a perceived problem. What is happening is that the pilot's memory is not total, and it is playing *tricks on him. This is not as far fetched as* one might imagine. Doing away with "tweaking" by making direct objective comparisons, however, easily overcomes this shortcoming of the "tweaking" approach to testing.

c. Third, "tweaking" does not produce highly repeatable results. The lack of repeatability is due in part to individual differences among pilots and in part to the complex interrelationships of the various facets of simulation.

A DIFFERENT APPROACH

In order to avoid the problems inherent in "tweaking," the verisimilitude approach to testing has been developed. This approach has both a quantitative side and a subjective side. Since the Air Force was virtually guaranteed satisfaction on the UPT-IFS program through the contractual requirement specifying no perceptible differences in performance characteristics, neither a quantitative nor a subjective evaluation could be ignored.

The objective portion of verisimilitude testing begins with a test plan written to allow extracting data from the simulator that can be compared to available aircraft data. This test plan must be written to take advantage of flight test results from the simulated aircraft. For the UPT-IFS T-37 simulator, a short series of flight tests was necessary to fill in some required data that was not available.

Test techniques were not specified in the test plan. The test techniques used, however, were consistent with those used in USAF aircraft flight testing.[3] The test pilots detailed to this project were thoroughly familiar with the use of these test techniques and had flown test profiles in an instrumented T-37 aircraft to collect flight test data.

Quantitative data collection in the simulator was somewhat of a problem. It was found that the computer output line printer was best used for some tests while a strip recorder was required for others. The line printer was extremely useful for testing level accelerations and decelerations, climbs and descents, and longitudinal and lateral directional static stability. The strip recorder had to be used for all dynamic tests, maneuvering flights, stalls, and spins. Since *the strip recorder was run in real-time, it* could be set up to record the desired parameters with optimum scaling and speed. This capability greatly facilitated reading data from the strip chart.

Once collected, the data were reduced and plotted for comparison with aircraft flight test results. Reduction of simulator data was greatly simplified by programming the simulator to duplicate aircraft test conditions. Reduction to standard day conditions was eliminated by setting the simulated environment to standard day conditions. For comparison purposes, flight test results were scaled and plotted on graph paper so as to allow the direct comparison between simulator and aircraft flight test results. The comparative plots were the heart of the quantitative test results since these plots were used to pinpoint and describe simulator inaccuracies.

The greatest difficulty in using this approach was interpreting the results. One must remember that the products of standard aircraft flight test techniques are only reflections of the equations that describe an aircraft's flight. The effects of various components of these equations are frequently interrelated, making it difficult to relate the results of standard flight tests to the equations of motion that are used to program a simulator. This difficulty can be largely overcome by using an experienced test pilot and flight test engineer to collect and interpret the aircraft and simulator flight test data. Using conventional flight test techniques, applied by an experienced test pilot and flight test engineer, therefore, allows us to move from "tweaking" to the more orderly and scientific approach of simulator verisimilitude testing.

Even with the introduction of quantitative tests for comparison purposes, the use *of subjective testing has not been ignored.* Subjective testing was planned early in the UPT-IFS program by the Air Training Command.

---

[3]USAF Test Pilot School. Stability & Control, *AFFTC-TIH-74-2, Edwards AFB, CA,* July 1974. USAF Test Pilot School, Performance, and FTC-TIH-70-1001, Edwards AFB, CA, January 1973.

Valuable assistance in this area has been provided by an Air Force Institute of Technology student who accepted this project for his Master's thesis. In order to produce meaningful results, it was necessary to achieve response repeatability, to eliminate random responses, and to insure that the test subjects represented a fair cross section of Air Training Command instructor pilots.

In preparation for the subjective testing, a questionnaire was prepared by the Air Training Command to insure that each pilot responded to the same questions. Two separate test missions were planned for each pilot participating in the subjective evaluation. Each of these missions was broken down into specific maneuvers, and questions regarding the perception of fidelity were written for each maneuver. The maneuvers planned included all maneuvers that might be trained in the simulator. Questionnaire responses were structured so that each pilot would be required to make one or more dichotomous decisions in order to rank the simulator's performance.

When the questionnaires have been completed and the results are compiled, analysis of the results will be accomplished by use of a computer program. The computer program will aid in the analysis of the results by illuminating areas where correlations can be made. The correlations, where noted, will be used to determine where significant problem areas may exist. Although few, if any, unknown problem areas should exist following the quantitative tesing, the results of the subjective testing should corroborate the quantitative findings.

The subjective tests should be of greatest utility in determining where problems exist with the integration of visual and motion cues in the total simulation. The subjective tests may also provide a valuable input to the Air Training Command in structuring their simulator training program.

Foremost to note in the structuring of this approach to simulator testing is that "tweaking" has been carefully avoided. This is because "tweaking" is an iterative process and is not only time consuming, but may also result in making changes that could affect the results of tests that have already been completed. This does not mean that some obvious problems cannot be corrected on the spot, but problems that are corrected on the spot must not affect data that have already been collected. For that reason, it is fairly critical that the full series of quantitative tests be completed prior to the correction of any deficiencies. Because of the nature of the subjective tests, it is even more critical that the full series of subjective tests be completed prior to the correction of any deficiencies.

LIMITATIONS

Although this balanced approach to simulator testing is a great improvement over "tweaking" the simulator to achieve verisimilitude or fidelity, there are also some disadvantages. First, there are no tolerances to use as guidelines. Although some tolerances should probably be developed, applying tolerances indiscriminately would be almost meaningless. Therefore, the opinion of a trained test pilot must be used in lieu of arbitrary tolerances when evaluating the performance and handling qualities of a simulator.

Second, this approach to testing a simulator cannot accurately probe the pilot's perception of the programmed math model. What is meant is that the perceived stability and control derivatives may be different than those that are programmed. Because of the time delays that exist from pilot input to total system response and because of such factors as motion system washout, visual system scaling, and control loading system inaccuracies, it is likely that there is a considerable difference between the model that the pilot perceives in a simulator and what he would perceive in the aircraft with virtually the same equations of motion. For this reason, an accurate math model may not always produce an accurate simulation. A meaningful test that would yield the components of the perceived math model is not available. Therefore, an impasse could be reached if the subjective feel is that the fidelity of the simulator is not satisfactory but the quantitative tests yield no significant problems.

The above two limitations appear to be the most significant, but by no means the only limitations of verisimilitude testing. For example, tests requiring outside references for optimum performance are constrained by the limitations of the visual system. These limitations on verisimilitude testing, however, should not be overemphasized.

TEST PROGRESS

The first round of quantitative tests have already been completed on the UPT-IFS T-37 simulator and the data have been analyzed. The analyses detected evidence of common problems turning up in more than one test. This recurrence of certain common problems helped to pinpoint specific problems. Once the deficiencies have been noted and corrected, round two of the quantitative testing can take place. Following the correction of the deficiencies noted in round two of the quantitative testing, the subjective tests can begin. Should any deficiencies remain at the commencement of the subjective testing, the results of the subjective tests should either confirm or deny the severity of any such problems. This should provide a form of check and balance concerning the necessity of certain corrective actions.

Where will verisimilitude testing go after it has been conducted on the T-37 and T-38 simulators being procured under the UPT-IFS program? Thus far this approach to simulator flight testing has shown great promise. Given its continued success as an efficient means to test the fidelity of simulator performance and handling qualities, it will likely be used on other simulator procurements now in progress. Verisimilitude testing is a very flexible approach to simulator flight testing and can be easily tailored to meet the demands of the situation.

ABOUT THE AUTHOR

MAJOR JAMES A. RICHMOND is currently assigned to the Air Force Aeronautical Systems Division, Simulator Systems Program Office, Wright-Patterson Air Force Base, Ohio. He was assigned to the Test and Deployment Division of the Simulator Systems Program Office following graduation from Air Force Institute of Technology. Major Richmond's previous military experience includes service as a pilot with the Military Airlift Command, as a forward air controller in Southeast Asia, and as a test pilot with the Air Force Systems Command. He received the B.S. degree from the Air Force Academy and the M.S. degree in systems management from the Air Force Institute of Technology at Wright-Patterson Air Force Base. He received pilot training at Laredo, Texas, and is a graduate of the Air Force Test Pilot School at Edwards Air Force Base, California.