IDA

INSTITUTE FOR DEFENSE ANALYSES

Unmanned Aerial Vehicle Operational Test and Evaluation Lessons Learned

Lee Carr, Project Leader
Kristen Lambrecht
Scott Shaw
Greg Whittier
Catherine Warner

December 2003

Approved for public release;
distribution unlimited.

IDA Paper P-3821

Log: H 03-001901

This work was conducted under contract DASW01 04 C 0003, Task BD-9-2299, for the Director, Operational Test and Evaluation. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-3821

Unmanned Aerial Vehicle Operational Test and Evaluation Lessons Learned

Lee Carr, Project Leader
Kristen Lambrecht
Scott Shaw
Greg Whittier
Catherine Warner

PREFACE

This paper was prepared by the Institute for Defense Analyses (IDA) for the Director, Operational Test and Evaluation (DOT&E), in partial fulfillment of the Operational Test and Evaluation Program tasks for Tactical Aerial Reconnaissance Systems. This paper presents lessons learned from the operational testing of Unmanned Aerial Vehicles from 1986 to 2002.

The IDA Technical Review Committee was chaired by Mr. Robert R. Soule and consisted of Dr. Michael Shaw, Mr. John Kreis, Mr. Al Wallace, Dr. Don Richardson, and Dr. Andrew Atwell.

UNMANNED AERIAL VEHICLE OPERATIONAL TEST AND EVALUATION LESSONS LEARNED

List of Figures

List of Tables

SUMMARY

A. GENERAL

This paper had as its genesis the Army Tactical Unmanned Aerial Vehicle (TUAV) Initial Operational Test (IOT) conducted in 2002. At the conclusion of the report-writing phase of that test, significant lessons had been learned that should be documented and applied to future Unmanned Aerial Vehicle (UAV) testing. In the course of preparing this paper, a review of previous UAV testing was conducted. This review identified shortcomings in numerous aspects of UAV operational testing (OT) that seemed to occur repeatedly from the OT of one system to another.

This paper is meant for DOT&E action officers (AO) with purview over Intelligence, Surveillance, and Reconnaissance (ISR) UAV systems. The contents highlight areas of UAV testing that have proven problematic from DOT&E's perspective in past UAV OTs. Armed with this knowledge, DOT&E AOs should be better able to positively influence the scope and conduct of future UAV OTs.

This paper should also prove of value to UAV program offices and Operational Test Agencies (OTAs). The lessons learned presented herein should be applied well in advance of drafting Parts IV and V of the Test and Evaluation Master Plan (TEMP), and certainly prior to the start of test planning for the first OT period.

B. REQUIREMENTS DEVELOPMENT

While DOT&E is not a stakeholder in the requirements generation process, there are occasions where DOT&E might be able to influence the evolution of Operational Requirements Documents (ORD) as a system matures. A thorough review of UAV ORDs as early as possible is critical to ensuring that testable UAV mission statements and requirements form a solid foundation for adequate testing.

Mission Statement. A disturbing trend in recent UAV requirements documents has been to focus the ORD on technical characteristics rather than describing the operational environment and missions that the UAV is expected to perform. Besides driving the design of individual test events, the UAV mission statement determines the overall OT program leading to a full-rate production decision. A clear mission description also supports a spiral development process if required to meet all of the users' needs. Finally, a clear mission statement assists the development of the test resources and test funding sections of the Test and Evaluation Master Plan (TEMP).

Effectiveness. Setting clear, measurable effectiveness criteria is critical to the success of posttest data analyses. At the highest level is Critical Operational Issues (COI). DOT&E should seek to influence COIs to ensure that the overall contribution of the UAV system toward mission accomplishment can be evaluated. At a lower level, it has been difficult to define metrics for the value of intelligence. Without clearly defined criteria, it is difficult to reach consensus on system effectiveness.

Suitability. To date, every UAV system that has undergone OT has been deemed not operationally suitable. When evaluating "systems of systems" such as UAVs, it is difficult to determine reliability requirements. During the requirements generation phase, suitability criteria should be defined and understood. Prior to the start of a test, system operations and the effects of component failures should be understood as fully as possible. Documents supporting the rationale for each of the reliability, availability, and maintainability metrics should be published. Reliability Block Diagrams are important in understanding and defining system failures.

C. TEST DESIGN

Many of the problems encountered during UAV OT could have been prevented during the test design phase. Test designers should ensure that the appropriate data can be collected in sample sizes large enough to support definitive conclusions regarding UAV effectiveness.

Intelligence, Surveillance, and Reconnaissance Cycle. The Intelligence, Surveillance, and Reconnaissance (ISR) cycle describes the manner in which a UAV unit is tasked and depicts the flow of information from the UAV system to the end user(s). Of necessity, the ISR cycle for each UAV differs depending on the capabilities of that particular UAV system and the type of data provided. The ISR cycle applicable to the UAV unit determines the scope and depth of test support assets required to conduct an adequate OT. If there are units that provide support to, task, exploit, or receive imagery from the UAV unit under the envisioned ISR cycle, then these units should be included in the operational test.

Test Environment. The operational environment under which testing is conducted should be representative of that encountered during combat operations. The test scenario, test sites, target sets, and launch and recovery sites should provide an adequate test of the UAVs capabilities, and identify shortcomings. Most importantly, the test scenario should allow for realistic use of the UAV system by the test unit. Two other key areas that should be addressed in detail during test design are the use of threat air defense systems, and Real Time Casualty Assessment (RCTA). A representative threat air defense system would enable commanders participating in the operational tests to be confronted with realistic operational trade-offs between losing air vehicles and imaging high value targets. An effective RTCA system could produce insights regarding air vehicle losses during routine combat operations.

Nonoperational Test Data. In order to reduce the time and cost of OT, many programs propose the use of data gathered during non-OT events. However, a careful review of the data should be conducted prior to approval of any test strategy that seeks to employ non-OT data.

Survivability Testing. The scope of survivability testing for any UAV should be based upon the acceptable level of attrition for that UAV as well as the value of the data provided by the UAV. Expendable UAVs may require no survivability testing. It is expected that these UAVs will be launched with little expectation of their returning safely. Attritable UAVs, while expected to suffer losses, should complete some level of survivability testing if only to provide information regarding expected losses. As a general rule of thumb, the more expensive the UAV (and its sensors), the greater the requirement to conduct detailed survivability testing.

D. TEST EXECUTION

A carefully planned OT should be adequately executed during the actual execution phase. There are certain aspects of the test that should be monitored, to ensure that the UAV is employed and evaluated in a manner representative of the users' desires.

Data Collection. Ensuring that the correct data to support posttest analyses are collected is critical during the test execution phase. Emphasis should be placed on tracking the taskings assigned to other ISR assets as well as all taskings assigned to the UAV unit. This information is required in order to determine the reliance of the unit on UAV operations as well as the overall tasking success rate of the UAV.

Cueing. Target cues provided to the UAV operators should be monitored to ensure that they accurately reflect the capabilities of the system that generated the cues. Grid locations or target descriptions that are too detailed should be screened and future cues adjusted accordingly.

Test Control. To the greatest extent possible, the test agency should not interfere with the test unit during OT; however, there are exceptions to this philosophy, most notably the need to balance test interference with data collection. The use of Observer/Controllers (OCs) during OT is one area where test team interference on test unit operations is acceptable. OCs could be used to influence test unit behavior to ensure that test objectives are met and to ensure that the test unit employs UAVs as envisioned by the Concept of Operations.

E. APPENDICES

Two appendices at the end of the paper describe operational testing on UAVs conducted since 1986. Appendix B covers formal operational test periods while Appendix C covers Advanced Concept Technology Demonstrations. The recommendations and lessons learned presented in this paper were derived from the testing described in these two appendices.

CHAPTER I

INTRODUCTION

I. INTRODUCTION

A. GENERAL

This paper had as its genesis the Army Tactical Unmanned Aerial Vehicle (TUAV) Initial Operational Test (IOT) conducted in 2002. At the conclusion of the report-writing phase of that test, significant lessons had been learned that should be documented and applied to future Unmanned Aerial Vehicle (UAV) testing. In the course of writing this paper, a review of previous UAV testing was conducted. This review identified shortcomings in numerous aspects of UAV OT that seemed to occur repeatedly from the OT of one system to another.

B. PURPOSE

The purpose of this paper is to provide the Director, Operational Test and Evaluation (DOT&E) with lessons learned from past UAV testing that could be applied to future UAV OT. These lessons learned take the form of requirements development, test design, and test execution issues that should be recognized and resolved in order to ensure the adequacy of future UAV OT.

This paper does not contain an in-depth discussion of the programmatics and developmental test history of UAV systems. It is felt that the myriad of developmental problems (often repeated) encountered by UAV systems could be the subject of a separate paper in and of itself. This paper is not intended to be a "cookie-cutter" for planning an OT or evaluating UAV performance. Rather, it discusses problems and shortcomings encountered during previous tests in an attempt to prevent them from occurring in future tests.

The focus of this paper is on UAVs acquired to perform reconnaissance, surveillance, and target acquisition missions (RSTA).[1] Since all the UAVs that the Services have tried to field since Aquila in the 1980s were to be employed as RSTA

---

[1] The paper may also have limited applicability towards the Navy and Air Force Unmanned Combat Air Vehicles. However, the degree of applicability will depend on the level of focus these UAVs place on the RSTA mission versus the target attack mission.

assets, all of the lessons learned presented here are applicable to future RSTA UAVs. As such, this paper may be applicable to the Navy's vertical takeoff and landing TUAV, the Air Force Global Hawk, and the Army Extended Range Multipurpose UAV.

Finally, this paper will concentrate on the major OT conducted just prior to DOT&E issuing a Beyond Low Rate Initial Production report (B-LRIP).[2] This is the point at which a final determination is made regarding resolution of the system's operational effectiveness and operational suitability. Prior to this point, it is expected (and understandable) that system performance may be somewhat less than desired due to system maturity. Therefore, many of the lessons learned may not be applicable prior to that B-LRIP testing.

C. TARGET AUDIENCE

This paper is primarily meant for DOT&E action officers with purview over UAV systems. The contents highlight areas of UAV testing that have proven to be problematic from a DOT&E perspective during past UAV OT. Armed with this knowledge, DOT&E action officers may be able to positively influence the scope and conduct of future UAV OTs.

This paper could also prove to be of value to UAV program offices and Operational Test Agencies (OTAs). The lessons learned presented herein could be applied well in advance of drafting Parts IV and V of the Test and Evaluation Master Plan (TEMP), and certainly prior to the start of test planning for the first OT period.

We hope that in presenting specific examples and trends from past UAV testing we will make program managers and testers aware of specific test elements of interest to DOT&E. It should be stressed that all of the critical elements of UAV testing identified in this document should be satisfactorily addressed prior to DOT&E approval of future UAV system TEMPs and OTA test plans.

D. ORGANIZATION

Chapter II presents issues that should be resolved during the requirements development phase of a program's life cycle. These issues affect contents of the Operational Requirements Document (ORD) and the TEMP that are of importance to the

---

2 For Army programs, this test is referred to as IOT. For the Navy, this period is called "Operational Evaluation" (OPEVAL).

operational tester. Issues such as quantifiable and measurable metrics, test resources, and the test schedule have an influence on test adequacy and execution.

Chapter III discusses lessons learned related to test design that should be resolved during the design of the OT. Issues related to the test environment, instrumentation, and data collection could affect the scope of the test and thus need close scrutiny during the test design phase. These issues should be reviewed and resolved prior to submitting of the OTA test plan for DOT&E approval.

Chapter IV presents lessons learned during the test execution phase of OT. The day-to-day activities of the OTA, the test unit, and test support units should be closely monitored to ensure that an accurate reconstruction of the test is available to support posttest analyses. Addressing these issues while the test is being conducted ensures that adequate data are available from which a strong statement regarding system effectiveness, suitability, and survivability can be made.

Chapter V presents describes methodologies used during past operational testing to assess the mission effectiveness of UAVs.

Two appendices at the end of the paper describe operational testing on UAVs conducted since 1986. Appendix B covers formal operational test periods while Appendix C covers Advanced Concept Technology Demonstrations (ACTD). The recommendations and lessons learned presented in this paper were derived from the testing described in these two appendices. Table I-1 lists the UAV system and the testing reviewed during the development of this document.

Table I-1. UAV Testing Reviewed for This Paper

System	Test Dates	Testing Conducted
Aquila	1986, 1987	Operational Test, Live Fire Testing
Pioneer	1986	Fielded without an Operational Test
Hunter	1992	Limited Users Test
Predator	1995-96, 2000	Military Utility Assessment, Initial Operational Test
Outrider	1998	Military Utility Assessment
Global Hawk	1999-2000	Military Utility Assessment
Shadow	2001, 2002	Limited Users Test, Initial Operational Test
Fire Scout	2001	Early Operational Assessment

CHAPTER II

REQUIREMENTS DEVELOPMENT

II. REQUIREMENTS DEVELOPMENT

Successful operational testing of UAV systems is supported by a clearly stated mission definition, system performance criteria and criteria threshold values. Past testing has shown that system performance criteria are often times ill defined, contradictory, and not operationally relevant. From past experience, UAV systems with ill-defined missions and requirements have usually encountered problems with test design, test execution, and posttest analyses.

While DOT&E is not a stakeholder in the requirements generation process, there are occasions where DOT&E may be able to influence the evolution of requirements as a system matures. Early in the life of a program, there is usually intense discussion regarding the meaning and intent of specific aspects of the ORD. At the same time, the program manager may be seeking to gain approval to "trade-off" capabilities in order to gain acceptance by the user community as the system matures and its actual capabilities (and limitations) become known. During these discussions, DOT&E's opinions regarding system requirements usually carry considerable weight. During this period, system requirements should be reviewed from a testability viewpoint since strong DOT&E input may be needed to shape the process.

A thorough review of UAV ORDs is needed to ensure that testable UAV mission statements and requirements form a solid foundation for adequate testing. Additionally, UAV requirements should be reviewed to assess how well they support definitive statements regarding system operational effectiveness and suitability.

A. MISSION DESCRIPTION

A disturbing trend in recent UAV ORDs has been to focus the ORD on technical characteristics rather then describing the operational environment and missions that the UAV is expected to perform. In order to fully inform decision makers, operational testing should be conducted in an operationally realistic environment. In order to do so, a clear mission description is called for prior to the start of test design. Besides driving the design of individual test events, the mission statement shapes the overall OT program leading to a full rate production decision. A clear mission description also supports a spiral development process, if required, to meet all of the users' needs. Finally, a clear mission statement assists the development of the test resources and test funding sections of the TEMP.

Without a clear mission description, disconnects between the requirements generator and the test community could occur. At a minimum, this will create friction during the test planning process as time and effort will be spent determining the users' actual needs versus those reflected in the ORD. This may result in an OT design that does not reflect the missions and operating environment envisioned by the users.

B. EFFECTIVENESS CRITERIA

There are two levels to evaluating the operational effectiveness of UAV systems. At the highest level are the Critical Operational Issues (COIs). At the lowest level, there are specific performance characteristics, such as time on station at a given range, the probability of detecting a target, or the level of target location error. Specific performance characteristics are usually utilized within requirements documents due to the perceived ease of measuring and evaluating these types of metrics; however, this approach presents its own set of unique problems as described below.

The following sections review frequently used effectiveness metrics and discusses issues that have proved troublesome during past OT events.

1. Critical Operational Issues

COIs are used to measure the overall contribution of the system towards mission accomplishment. Defining an appropriate COI, and then choosing operationally relevant criteria is a method that has rarely succeeded in the case of UAV testing. Normally Service recommended COIs are limited to minimum range, on-station time or payload requirements. While easily tested, these types of COIs may not provide the decision maker with a true sense of the contribution of the UAV towards overall unit effectiveness.

It is important to note that, in comparison to ORD requirements in which DOT&E has little influence, COIs can be influenced by DOT&E. COIs are developed by the Services and defined in the TEMP. As the overall approval authority for TEMPs, DOT&E has strong influence over the COIs used during OT. The TUAV effectiveness

COI[1] was heavily influenced by DOT&E during the TEMP approval process, and as a result, there was a quantifiable, high-level metric available to evaluate overall TUAV performance. Without DOT&E input, such a metric may not have been available during posttest analyses.

2. Value of Information

Measuring the contribution of information gathered by any UAV system has long been problematic. Attempting to quantify the "value" of information gathered by the UAV as it contributes to overall mission success has rarely been attempted. While other intelligence systems such as the Joint Surveillance and Target Acquisition Radar System (JSTARS) have tried to assess operational effectiveness this way, only recently has this been attempted for UAV systems.

Overall, assessing the operational effectiveness of a UAV based solely on performance characteristics instead of its overall contribution to mission success is questionable. For instance, suppose that a given UAV is required to have a target detection rate of 60 percent, and that during OT, the UAV achieves a target detection rate of 39 percent. Is the system not operationally effective because it failed to meet the 60 percent target detection criteria? The degree of effectiveness should be dependant upon what targets were detected and when, in other words, the value of the information provided.

During the TUAV IOT conducted in 2002, an attempt was made to score, in real time, the value of information provided by the TUAV. At the time each UAV report was received, an intelligence officer was tasked to evaluate its contribution. The problem with this methodology was that the intelligence officer had no inkling of the ground truth situation, and as a result, the intelligence officer was unable to determine the degree of accuracy associated with each report. Posttest analysis revealed that many reports containing inaccurate, or even worse, misleading, information were given passing contribution scores. This, combined with other issues, may have lead to overestimating the operational effectiveness of the TUAV based on the contribution scores.[2]

---

1 Does the TUAV system contribute to the commander's requirements for timely and accurate reconnaissance, surveillance and target acquisition information?

2 Chapter V, Evaluating Mission Effectiveness, presents a detailed discussion of the report success templates used during the TUAV IOT.

An alternate method would have been to adopt the one used during the JSTARS Common Ground Station (CGS) OT. During the JSTARS CGS test, the contribution of each report was assessed at the conclusion of the exercise. In this way, the ground truth associated with each report was understood and inaccurate reports could be assigned proper contribution scores.

The value of information provided by the UAV should be assessed during realistic test scenarios and not during developmental test (DT) sorties. The scenario should be such that users are forced to employ the UAV to develop or confirm intelligence vital to the successful completion of their mission. Additionally, the scenario should contain elements that force users to react to data collected or confirmed by the UAV.

Regardless of the methodology chosen, the value of information collected by the UAV should be measured so that the commander can discern the value of the system and an informed decision regarding the cost benefits of the system can be made.

3. Timeliness

Metrics and methodologies used to measure the timeliness of UAV imagery is an area that should be addressed during the requirements process. The major area of discussion here revolves around the scope of the timeliness issue. In the eyes of the TUAV program manager, timeliness could be been measured from the time a task is received at the Ground Control Station (GCS) until the operator sends out a report related to that task. In this case, systems affecting timeliness are limited to those under the purview of the UAV program manager.

In the larger context, timeliness could be measured from the time a task is generated by the requesting unit to the time the requestor receives the report. In this case, systems beyond the purview of the UAV program manager may adversely affect UAV timeliness. At the tactical level, a GCS located in the Brigade Tactical Operations Center (TOC) is likely to prove responsive to Brigade imagery requests. The same GCS tasked to support Battalion level intelligence requests is likely to be less timely due to delays in transmitting requests from the Battalion, through the Brigade, and then into the GCS. Delays in the opposite direction, when GCS reports flow to the Battalion, are also to be expected. In this case, timeliness criteria are expected to be somewhat longer.

When setting timeliness requirements, the full scope of the requesting and reporting lines of communications should be examined and considered. Once this is done, the OT should be designed so that all lines of communications are replicated during the test.

4. Imagery Quality

An operationally realistic evaluation of imagery quality would ask the requester to rate whether the imagery provided was of sufficient quality to satisfy their essential elements of information. During previous OT, user ratings of imagery were not systematically collected, (i.e., ratings were not collected on a tasking-by-tasking basis).

A goal of the Predator Advanced Concept Technology Demonstration (ACTD) was to demonstrate National Imagery Interpretability Rating Scale (NIIRS) Level 6[3] imagery from 15,000 feet slant range; this quality of imagery provides a ground-resolved distance of 16 to 30 inches and allows for differentiation between distinct vehicles, aircraft, and ships. Imagery of NIIRS Level 6 quality was desirable to permit the Predator to identify and track tactical-sized, mobile targets such as tanks, personnel carriers, and artillery pieces.

NIIRS is not a typical operational measure; it is often used to provide objective imagery ratings against engineering targets from prescribed altitudes and slant ranges. However, NIIRS is more subjective against tactical targets in an operational environment.

A thorough evaluation of UAV effectiveness should include near real time user ratings of the imagery provided by the system. Care should be taken to ensure that the ratings are only applied to images acquired at operationally relevant slant ranges. For example, during the TUAV IOT only eight of 110 reports imaged targets at the required standoff distance or greater. If all of these images were used to assess imagery quality, the results may overstate the true capabilities of the system at the desired ranges.

In the absence of user ratings, selected samples of UAV imagery should be reviewed in order to determine the operational capabilities of the system under test. The samples evaluated should span all of the target sets involved in the test as well as those taken from various target slant ranges.

---

[3] The NIIRS ranges from Level 0 to Level 9 where Level 0 quality precludes any interpretation and Level 9 quality provides the highest level of ground resolution.

5. Detection, Classification, Recognition Criteria

Different UAVs are expected to demonstrate varying capabilities regarding the detection, classification, and recognition of targets. For the sake of brevity, the following section will focus on target detection criteria; however, the same examples and comments are applicable when testing the classification and recognition capabilities of UAV systems.

Prior to testing, all detection criteria should be understood. During Aquila testing, criteria used to measure system success for detection probability were incomplete and confusing. In attempting to quantify detection probabilities, the two criterion statements (50 percent detections of moving target arrays and 30 percent detections of stationary target arrays) presented a major problem: the area of search and the time of search were not specified. Intuitively, one would expect a higher detection rate given a smaller search area and/or a longer search time for a given area. Further, stationary targets were either in the open, under hasty camouflage, or fully camouflaged, but the percentage of targets to be detected in each situation was not specified.

In light of posttest difficulties in measuring UAV detection performance, future detection criteria should include the following:

- A threshold
- A description of the expected target sets (camouflaged/non-camouflaged, hot/cold/stationary/moving)
- Conditions in which a "success" is achieved
- Description of the conditions in which the criteria will be compared (a time limit and search area size).

It is possible that a matrix of criteria will be required in order to fully delineate the desired performance of the UAV system. These criteria should state the percentage of targets to be detected, recognized, or classified for various target characteristics.

6. Target Location Criteria

Criteria used to evaluate target locations provided by the UAV should be clear, quantitative, and operationally relevant. Target position criteria should be based on an overarching metric (with justification) and not be limited to a simple measure of target location error (TLE). As a good example, the TUAV ORD states that the system would have a TLE of less than 80 meters in order to facilitate second-round fire for effect artillery missions. The Aquila requirement for 85 percent of the Mean Point of Impact

(MPI) of fire for effect falling within 50 meters of the target is another example of quantifying the target location capabilities of the UAV.

In both the TUAV and Aquila requirements, problems with the fire support criteria arose. In the case of Aquila, the stated criterion of 85 percent was very stringent compared to existing and planned systems. Additionally, there was no limit on the number of rounds that could be fired in adjustment prior to the fire-for-effect salvo. Finally, the inherent Circular Error Probable (CEP) of the artillery rounds themselves could cause the Aquila UAV to fail to meet the ORD requirement.

In the case of the TUAV criterion, there did not appear to be an operational basis for the 80-meter TLE requirement. The percentage of fire missions with an 80-meter or less TLE is omitted.[4] It is unrealistic to think that TUAV would meet the 80-meter requirement TLE on 100 percent of the fire missions. It is more realistic to expect that the 80-meter TLE criterion might be met for some percentage of the fire missions conducted.

For artillery adjustment missions, an acceptable criterion should state the desirable effects on target for a given percentage of the fire missions. There should also be some timeline (focused on the UAV portion of the sensor to shooter chain and the number of rounds) associated with each fire mission.

Precision-guided munitions should be addressed using separate criteria. This criterion should focus on the accuracy of the target location and the ability of the UAV to designate for the precision munitions or to quickly report an accurate location to the actual shooter (depending on the capabilities of the UAV). The degree of accuracy required may be directly related to the capabilities of the weapons system being employed. An additional requirement could be that target coordinates be reported in a format that is interoperable with the fire support system to be employed. This interoperability chain should extend from the UAV sensor through the actual firing unit without a need for data conversion along the way.

A target location criterion for intelligence purposes is somewhat harder to measure. The degree of error in the reported location should be tied to the value of the target, the way in which the information will be used, and the risk involved with an

---

4 It was assumed that the 80-meter TLE requirement was, in terms of fire support, an 80-meter CEP requirement. This would imply that 50 percent of reported fire support target locations would be less than 80 meters from the actual location. However, the ORD did not specify such a requirement.

erroneous location. The overall concern when setting criteria is to tie target location criteria to the needs of the commander. To date, no UAV planning documents have addressed this issue.

7. Survivability Criteria

Recent UAV requirements documents such as the TUAV have attempted to characterize survivability criteria in terms of infrared (IR), radar cross-section, and/or aural/visual signatures. In terms of OT, these criteria may not allow for an assessment of the operational impact of UAV survivability, whose criteria should be stated in terms of probability of mission success, expected losses, or expected operating hours, given an operationally realistic threat air defense system.

8. Effective Time On-Station

One key capability of any UAV system is its Effective Time On Station (ETOS). ETOS is a measure of the amount of time on-station (with functioning sensors) the system is able to provide to the war fighter. ETOS is a function of the number of air vehicles in a UAV system, the distance from the UAV base to the operating area, UAV airspeed capabilities, and the system reliability and maintainability characteristics.

Since the UAV could be called upon to operate almost anywhere in the world, the distance from the UAV operating base to its operating area could be any distance imaginable (limited only by its data link). However, for analytical purposes, most UAV operational documents, and the characteristics of the AV itself, provide a desired range and on-station requirement. This desired range normally forms the basis for the ETOS calculation.

UAV requirements documents should contain reliability and maintainability thresholds that, when combined with the desired distance and airspeed characteristics, support the desired ETOS. A difficulty arises when the presupposed values used to estimate the ETOS for the system are not demonstrated during OT. A determination must then be made as to whether or not the required ETOS can be met when some of the required parameters have not been. As an example, the demonstrated values for TUAV reliability metrics during IOT were below those required in the ORD. It was determined that the ETOS requirements could be met given the inherent reliability of the system provided by multiple air vehicles.

C. SUITABILITY CRITERIA

To date, every UAV system that has undergone OT has been deemed not operationally suitable. During the requirements generation phase, suitability criteria should be fully defined and understood. Also important is that the operational impact of each suitability requirement is identified, and the implications from not meeting a given threshold are understood.

1. Reliability Criteria

Because UAVs are essentially "systems of systems" (air vehicles, GCS, ground support equipment, etc.), ORD reliability criteria can be especially ambiguous. All components of desired reliability metrics should be clearly defined and understood by all parties involved with future UAV testing, and the criteria should be operationally relevant. While the intent of reliability metrics is usually understood, it should be translated into a clear, easily understood, and quantifiable measure.

An example of this is the definition of "operating time" from the TUAV IOT. The TUAV ORD required a 20-hour mean time between system abort that explicitly defined the "time" measure as flight hours plus mission planning time. But according to the TUAV failure definitions (a separate document), the system is in continuous operating mode for the duration of an operation. As a result, the Army defined failure modes and operating times that were in direct conflict with the ORD requirement, which caused difficulties and friction between the OTA and DOT&E during posttest analyses.

2. Supporting Documents

During the requirements generation phase, several source documents related to reliability criteria should be examined. Without these documents, it may not be possible to attach threshold and objective criterion to each requirement. Similarly, the operational impact of failing to meet a required threshold may not be accurately assessed during post test analyses without these supporting documents.

a. Reliability, Availability, and Maintainability Rationale

A source document that drives the development of reliability, availability, and maintainability (RAM) metrics should provide the analytical underpinnings for the threshold values associated with each RAM metric. It is possible that a system could fail to meet several RAM metric thresholds yet still be deemed operationally suitable. For example, the TUAV failed to meet its mean time between system aborts criteria.

However, the inherent reliability built into the system (multiple AVs, multiple launchers) mitigated many of the suitability shortcomings of the system. With knowledge of the source and value of each threshold, the analyst is better able to judge the operational impact should one or more thresholds not be met during OT. In past testing, the reasoning behind various RMA metrics has been unknown. When a system failed to meet the established threshold, it was not possible to judge the impact of this shortcoming; therefore, the overall suitability of the system could not be ascertained. Producing a RAM rationale report may be in the best interest of the program manager since it could reduce the risk of finding a system unsuitable should it fail to meet a given metric.

b. RAM Failure Definitions and Scoring Criteria

Prior to the start of testing, a document that outlines specific RAM definitions and scoring criteria should be developed. Such a document is used to determine the classification and chargeability of reliability and maintainability incidents during the course of the test. These definitions are then used to establish a database to compute point estimates for RAM values to assess operational suitability and ownership costs.

This document should define all RAM terms to be used during the test.[5] Each failure type used during the test (i.e., mission essential failure, system abort, operational mission failure) should be clearly defined and agreed upon. Ideally, a flow chart will be included to provide a step-by-step methodology to determine the failure type, chargeability, and effect of each failure. Other metrics required to make a finding regarding system suitability (such as built-in-test) should also be included in the definitions and flow charts. Each failure can then be scored using RAM definitions set forth in this document.

A major feature of such a RAM scoring document is a mission essential sub-systems list (MESL). This MESL shows what subsystems are needed to successfully perform each mission/function of the UAV. The MESL, derived from reliability block diagrams, can be used to determine the severity of subsystem failures and the missions affected by that failure.

It is highly improbable that a given RAM scoring document would cover every situation that could arise during testing. Many failure modes cannot be identified until

---

[5] The TUAV IOT definition for *operating time* presented earlier is an example of the difficulties with ambiguous definitions.

systems are in the hands of operational users and employed in a manner not envisioned by engineers sitting in air-conditioned buildings. As an example, on several occasions during one Global Hawk ACTD deployment generators powering ground stations ran out of fuel. While a Global Hawk FDSC would probably have covered power failures, it may not have covered power failures due to operator errors such as this.

In these cases, it may be necessary for representatives of the test agency, the end user, and the program office to make a final determination of RAM scoring. When omissions to established RAM scoring procedures are discovered, a change to the procedures should be documented to ensure standardized scoring for similar failures.

c. Reliability Block Diagrams

The use of reliability block diagrams is critical during the development of RAM failure definitions and scoring. Any UAV system is a system-of-systems consisting of air vehicles, ground stations, launch and recovery systems, and imagery systems. Normally, a UAV system offers a great deal of redundancy to ensure that imagery can be provided on demand with little interruption. In order to evaluate the operational suitability of a UAV system, it is necessary to understand the function of each subsystem, when each one is required during a mission, and the effect of a failed subsystem in any phase of a mission. This is required knowledge to judge the affect of failures to determine the overall reliability of the system.

In order to make sense of this redundancy, reliability block diagrams for each phase of a UAV mission could be developed. As an example, Figure II-1 presents the reliability block diagram for the TUAV launch stage of a mission. Note the system redundancy as evidenced by multiple air vehicles and two separate paths from which command and control of the Air Vehicle (AV) may be affected.

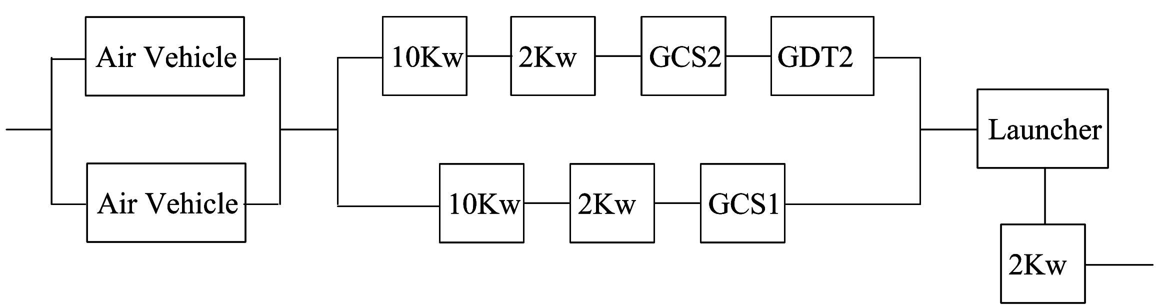

Figure II-1. TUAV Launch Stage Reliability Diagram

Prior to testing, block diagrams developed for UAVs should be reviewed to ensure operational relevance. Figure II-1 is misleading in that only one AV is mounted on the launcher. Should that AV fail its preflight checks, it must be removed from the launcher and another AV mounted in its place. The delay in replacing AVs could cause the system to miss a launch time and result in a mission abort. As another example, Figure II-2 presents the TUAV reliability diagram for the recovery stage of a TUAV mission. Note that this diagram shows the presence of two TUAV Automatic Landing Systems (TALS). However, normal employment of the TUAV system has only one TALS deployed and the other stowed away in the transport vehicles. It could be argued that, depending on the fuel state of the TUAV and the time required to set up the second TALS, there could be a single point-failure mode at the TALS. In theory, the TUAV system was redundant within the TALS segment of the TUAV recovery reliability diagram. Operationally, this was not the case.

In contrast, the two command and control paths in Figure II-2 are always active and either could assume control of the TUAV recovery at a moment's notice.

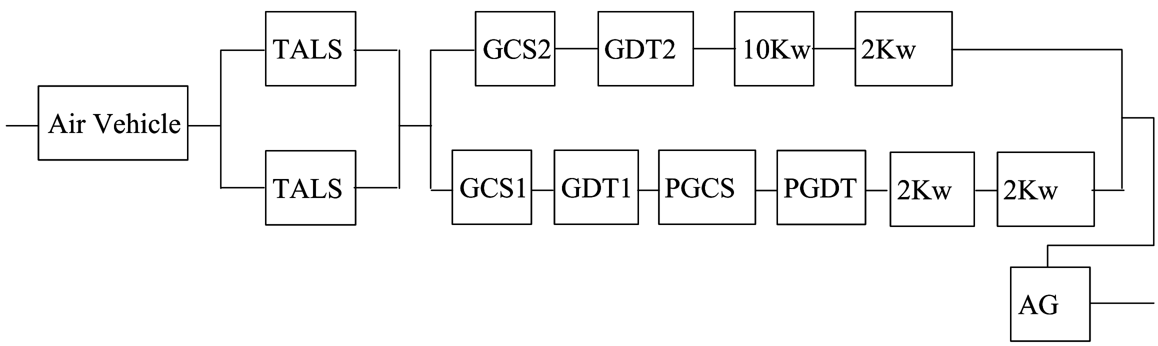

Figure II-2. TUAV Recovery Reliability Diagram

d. Operational Mode Summary/Mission Profile

The Operational Mode Summary/Mission Profile (OMS/MP) provides a description of the missions most likely to be performed by the UAV unit.[6] In UAV terms, this document outlines the expected operational tempo (OPTEMPO) of the UAV unit over a given time span. The number of launches, the number of operating hours, and the frequency at which the unit displaces are all delineated within this document. Knowing the intended use of the UAV system is needed in order to evaluate the design of

6 OMS/MP is the term used in Army documents. Other Services provide similar information in a Concept of Operations or a Concept of Employment.

the OT. The design of the test should ensure that the UAV system is flown and operated to the level outlined in the OMS/MP.

According to the TUAV OMS/MP, the TUAV platoon was required to be able to provide flying times of 12, 18, 18, 18, and 8 hours on 5 consecutive days. As a result of this requirement, the first week of the TUAV IOT was dedicated to demonstrating the capability of the system to perform at the level of operations set forth in the OMS/MP. It also described the frequency and reasons that the UAV platoon is expected to displace, move, and emplace (DME). According to the IOT scenario, the TUAV platoon was not forced to displace at the frequency or for the reasons stated in the TUAV OMS/MP. As a result, a definitive evaluation of the DME capabilities of the platoon could not be made due to the poor test design.

D. INTEROPERABILITY

The Joint Interoperability Test Command (JITC) is responsible for certifying that all DoD C4I systems are interoperable. DoD Directive (DoDD) 4630.5 and DoD Instruction (DoDI) 4630.8 mandate joint and combined interoperability certification testing for "all Command, Control, Communications, and Intelligence (C3I) systems developed for use by U.S. forces." This certification is required prior to fielding a new system.

The key to a successful interoperability certification is the early and frequent involvement of JITC. The JITC should be consulted during the requirements development phase. By doing this, JITC staff can review concept of operations diagrams and architectures, and identify key interoperability requirements.

The JITC has used the TUAV program as a model in which early involvement led to a good working relationship between the program office and the JITC test team. JITC personnel worked hand-in-hand with TUAV program office personnel early on in the life of the program. This close working relationship helped JITC understand the TUAV operating environment and allowed Program Manager (PM) personnel to fully understand the role and requirements of JITC certification. However, this working relationship was unable to obtain a TUAV interoperability certification prior to the end of IOT due to a non-standardized use of software at the unit level during the TUAV IOT.

E. UAV ADVANCED CONCEPT TECHNOLOGY DEMONSTRATIONS

A major shortfall of the ACTD system is that much of the early requirements generation work associated with formal acquisition programs has not been completed. Upon transitioning to a formal acquisition program, many of ACTD systems lack formal requirements and supporting documentation. Additionally, the time allocated for planning for and executing OTs becomes compressed, often resulting in less than optimal testing and results.

It has been suggested that once an ACTD program has demonstrated "military utility" and is deemed successful, the asset should not be deployed until the post-ACTD preplanned product improvement program has been completed and the system is deemed supportable (Ref. 2). A case in point is the loss of a Global Hawk AV during an operational deployment that resulted in the destruction of the last EO/IR payload. While other air vehicles were available, further development of the system was delayed until additional sensors could be produced. Although it is understandable that a system with proven military utility is in demand, immediately deploying assets upon ACTD acceptance makes it extremely difficult to correct or implement identified equipment modifications and upgrades, and as a result, the system might be unable to reach its full potential for some period of time.

CHAPTER III

TEST DESIGN

III. TEST DESIGN

Many of the problems encountered during UAV OT could have been prevented during the test design phase. Data to support definitive conclusions regarding system effectiveness and suitability could have been collected had certain aspects of the test design been better implemented and executed. The design of the test itself should ensure that the appropriate data is collected in sample sizes large enough to support definitive conclusions regarding UAV effectiveness.

This chapter examines test issues that should be resolved prior to the start of detailed test planning. Decisions made regarding these issues could affect the adequacy of planned testing. Additionally, the true capabilities and limitations of a UAV system operating in a combat environment may not be adequately evaluated based upon a less than optimal test design.

A. INTELLIGENCE, SURVEILLANCE, AND RECONNASSAINCE (ISR) CYCLE

The Intelligence, Surveillance, and Reconnaissance (ISR) cycle describes the manner in which a UAV unit is tasked and depicts the flow of information from the UAV system to the end user(s). Of necessity, the ISR cycle for each UAV differs depending on the capabilities of that particular UAV system and the type of data provided. The ISR cycle applicable to the UAV unit may help determine the scope and depth of test support assets required to conduct an adequate OT. If there are units that provide support to, task, exploit, or receive imagery from the UAV unit under the envisioned ISR cycle, then these units may be needed to ensure the success of the OT.

A minimal UAV ISR cycle might consist of a single tasking UAV unit/user. Figure III-1 presents a simple ISR cycle for a small UAV. In this case, there is only one source of tasking for the UAV as well as one user of the information generated by the UAV unit. An example of a UAV that employs such a limited ISR cycle would be a small handheld UAV that supports company- or battalion-sized operations. In this case, it may be possible to limit the OT to the company- or battalion-sized unit that employs the UAV. However, depending on the level of coordination required between adjacent and higher echelon units employing UAVs, a larger-sized test may be appropriate.

In contrast to the simple ISR cycle presented in Figure III-1, Figure III-2 represents the ISR for a tactical UAV. In the tactical UAV ISR cycle, the unit is tasked by its parent headquarters element. However, imagery from the UAV unit may be piped directly to ancillary units, such as an artillery battery. In this case, ancillary units are usually provided with a remote video terminal (RVT) where they receive imagery but exercise no control over the UAV itself. While these ancillary units are unable to directly task the UAV unit, they are able to request specific tasking though the controlling unit. Testing for this type of ISR cycle may require that the communications nodes from the ancillary units and the RVT be included in the test design.

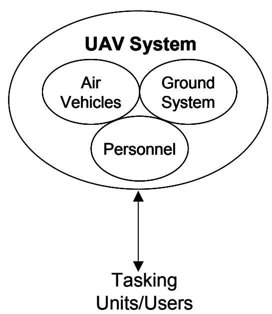

Figure III-1. Simple ISR Cycle

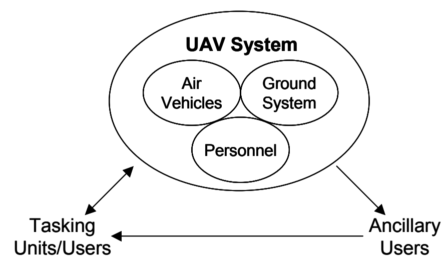

Figure III-2. Tactical UAV ISR Cycle

A more complicated ISR cycle is presented in Figure III-3. In this instance, a strategic level UAV asset receives tasking from a controlling headquarters; however, it does not receive imagery directly from the UAV. Instead, an imagery exploitation unit produces imagery and intelligence summaries for the controlling headquarters. This intelligence product may also be sent for internal use or promulgation to various ancillary units. In a test of this UAV system all units shown in Figure III-3 should be represented during the test, and assets that support these units should also participate in the testing. Finally, all communications links should be representative of those employed during operational deployments.

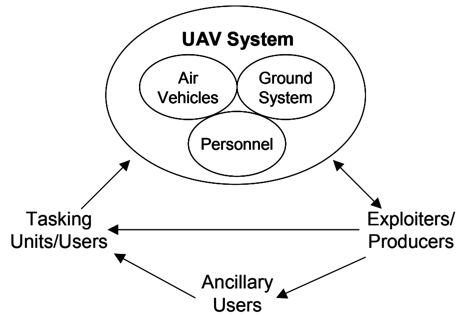

Figure III-3. Strategic UAV ISR Cycle

B. OPERATIONAL ENVIRONMENT

The operational environment in which a UAV system is tested greatly affects an evaluation of the systems operational effectiveness. Operational environment as used here is not limited to the weather and terrain conditions encountered during the test. Instead, operational environment is used to describe the scenario, targets sets, threats, OPTEMPO and other factors that affect an evaluation of system effectiveness.

An operational environment that is not representative of one likely to be encountered by the UAV may misstate the true capabilities of the UAV during combat operations. Therefore, great care should be taken to ensure that the operational environment employed during testing represents the one we anticipate the UAV will encounter. Test design problems may arise should the UAV be procured to operate in several operational environments. In this case, testing should vary the operational environment to the greatest degree possible under existing cost and time constraints.

The following sections outline some key aspects of the operational environment that should be reviewed during the test design phase of OT.

1. Test Scenario

Past testing has indicated that the scenario used during the course of testing can affect the assessment of UAV effectiveness. The set of targets employed, reflective of those likely to be encountered in the chosen scenario, may overestimate system performance. For example, during the Hunter Limited User Test (LUT) two scenarios were used, each with a different set of targets (Ref. 5). System performance, as measured by target detection rates, varied significantly between the scenarios. In the Mid-Intensity Conflict (MIC) scenario, a detection rate as high as 92 percent was demonstrated.

However, in the Low Intensity Conflict (LIC) scenario, a detection rate of 40 percent was observed.[1]

By contrast, the TUAV IOT consisted of a single scenario with a limited number of target sets and types. Specifically, there was no attempt to cover or conceal any of the target sets. While the capabilities of the TUAV may be clearly understood under these conditions, its performance in a more stressing operational environment is unknown.

In order to accurately gauge system performance, multiple scenario types should be employed during UAV OT. Within each scenario, targets sets should be composed of those elements likely to be encountered during that particular scenario. For example, in a MIC scenario, one would expect to see command posts (of various sizes), logistics convoys, mechanized units, infantry units, surface-to-air missile sites, artillery units, armored vehicles, and troop formations. In a LIC scenario, likely target sets might include guerilla headquarters, bivouac sites, logistics points, ambush points, small troop formations, and roadblocks.

Test scenarios should be representative of the threats likely to challenge U.S. forces in the future, so all test scenarios must be based upon current Defense Planning Guidance scenarios. The number and types of target sets to be included in each test scenario should be sanctioned by the Defense Intelligence Agency (DIA) prior to DOT&E approval of the OTA test plan.

2. Test Sites

The land/air space available, terrain, vegetation, and weather at the proposed test site all have the potential to impact the OT results. Among those items that could be affected are assessments of target detection and recognition, system survivability, navigation, aerial relay (if required), imagery performance at maximum range, response time, and target coverage. Available ground space should allow for ample separation of target sets to accurately define system performance. The size of the operating area during testing should mirror that of the one stated in the system's concept of operations.

The use of a single test site could restrict our understanding of how well the system would work in other environmental conditions. Of particular concern is an assessment of the payload operator's ability to detect and recognize targets in different visibility, clutter, and cover conditions. Most importantly, the test site should not be

---

1 Detection rates varied among type of target set and the characteristics of each target set.

located at the home base of the test unit. Familiarity with the local operating area may not enable the test unit to fully utilize the systems capabilities and could skew test results.

As an example, consider a depiction of the flight path from a Predator test mission presented in Figure III-4. Note that most of this mission was conducted in the vicinity of the Indian Springs base, where the Predator is based. It may not be possible to objectively evaluate the search, detect, and target locating capabilities of the UAV when the operators are overly familiar with the operating area in which the test is conducted.

In order to truly evaluate system performance, testing should be conducted at multiple sites. For example, testing in a high desert environment could lead to optimistic conclusions about target detection and recognition performance, and to pessimistic conclusions pertaining to survivability. In order to balance the evaluation, additional testing should be conducted in an environment characterized by higher humidity, denser vegetation, and rolling hills. Testing should also be conducted in coastal and maritime areas, as required.

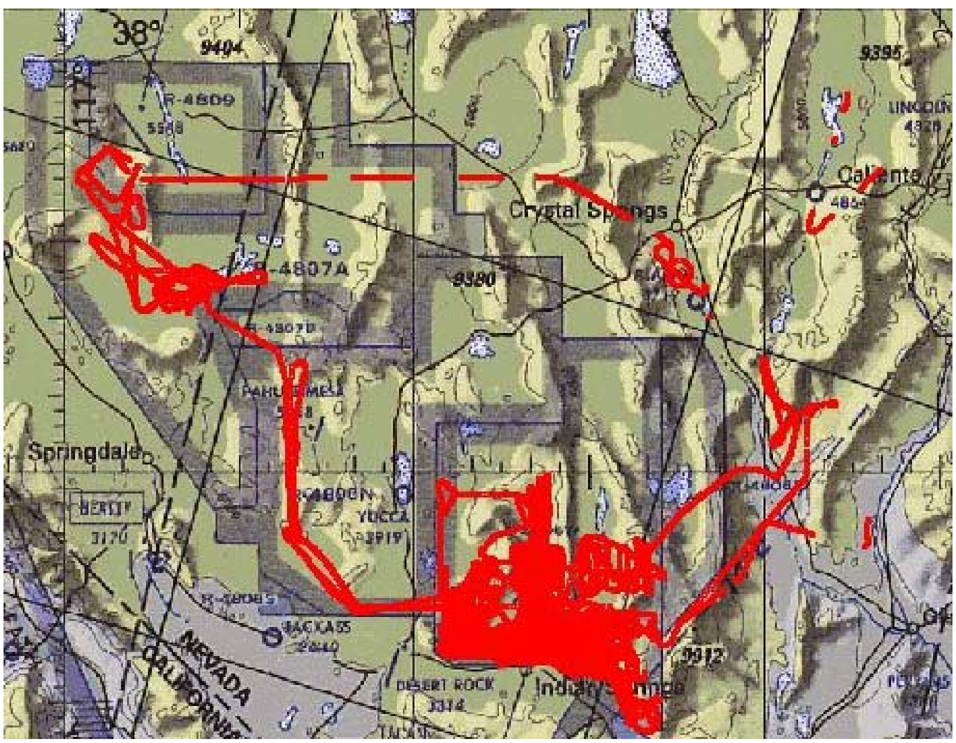

Figure III-4. Sample Predator Flight Path

3. Launch and Recovery Sites

There has been a trend (TUAV and Hunter) to conduct UAV launch and recovery operations from surfaces that do not meet the criteria set forth in requirements documents. The importance of operating from surfaces defined in the ORD was illustrated during Operation Desert Storm. At that time, the prerequisite of constructing hard surface runways was a severe constraint to operating the Pioneer UAV and required considerable cost, time, and manpower. The Army found that rather than using up scarce engineering equipment to construct UAV airfields in a combat environment, it needed a UAV system that could operate from unimproved surfaces (Ref. 2).

In the case of Hunter, ORD requirements specify that the system be capable of operating from an unimproved flat grass or dirt surface measuring 200 meters by 75 meters. Nevertheless, two runways with improved surfaces were used during the Hunter LUT. While unpaved, the runways were graded and packed by heavy road construction equipment. Furthermore, the dimensions of both runways exceeded the required distances. One runway measured 1,500 meters by 30 meters while the other measured 300 meters by 75 meters. Both of these strips could be considered "prepared" surfaces, to the extent that one was used to support C-130 operations. Out of safety concerns, only rocket-assisted take-offs were conducted from the shorter, 300-meter strip. Test results showed that the distance required for most recoveries of the air vehicle exceeded the 200-meter limitation (Ref. 3). No operations were conducted from unimproved areas.

Similarly, the TUAV ORD sets a threshold requirement that the system be capable of launch and recovery on an unprepared soccer field-sized surface[2] without engineering equipment to prepare or maintain that surface, at a density altitude of 9,000 feet. However, the primary landing site used during the TUAV IOT test was Shadow Field, a TUAV operations facility owned by the 104th Military Intelligence Battalion (Figure III-5). This field was prepared with Corps engineering assets prior to the LUT conducted in 2001. At that time, this strip did not meet ORD requirements and was not operationally representative of areas that would normally be available to a maneuver brigade commander; however, through erosion and lack of upkeep, it has deteriorated significantly. While not considered an improved site at the time of the IOT in 2002, the strip was considered better than average in an operational environment (Ref. 4).

After the conclusion of the TUAV IOT, the program manager conducted an excursion on an unprepared site selected by the TUAV platoon (Figure III-6). This site was considered closer to the intent of the ORD requirement. During the excursion, two air vehicle recoveries were conducted on the site, but no data were collected on damage

2 Further surface specifications include rocks no larger than 2 inches in diameter, stumps higher than 2 inches, and level within a 5 percent grade.

to the air vehicle. While this site allowed the TUAV to demonstrate an unprepared surface recovery capability, the small number of landings and lack of data collection make it impossible to determine the long-term effect upon the system of such operations.

Figure III-5. TUAV IOT Primary Recovery Site

Figure III-6. TUAV IOT Excursion Recovery Site

Granted, operations from extreme launch and recovery sites are a high-risk proposition. It would be counterproductive to lose all air vehicles early in a test and prematurely terminate the test period; however, it is also counterproductive to conduct a limited "demonstration" at the end of the test period and claim success. Future UAV OT should conduct the majority of the flight operations from ORD-compliant launch and recovery surfaces. AVs that are not able to reliably operate from such surfaces cannot effectively support the commander in the envisioned operating environment.

4. Air Defense Threat

An operationally realistic air defense threat should be portrayed during testing. There are several reasons to include such a threat. First, a viable air threat restricts the amount of data generated by the UAV to an expected level given a combat deployment. If there is a viable threat, the UAV will not be flown (in most cases) within the engagement envelope of the threat systems; however, if a target within the air defense zone is of such high value, the commander should be presented the opportunity to make a conscious decision to risk the UAV during the mission. A viable air threat during the test allows the commander to make these decisions.

Lack of an air threat allows operators to image targets at a much closer range than they would normally do. Many ORDs have specific standoff altitude and distance requirements so the UAV is able to image targets while reducing the threat to the UAVs themselves. Without an air threat, UAVs can get close to targets and possibly increase target detection rates and decrease target location error.

As an example, during the TUAV IOT, only seven percent of the reports imaged targets beyond the daytime standoff ranges specified in the ORD (18 percent occurred beyond the nighttime range). Throughout the course of the test, the UAV was observed to fly and loiter directly over threat air defense systems. In fact, there was only a 6-hour period (out of 242 operating hours) during the 2-week test when the UAV honored the air defense threat, which resulted in 15 unexecuted tasks. Based on these observations, it was felt that the results obtained using IOT data probably overestimated the capabilities of the UAV in terms of TLE and detection rates (Ref. 4). A similar observation was made at the end of Hunter testing (Ref. 3).

Another reason to employ a viable air threat is to enable the UAV unit to employ the entire spectrum of mission planning and coordination links required during real world operations. If the UAV is allowed free reign, there may be no need for the unit to preplan missions and track the location of identified air defense systems using other intelligence assets. During operational deployments it is imperative that the locations of new air defense systems be coordinated with the UAV unit so that AVs are not needlessly lost. The portrayal of a viable air defense threat allows the unit to plan and execute missions while obviating existing threats and conducting real time avoidance of newly identified threat systems.

5. Target Sets

The targets used during UAV testing directly affect the demonstrated capabilities of the UAV sensor system. Targets used during UAV OT have evolved from the earliest test to the latest UAV testing. Unfortunately, this evolution has been in a negative direction, resulting in less complicated, and potentially less challenging, target sets. During Aquila testing, targets consisted of a mix of vehicles – camouflaged/non-camouflaged, moving and stationary, hot and cold. The recent TUAV IOT test consisted of stationary targets located in open areas with no attempt to cover or conceal the targets. Exposed targets set in wide-open spaces are easier to detect than covered or concealed targets set in wooded areas. Future targets sets should revert back to the standards set by early OT.

Several terms must be defined prior to discussing targets employed during OT.

- Target site: The physical location of a target element on the ground.

- Target area: The area encompassed by a target set.
- Target element: A single entity, such as a single tent, vehicle, or person.
- Target set: Any set of target elements. The site location, number, and types of target elements, and the target area define a target set.

a. Target Sites

Target sites should accurately reflect the locations that will most likely be occupied by each type of target set since the number of targets detected and recognized can be significantly affected if careful consideration is not given to target sites during the test planning process.

Great care must be given when choosing target sites. For example, during the Aquila OT, the UAV took up a single loiter position off to the side of the area to be searched. Target sets in the tasked area included vehicles next to stands of tall trees. If that UAV loiter position was on the wrong side of the trees, the vehicles would be obscured. Post test analysis revealed that only 60 percent of the uncamouflaged targets and 30 percent of the camouflaged targets were detectable from the chosen loiter positions.[3]

During the TUAV IOT, targets were located exclusively in open areas near roads. Table III-1 shows the percentage of target sites located within a given distance of a primary or secondary road. For example, given an air vehicle altitude and mission payload field of view combination that allows the Mission Payload Operator (MPO) to view 100 meters on each side of the aimpoint, 58.2 percent of the target sites would be visible to the MPO while focused on a primary road. Given 600 meters on each side of the aimpoint, 80 percent of the targets would be visible to the MPO while focused on a primary road. In all, 69.1 percent of the target sites are located within 100 meters of a primary or secondary road.

The close proximity of each target site to a primary or secondary road, while representative of the scenario, provides a significant advantage to MPOs tasked to locate targets at specific geographic points. This could significantly reduce the time required for the TUAV system to complete mission tasking and may not be representative of performance in the operational environment envisioned by the ORD.

---

[3] It could be argued that this is a failing of the UAV operators and not the test design. One hypothesis held that better detection rates could have been achieved had the operators observed tasked locations from multiple viewpoints instead of a single loiter point.

Table III-1. Percentage of Target Sites Within a Given Distance of Primary or Secondary Roads

Attack Point	Distance from Attack Point					
	100m	200m	300m	400m	500m	600m
Primary Rd	58.2	65.5	74.5	80.0	80.7	80.0
Secondary Rd	60.0	76.4	87.3	90.9	92.7	92.7
Intersection	7.3	12.7	21.8	25.5	27.3	29.1
Primary or Secondary Rd	69.1	80.0	89.1	92.7	92.7	92.7

Note: The distances chosen are merely for discussion and do not represent actual altitude and zoom combinations.

Another example from the Shadow IOT showed a disconnect between the test sites chosen by the test agency and the method the test unit chose to employ the TUAV. In this case, the TUAV was tasked to observe several Named Areas of Interest (NAI). A review of the target sites reveals that only eight of 62 target sites (13 percent) were located within the NAIs of interest to the test unit. As a result, UAV operators spent large amounts of time observing areas without the opportunity to report on instrumented targets.

b. Target Area

The area occupied by any target set should be in accordance with the DIA-approved deployment for those target elements. The surveyed Global Positioning Satellite (GPS) location of each target element should be available for posttest data reduction and analysis.

c. Target Element

Each element of a target set should be consistent with the target elements expected of that particular target set. The number and types of each element should reflect expected ones given the target set and level of conflict.

Target elements should include personnel, tracked and wheeled vehicles, general-purpose tents, bridges, revetments, trench lines, antennas, air defense systems, artillery tubes, and boats. Additionally, target elements should be camouflaged, not camouflaged, hot and cold (from an IR perspective), moving and stationary.

d. Target Set

The characteristics of each target set are critical to adequate testing of the UAV system. Whether a target set is moving or stationary, hot or cold, camouflaged or not, and the number and type of target elements within the target set, all have an effect on UAV performance. Test design must consider all of these characteristics and ensure that sufficient target sets are available and that the UAV has the opportunity to detect an adequate number of targets of each type to ensure statistical significance.

Prior to approval of OTA test plans, precise details of the makeup of proposed targets sets should be reviewed. Table III-2 shows the target sets presented during the Hunter LUT (Ref. 5). This table reveals the level of detail required to make an accurate assessment regarding the adequacy of proposed target sets. The goal here is to provide the UAV with an opportunity to detect an adequate number of target sets of each type or configuration to ensure that strong conclusions regarding system effectiveness will be derived from the test data. Additionally, this detailed review may ensure that adequate test resources have been allocated for the test. Ideally, this level of detail should be developed to support TEMP Part V inputs.

Table III-2. Sample Target Arrays to be Presented to UAV

Target Mode	Reconnaissance Mode			Surveillance Mode		Total
	Area	Route	Point	Location	Moving	
Moving	144	144	144	24	24	480
Stationary						
No Camouflage	144	144	144	14	8	454
Hasty/Natural Camouflage	72	72	72	24	10	250
Full Camouflage	78	78	78	14	12	260
Total	438	438	438	76	54	1444

Note: Table repeated for day and night operations.

The composition of the target sets, in terms of number and type, utilized during OT should reflect those likely to be found during the actual scenario. As mentioned previously, multiple scenarios should be employed, and multiple families of target sets would then be presented to payload operators. These scenarios should be played sequentially so that the target sets would "make sense" and not further confuse the data analysis process.

6. Operational Tempo

The OPTEMPO of the test unit should reflect the one the unit anticipates encountering during operational commitments. Otherwise, the system and operators performance might not be consistent with performance during combat deployments. Examples from past UAV testing highlight the proper use of projected OPTEMPO.

An assessment of Predator capabilities during combat deployments revealed that the mission availability of the Predator system was quite good; however, the unit was equipped with at least three air vehicles, and was typically tasked to fly only six missions per week, many of which were cancelled due to weather, which meant the maintenance team was not severely stressed.

The Hunter system in the LUT was designed, equipped, and staffed to conduct operations on a 16-hour shift basis utilizing 20 people. The actual execution of the LUT only included three missions per week: one on Monday and two on Wednesday; Tuesdays and Thursdays were set aside for system maintenance and crew rest. Several of the scheduled test flights were terminated short of the scheduled completion time, which again left the operators and maintainers not operationally stressed.

The TUAV IOT poses an excellent example of the system being stressed as it would be during operational deployments. During the 2-week test period, the unit flew 226.9 hours, demonstrating the capability to sustain operations as defined in the TUAV OMS/MP. The number of flight hours during the test was a tribute to the unit, and also test design, which sought to stress the unit as envisioned by TUAV operational concepts.

Future OTs should maintain an OPTEMPO that reflects the projected system use during contingency operations. Anything less could underestimate manning requirements and AV usage rates.

C. TEST UNITS

It is important that the test unit operates and is stressed at levels envisioned during operational deployments. Of particular concern is the amount of available upper echelon maintenance support and spares and the requirement for split-site operations.

If spare air vehicle and upper level maintainers are located at a higher level common to multiple UAV units, then these assets should be portrayed during the test as if they were supporting all of the UAV units simultaneously. For example, under the TUAV concept of operations (CONOPS), multiple TUAV platoons depend upon the Division Mobile Maintenance Facility (DMMF) for spares and upper echelon maintenance support; however, during the TUAV IOT, the DMMF was only required to support the test platoon. Provisions were made for administrative delay times, but this cannot adequately stress the DMMF, and might possibly delay TUAV operations, and so doesn't offer a true idea of the maintenance requirements of the TUAV system.

Some of the systems tested to date have been required to conduct split-site and concurrent operations. During testing, only a "slice" of these units participated, ostensibly representing one operating site worth of personnel and equipment; however, only evaluating a slice of a UAV unit does not fully replicate the stress placed upon the units leaders, the command and coordination requirements, and the logistic burden upon the unit. Past testing has also shown that "slice" testing adds personnel and equipment that would not normally be present. Testing a slice could also result in a "golden crew" because the slice of the unit being tested is stacked with a higher level of experience than might be found in a unit forced to operate from several sites separated by a great distance.

Future OT should employ a full UAV unit manned and operated according to standing operational concepts. The additional tactical tasks expected of the UAV units (such as perimeter security) should be included in the test design. If they are not, then the manning levels of the unit, determined in part by OT, may be set so low that the operational effectiveness of the unit is adversely affected during contingency operations. By the same token, testing a slice of a UAV unit in isolation does not adequately stress the unit or its resources.

Also requiring careful consideration is the size and composition of the tasking unit during the operational test. The tasking unit should be composed and operate like the tasking unit in the envisioned operating environment. All staff personnel and tactical operations center connections should be established and exercised. Furthermore, the tasking unit should have a commander whose actions under the test scenario can be influenced by ISR data gathered by the UAV unit.

The target set planned for operational testing should be designed to meet thresholds established by the commander's priority information requirements. During the Shadow IOT one of the Brigade commanders priority information requirements was to identify occasions where 100 heavy equipment transport trailers were gathered. However, the target set was limited to sets of two to eight utility vehicles. As a result, the Brigade commander was never forced to react to data gathered by the Shadow platoon. This makes it difficult to determine the contribution of the UAV towards unit effectiveness.

D. INSTRUMENTATION

Test instrumentation and data collection are critical to the successful resolution of operational effectiveness, operational suitability, and addressing test adequacy. Without proper instrumentation, it may not be possible to conduct posttest analyses and resolve UAV performance as measured against the thresholds delineated in requirements documents.

1. Test Instrumentation

Test instrumentation should allow for precisely recording the AV flight path and sensor position. From this data, it's possible to approximate the field-of-view of the payload to assist in data reduction efforts. Data required for this effort includes AV heading, altitude, position, pitch and roll, payload depression angle, and payload azimuth. For instrumented targets, data regarding location, elevation, target name and type, and target characteristics must be collected.

Each target should contain a mechanism for recording target locations at all times during UAV test missions. This location data is critical in order to estimate TLE, percent of targets detected, and to understand why targets went undetected.

Operations within the GCS should also be recorded. Video cameras should monitor the operator's computer screens and record all inputs. This proved useful during the TUAV IOT because it allowed analysts to identify operator error several times during normal operations. For example, when tasked to search grid coordinates 123 456, a review of the GCS video shows that the operator mistakenly input 123 546 into the auto-point dialogue box. As a result, the operator searched the wrong area and submitted an erroneous report.

Also required is a personal computer-based "playback" capability so that AV and instrumented vehicle positions as well as the approximate payload footprint, can be synchronized and reviewed. This playback feature has proven useful during past tests for two major reasons. First, analysts are able to determine the amount of time spent searching for targets and whether the operator searched in the correct location. Second, by watching the approximate payload footprint, analysts are able to identify targets that went unreported by the UAV operator. Additional analysis can then focus on these targets to identify system or operator limitations.

2. Non-Instrumented Targets

Depending on the test site, it might be almost impossible to totally eliminate non-instrumented targets from the operating area. While battlefield clutter is expected during operational deployments, non-instrumented targets add confusion for the operator and the analyst during OT. The time spent imaging non-instrumented targets reduces the time spent looking for the cued instrumented targets. This artificial distraction could adversely affect search and target detections rates.

During the Aquila testing, 78 percent of the targets reported by system operators were non-instrumented targets. During the TUAV LUT, too few instrumented targets were imaged and reported to draw any conclusions regarding TUAV effectiveness. As a result of concerns expressed by DOT&E after the TUAV LUT, its design was adjusted to ensure enough instrumented targets would be imaged and reported. Through prior arrangement, the TUAV test team obtained (almost) sole use of a portion of the Fort Hood training area; yet despite this effort, nearly half (44 percent) of the reports of imaged vehicles were of non-instrumented targets.

To the greatest extent possible, future UAV testing should have exclusive control of the test area. This should not preclude other units from using the training area (provided they fit the scenario in play). Rather, these units should be instrumented and the data harvested so that the data collection analysts have an accurate ground truth picture. If need be, the test team can insert battlefield clutter in the form of adequately instrumented civilian vehicles and personnel.

E. TARGET LOCATION ERROR

For ISR UAV systems that are required to provide precision coordinates, one of the most important metrics available to evaluate effectiveness is TLE. In UAV ORDs, TLE is either defined in terms of CEP or Spherical Error Probable (SEP). The difference between the two is that SEP contains an elevation element, whereas CEP does not.

Extensive use of instrumented targets and ground truth data is required to measure TLE during OT. After the UAV unit submits a report (containing a target location), data collectors must review ground truth data in order to determine the ground truth location of the reported target set. During this process, care must be taken when the subject report involves a set of targets, in which case, the data collector must ensure that the vehicle in the center of the image is the same vehicle for which ground truth data is being derived. In the TUAV IOT, where the required CEP was less than 80 meters, the distance between two vehicles in a target set could result in an adverse effectiveness analyses should the data collector use ground truth data for the wrong vehicle.

In some reports, TLE may appear to be abnormally large. For instance, there were target reports during the TUAV IOT where the observed TLE exceeded 1 kilometer. In these cases, the cause of the large TLE should be determined. In some of the TUAV reports, the wrong ground truth data were used during the data reduction process. In others, UAV operator error was to blame. Nevertheless, reports that involve excessively large TLE values, in particular, should be fully understood in order to identify potential system limitations.

Before the start of OT, analysts must understand the subsystems and algorithms employed by the UAV in order to compute target location. This knowledge can be helpful in setting an expected TLE during the test. For example, the subsystems used in the test UAV could be compared to subsystems in existing UAVs with known TLE capabilities. This review should allow analysts to determine if the TLE values shown during testing are close to those expected. If not, there may be systemic issues, such as subsystem integration or operator training or proficiency that need to be explored.

F. ALTERNATIVE APPROACHES TO OPERATIONAL TESTING

In order to reduce the time and cost of OT, many programs propose the use of data gathered during non-OT events. Combat deployments and non-OT testing provide viable assessments of UAV capabilities; however, a careful review of the data should be conducted prior to the approval of any test strategy that seeks to employ non-OT data. Besides providing little if any control over the elements of the operational environment described above, the use of alternative approaches to OT present unique pitfalls that should be carefully examined. Any data collected outside of formal OT periods should pass an “operational realism” litmus test.

Non-OT events present excellent opportunities to collect qualitative and anecdotal information; however, these events should only be relied upon to assess the *potential* effectiveness and suitability of UAV systems and not to draw definitive conclusions on which major programmatic decisions are then based.

1. Use of Non-OT Data

The conditions under which the non-OT data was collected and the nature of the targets employed and missions flown should be examined. Also, the operators’ experience level should be judged representative of what will be encountered with fielded units. The conditions in which the data was collected as well as the target sets employed should be diverse and reflect those anticipated during operational use.

The quality of the data should also be examined. During the Predator IOT, it was expected that previously collected RAM data could be used to augment that collected during the IOT. However, this data, stored in the Core Automated Maintenance System (CAMS) suffered from numerous inconsistencies that rendered it unusable for post-IOT analyses.

Before embracing a test strategy that relies upon non-OT data, the proposed data set should be reviewed by DOT&E for accuracy and usability. Should doubts arise regarding the conditions under which the data were collected or the quality of the data itself, the test strategy should be rejected. Efforts should be made to ensure that only operationally relevant data be included in the scope of any OT.

As an example, during TUAV DT, the TLE CEP demonstrated by contractor personnel under ideal conditions was 115 meters (average 129 meters). During IOT, school-trained military operators in an operational environment achieved a demonstrated TLE CEP of 191 meters (average 224 meters). The disparity in demonstrated TLE CEP between the contractor and military operators shows that the use of non-OT data may not be a true indicator of system performance in an operational environment.

2. Combat Deployments

User satisfaction with developmental systems employed during combat operations has proven difficult to quantify and evaluate. The first and foremost shortfall with data collected during combat deployments is a lack of ground truth data. Without formal test and evaluation controllers and data collectors, it is almost impossible to harvest and reduce accurate ground truth data. Without such data, it is impossible to assess detection rates and reporting accuracy among other metrics of importance to the user.

The Predator early operational assessment (OA) utilized data collected during one training exercise and two real world deployments. While these data were valuable, no definitive conclusions could be reached regarding target detection rates since the total number of targets in the operating area was unknown.

Care must also be taken with combat deployments since the UAV system might not be tasked, operated, and report results as envisioned by the end user. Normally, events conducted early in a program's life require heavy use of contractor maintenance and support personnel, which, combined with unfamiliarity with system capabilities and limitations on the part of staff planners, could result in the UAV not being employed as envisioned by the requirements generators and system developers. As a result, gathering comprehensive effectiveness and suitability data may not be possible. Along these lines, the OPTEMPO of the unit during non-OT events may not stress the UAV unit to the level envisioned by operational documents.

During real world operations, access by imagery users may be limited. This would preclude assessing the contribution of the UAV towards overall mission success. It is also possible that a lack of system experience on the part of the tasking agencies would result in the payload being tasked to execute missions not commensurate with its abilities.

G. SURVIVABILITY TESTING

Current TEMPs say little about UAV survivability testing (not only the air vehicle but the data links as well).[4] There was no live fire testing planned for the Army TUAV, nor is it planned for the Predator-B or Global Hawk.[5] A reasonable strategy may be to purchase and field *some* UAVs without assessing their survivability simply to use them to exploit their capabilities, despite not knowing their vulnerabilities. However, survivability should be understood to assess both the costs and methods of planned improvements, and to understand the cost of employing UAVs in support of rapid maneuver forces on nonlinear battlefields. Any survivability assessment should be made in light of the modern weapons that the systems might be faced with (Ref. 10).

The scope of survivability testing for any UAV could be based upon the acceptable level of attrition as well as the value of the data provided by the UAV. Expendable UAVs may require no survivability testing because these UAVs will be launched and are not expected to return safely. Attritable UAVs, while expected to suffer loses, should complete some level of survivability testing in order to provide information regarding expected losses. As a general rule of thumb, the more expensive the UAV (and its sensors) the greater the requirement to conduct detailed survivability testing.

---

4 Past UAV testing has included survivability assessments (Refs. 6,7,8). Reference 9 describes UAV combat deployments to Bosnia and the Persian Gulf.

5 UAVs are not considered “covered” systems and so are not required by Title X to conduct life fire testing.

1. Susceptibility

UAV OT should seek to understand the AV's susceptibility against anticipated combat threats. Can these threat systems detect, acquire, track, and engage (i.e., launch, intercept, and successfully affix damage mechanisms to) the AV? Susceptibility reduction can be accomplished with mission planning, signature management, aircrew warning devices, speed, and altitude.

Threat systems employed during OT (using actual, surrogate, or simulated systems) should match those likely to be encountered by the UAV as described in the systems threat assessment report. Of utmost importance is that UAV operators honor the portrayed threat. During the TUAV IOT, the test unit was instructed to use tactics, techniques, and procedures to avoid known threats, but for the most part, the air vehicle was flown over threat territory with impunity and observed targets from less than the required standoff distances with increased loiter times over targets.

In order to measure UAV susceptibility, the following areas should be evaluated:

- Probability of Detection
- AV Signatures
  - Radar Cross Section
  - IR Signature
  - Aural and Visual Signatures
- Aircraft Survivability Equipment
- Tactics

2. Vulnerability

There has never been a systematic live fire test program conducted on a UAV. Arguments against live fire testing include the cost (in terms of dollars and air vehicles) of such testing relative to overall procurement costs; the cost to implement vulnerability reduction features; and the view that UAV systems are expendable to some degree. However, in light of the rising costs of such systems, the benefits of such testing should be explored.

Designed vulnerability reduction features can enhance air vehicle effectiveness during peacetime and wartime missions. Design features that minimizes the probability of losing an aircraft due to damage (accidental or threat induced) reduce the loss of potentially valuable airframes. During peacetime, the dollar cost of replacing damaged airframes drains limited resources from other programs. During wartime, besides the dollar cost to replace airframes there is the possibility that commanders will be forced to conduct combat operations without the benefit of a valuable ISR asset. The benefits to be derived from even a rudimentary vulnerability assessment have the potential to far outweigh the costs involved.

In the absence of live fire testing, an extensive analytical effort could be made to determine the vulnerability reduction characteristics of the AV itself. For example, a quick examination of the Predator AV conducted by members of the Joint Technical Coordinating Group on Aircraft Survivability (JTCG/AS) revealed that critical electrical and fuel lines were routed next to each other through the fuselage that severely increased the vulnerability of the AV. Routing of these aircraft components are a major design consideration in manned aircraft.

3. Electromagnetic Effects Testing

UAV systems should conduct electromagnetic effects testing for two reasons. First, such testing may identify vulnerabilities to the UAV system from threat weapon systems. Second, and just as importantly, this testing may identify the effects of co-located friendly equipment that can impact UAV operations. This testing is especially useful for those UAV systems that will operate in a shipboard environment.

The family of electromagnetic effects testing that should be considered includes:

- Electromagnetic radiation operations
- Intersystem electromagnetic compatibility
- Helicopter electrostatic discharges
- Hazards to fuel of electromagnetic radiation
- Hazards to personnel of electromagnetic radiation
- High altitude electromagnetic pulse
- Near-strike lightning.

CHAPTER IV

TEST EXECUTION

IV. TEST EXECUTION

A carefully planned OT should require little attention during the actual execution phase; however, certain aspects of the test should be monitored to ensure that the UAV is employed and evaluated in a manner that reflects the users' desires.

A. CHANGES TO THE APPROVED TEST PLAN

DOT&E is the ultimate approval authority for operational test plans. Testing conducted in the field should be monitored to ensure that the intent of DOT&E's approval is being followed to the greatest extent possible. Certainly, there may be circumstances that dictate adjustments to the existing test plan. Should such occasions arise, DOT&E should be notified and briefed on the change and its impact upon test adequacy. If there is the possibility that the change would prove detrimental to test adequacy, mitigation measures should be identified and instituted as rapidly as possible.

B. DATA COLLECTION

It is important that data necessary for posttest analyses be collected and made available in a useable format. A vital piece of data to collect is the total number of taskings to each of the ISR assets available to the test unit. This is needed in order to determine the relative importance of the UAV as an ISR asset in relation to all the assets that could be available to the test unit. During the TUAV IOT, these data were not collected. As shown in Figure IV-1, the number of entries in the column labeled *All BD Tasks* was not known, which makes it difficult to judge the contribution of the TUAV to the commander's ISR plan.

Another important set of data that should be collected is the complete list of tasks intended for the UAV. This list should include those tasks received by the UAV operator as well as those intended for, but never received by, UAV operators. The tasks assigned to the UAV unit should be recorded in real time and cannot be reconstructed posttest or at the end of each day. During the TUAV IOT, the methodology used by the tester to correlate tasks and reports was flawed when, after each report was generated, the tester correlated that report to a task. Under this methodology, every task in the performance database has a report associated with it, but the database did not contain those tasks the TUAV failed to complete. In Figure IV-1, the total number of entries in the column labeled *All TUAV Dynamic Tasks* is unknown due to the data collection methodology employed. Therefore, it was not known how many tasks were not completed by the TUAV platoon; this number could be significant and affect the perceived effectiveness of the TUAV system.

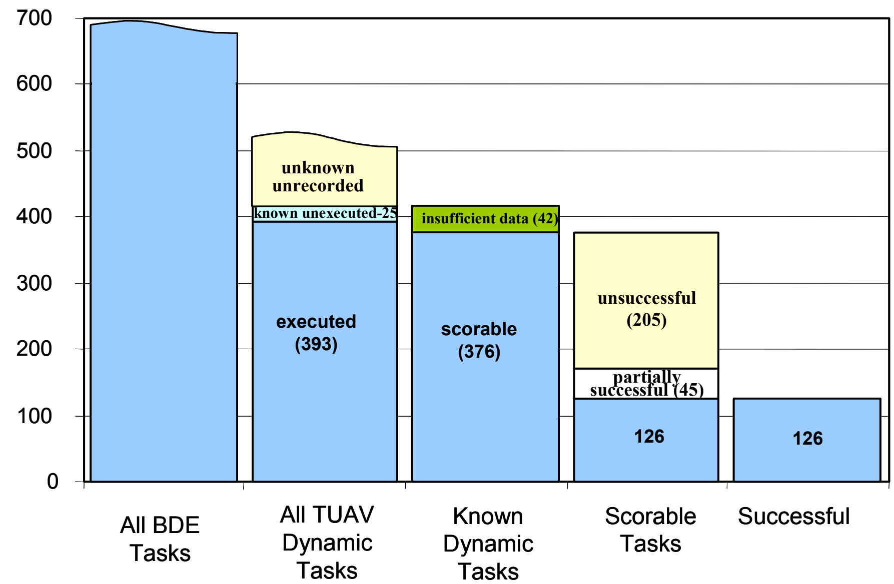

Figure IV-1. TUAV IOT Tasking Success

For every UAV imagery report, the operating mode of the sensor (Electro-Optic (EO), IR, Side Aperture Radar (SAR), etc.) should be recorded. The TUAV ORD states that the threshold requirement for Probability of Recognition, given detection, (Pr(R|D) is 70 percent for the IR and 80 percent for the EO sensor. During the IOT, the payload mode (EO or IR) at the time a target was detected was not recorded, which made determining the percentage of targets successfully prosecuted in each mode impossible.

C. TARGET CUEING

Most tasks assigned to a UAV during OT require the system to search a particular coordinate for a target. Near-perfect target cueing could result in optimistic estimates of the UAV operators' probabilities of detection and a reduction in the time required to detect the target sets. During the TUAV IOT, questions arose regarding the accuracy of the grid locations contained within taskings to the TUAV. Of the 128 reports containing instrumented targets, two of the corresponding taskings contained four-digit grid coordinates, 115 contained six-digit coordinates, and 11 contained eight-digit grid coordinates. The average distance from the tasked grid location to the ground truth location of the correlated report was 146 meters (the median was

71 meters). Similarly, during the Hunter LUT, testers made efforts to ensure that target sets were as close to the cued locations as possible.

It was not possible to derive the cueing source for these tasked grid locations from the TUAV and Hunter databases; however, the accuracy of the tasking appears to be too precise. As one member of the test team stated, “If they want to pass this test all they have to do is fly to the tasked grid coordinates and look down. They will find all of the targets.”

It is doubtful that the TUAV platoon would receive such accurate grid coordinates in its taskings in an operational environment. Such accurate gird coordinates may have reduced the time the TUAV required to search for targets and this, in turn, may have artificially decreased the system’s response time. The accuracy of the grid coordinates contained in the tasks was considered a flaw in the test design.

In future testing, attention should be given to the accuracy of cueing information to ensure that the cueing accuracy reasonably represents the correctness of the cueing device(s), the age of the information at the time the UAV enters the suspected target area, and the propensity for targets to periodically move on the battlefield. Determining system effectiveness should entail more than merely navigating to a known point and pointing the payload straight down.

D. REAL-TIME CASUALTY ASSESSMENT

During the IOT, testers authorized the TUAV to fly over threat territory even though threat air defenses were able to detect the TUAV AVs. On several occasions AVs were observed orbiting directly over air defense systems capable of downing them. This unrestricted ability of the AVs to fly where desired eliminated the operational requirement for them to observe targets from realistic slant ranges, and improved their opportunity to loiter and identify targets. This artificial invulnerability to threat systems may have resulted in a false sense of the operational usefulness of the system in a combat environment.

Currently there exists no real-time casualty assessment system (RTCA) for UAVs. As such, it is usually a subjective assessment on the part of the threat system operator as to whether or not an AV could be engaged and destroyed. Until such time as an RTCA system is available for use during OT, test personnel should monitor UAV operations at the GCS. Such an observer would be in radio contact with air defense system operators and instruct the UAV operators to return the AV to the recovery site should it be successfully engaged. In this way, UAV operators would appreciate the air defense threats and their impact on UAV operations.

E. NON-INTERFERENCE VERSUS DATA COLLECTION

To the greatest extent possible, the test agency should not interfere with the test unit during OT. However, there are exceptions to this philosophy, most notably the need to balance test interference with data collection. One prime example during the TUAV IOT was the need to fly air vehicles during overcast periods when the sensor was incapable of viewing targets. Significant flight hours were flown during which no imagery was collected and disseminated. Under normal circumstances, the test unit would not have flown air vehicles in order to preserve assets instead. In this case, the need to collect operating hours in order to determine system suitability outweighed interference with normal test unit activities.

Another possible reason to interfere with test unit activities is to collect data to obtain a complete understanding of UAV activities. As part of the data collection process during the TUAV IOT, video monitors were mounted throughout the TOC. The purpose of these videos was to capture taskings from the Brigade S-2 to the TUAV operator, so that at the end of each day all UAV taskings could be captured from the video; however, in practice, this did not work. A better method would have been to have the Brigade S-2 log all the taskings to the TUAV GCS. While this is not part of the units' normal operating procedures, the deviation would have been insignificant and resulted in a better pool of data for posttest analyses.

The use of Observer/Controllers (OCs) during OT is another area where test team interference on test unit operations is acceptable. The presence of OCs, or their equivalent, is common throughout Service-conducted training exercises and their appearance within an OT would not be considered out of the ordinary for the test unit. OCs could then be used to influence test unit behavior to ensure that test objectives are met.

During test operations, OCs would ensure that the test unit employs UAVs in the manner envisioned by the Concept of Operations. For example, during the TUAV IOT, only 7 percent of the reports imaged targets beyond the daytime standoff ranges as specified in the ORD (18 percent occurred beyond the nighttime range). Additionally, throughout the course of the test the UAV was observed to fly and loiter directly over threat air defense systems. The reduced imaging ranges may have increased the TUAV operators' detection and recognition capabilities and may not represent the results expected during combat operations. In this case, OCs could have modified TUAV employment by directing greater standoff ranges or administratively "downing" air vehicles. In this way, data collected during the test might be a closer reflection of performance expected during combat operations.

F. ADDITIONAL ISR ASSETS

UAV system testing should include the full spectrum of ISR assets available to the tasking unit. If the UAV is the only ISR asset available to the unit, then it should be expected that the UAV will be extensively utilized, probably to an unrealistic level. The absence of scouts, ground sensors, and aviation assets typically available to the unit might result in an unnatural dependence upon the UAV.

Because including additional systems may increase test costs, the test agency should account for the additional ISR systems during the planning process. The control cell may be able to mimic these assets, but in terms of capabilities and timeliness it should be as responsive as the real assets.

During OT, test unit activities should be monitored to ensure that an over reliance on the UAV doesn’t result in an inflated sense of the UAV value.

CHAPTER V

EVALUATING MISSION EFFECTIVENESS

V. EVALUATING MISSION EFFECTIVENESS

A. REPORT SUCCESS TEMPLATES

One of the more important aspects of effectiveness analysis is the measure by which UAV reports are deemed successful. Criteria for determining the accuracy, timeliness, and contribution of UAV reports may be used to assess the operational relevance and accuracy of UAV performance. Equally important is identifying whether the reason behind poor reports is due to system limitations, operator error, or unit training.

One effective means of evaluating UAV reports is through the use of Report Success Templates (RST) developed for use during the Shadow TUAV OT.[1] The following section reviews RST scoring as used in the TUAV IOT with adjustments for lessons learned regarding their use in future testing. Note that this example applies to surveillance and reconnaissance missions. Should the UAV be designed to conduct other missions, similar scoring should be established for timeliness, accuracy, and contribution criteria reflecting that particular mission.

1. Timeliness

The timeliness of a report is defined by the last time information is of value (LTIOV), which is given as part of the tasking and is based on the urgency of the information. Intelligence information loses value if not provided in time for the supported unit commander to act upon it. Assessing timeliness is done by comparing the time the report is sent to the LTIOV assigned by the tasking unit. The scoring criteria for timeliness during the TUAV IOT was as follows:

0 = Did not report

1 = Did not meet LTIOV, but the intelligence provided was of value

---

[1] During the TUAV IOT, Mission Success Templates were used to evaluate each report. These templates should have been referred to as “Report Success Templates” since they were applied to each individual report generated by the TUAV.

2 = Met LTIOV as required, or, if dynamically tasked, was on-station when required.

3 = Exceeded LTIOV, by some predetermined amount of time.

A successful report achieved a timeliness score of two or higher. If a task did not have an LTIOV associated with it, the report automatically received a passing score of two. Great care should be given during the test design to ensure that the vast majority of the tasks will contain an LTIOV. This may ensure that report success rates are not inflated due to an abundance of reports with a default timeliness score of two because of no LTIOV being assigned to the task.[2]

It should also be noted that the conditions under which a score of three is assigned should be agreed upon in advance of the test. During the TUAV IOT, Army analysts decided (without input from the program office or the user representative), that reports submitted at least one hour prior to the LTIOV should be given a score of three for timeliness. During future tests, the tester and user representative should agree to the cutoff time to ensure an operational foundation for that time cutoff.

2. Accuracy and Completeness

Prior to the test, a set of Minimum Analytical Thresholds (MAT) should be defined to determine the accuracy and completeness of a report. As implied by the definition, the standards reflect the minimum accuracy required for a report to support a correct intelligence analysis; and a report must be of adequate accuracy and completeness to prevent it from being misleading.

The MAT includes standards for size, activity, location, and equipment. As an example from the TUAV IOT, a MAT was developed prior to the start of the test for every type of target set in the scenario (Table V-1).

Of great importance in Table V-1 is the column labeled “Location” which refers to the target location error for the report. Prior to the test, the desired TLE must be extracted from the ORD or other operational documents. The 200 meter limit shown in this table was determined to be the minimum acceptable TLE of use to the user. During other OTs, the desirable accuracy may be greater (or less) then the 200 meters used in this case.

---

[2] During the TUAV IOT, only 16 percent of the reports had an associated LTIOV. This resulted in many reports receiving an automatic two for timeliness and overinflating the report success rate.

Table V-1. Sample Minimum Analytical Thresholds

Actual Target Set			Minimum Analytical Threshold Requirements		
Target Type	**Equipment Type**	**Equipment Quantity**	**Location**	**Equipment Type**	**Equipment Quantity**
SPF Element	CUCV SA-18	CUCV: 2-3 SA-18: 0-1	≥ 200 m	Light Wheel	CUCV: 2-3
SAM Site	CUCV SA-8/SA-9	CUCV: 2 SA-8/9: 0-1	≥ 200 m	Light Wheel	SA-8/9: 1
Radar Site	G-75	1	≥ 200 m	Radar	
Mechanized Company	M113	10	≥ 200 m	Track	6-10
Command Post	CUCV	10	≥ 200 m	Light Wheel	6-10
Logistics Point	CUCV	10	≥ 200 m	Light Wheel	6-10
Logistics Re-supply Point	HMMWV Fueler 5-Ton	HMMWV: 2 Fueler: 2 5-ton: 4	≥ 200 m	Light Wheel Heavy Wheel	HMMWV: 0-2 Fueler: 1-2 5-ton: 2-4
Bridging Activity	AVLB		≥ 200 m	Heavy Track	
Engineer Activity	ACE		≥ 200 m	Heavy Track	
Logistics Company	CUCV	10	≥ 200 m	Light Wheel	6-10

The accuracy and completeness score is determined by comparing the TUAV report to ground truth data and the MAT for that particular target set. The scoring criteria for accuracy and completeness are as follows:

0 = Report contained no accurate information

1 = Provided an accurate report less than the minimum analytical threshold (i.e., less than the required number of target elements or exceeded the maximum TLE value)

2 = Met the minimum analytical threshold or correctly reported negative contact

3 = Met the minimum analytical threshold and recognized the target (tank vs. armored personnel carrier).

Note that a score of two has two different and distinct interpretations. First, there was a target on the ground, it was detected, and the information correctly reported. The second interpretation is that there was no target on the ground, the area was searched, and the absence of a target was correctly reported. The downside to giving the same score for reports of threat targets and negative contact is that the two types of reports are weighted equally. Clearly, it is easier to provide a report of negative contact than one in which the operator is required to provide accurate location, number, and type of vehicles. In addition, if the majority of reports are of negative contact, the TUAV's ability will be inflated. Test design must limit the number of tasks that will result in reports of negative contact.[3]

3. Contribution

Besides timely and accurate reports, the information provided by the UAV should contribute to answering the tasking units' intelligence requirements. In many cases, the UAV will be part of a "system of systems" and not the only intelligence asset available to the tasking unit. The contribution score provides a measure of value added to the tasking unit by having the UAV. The score for the contribution is based on the utility of the report to answer the intelligence requirements (IRs) and priority-IRs (PIRs) of the tasking commander. The report receives increasing scores as its ability to be a sole information source increases; that is, if it answers the IR on its own, it receives a higher score than if it just confirms information from another sensor. The scoring criteria for the contribution are as follows:

0 = Did not contribute

1 = Provided supplemental information that confirms (data already known) situational awareness

2 = Provided supplemental information that enhances (data not previously known) situational awareness

3 = Contributed to answering the commander's intelligence requirement (IR, PIR, Dynamic Tasking)

4 = Contributed significantly to a commander's decision point.

---

[3] During the TUAV IOT, 33 percent of the tasks resulted in negative contact reports.

The user of the information assigns the contribution score for a given report; however, this methodology presents a problem. If the contribution score is given in real time, as the information is received, the contribution will be made without knowledge of the ground truth, in which case it is possible for a particular report to receive a high contribution score for inaccurate, and possibly misleading, information. This would inflate the total report score and overstate the effectiveness of the system. On the other hand, it is entirely possible that a key piece of information could be given a low contribution score at the time it was received, in which case the contribution score would underestimate system performance.

A possible improvement of the methodology would have each report scored for contribution during an after-action period at the completion of each mission. In this way, the true contribution of a report could be established with the aid of ground truth data and a review of mission outcome.

Of the three scores assigned to a report, contribution is far and away the most subjective. Regardless of how contribution scores are derived, great care should be taken to ensure that this score accurately portrays the contribution of that report to the tasking unit's overall intelligence effort.

B. EVALUATING TASK EFFECTIVENESS

The success, or quality, of individual reports is based upon the RST score for each report. The respective RST score for each report is helpful in determining the quality of the reports, but just as importantly, low scores can be used to identify why the quality of a report was less than desirable. Low timeliness scores could point to bulky or complex operating or reporting procedures, and low accuracy scores could highlight system or operator deficiencies.

While overall report success rates may identify system capabilities or limitations, they should not be used in and of themselves as measures of mission success. More important than report success rate is the successful task completion rate. This metric measures the ability of the UAV to successfully accomplish assigned tasks. The difference between report success and task success is as follows: It is possible that several reports could be submitted in response to a single task. It could be the case that each of the reports, taken individually, would score low using the MST criteria, but as a group and applied to a task, it's possible the information was enough to consider the total of the sum of the responses to a *task* successful.

1. Executed Tasks

During data analysis, reports associated with a given task should be reviewed and a binary scoring system employed. If the reports, taken as a whole, were timely they would receive a score of 1 (0 for a non-timely set of reports); if they were accurate, they would receive a score of 1 (0 for inaccurate or misleading reports); if they contributed to the tactical picture, they would receive a 1 (0 for non-contributors).[4] The product of the timely, accuracy, and contribution scores would then be used to measure task success. A product of 1 represents a successful task; a product of 0 represents an unsuccessful task. To be considered successful, a task must result in timely and accurate reports that contribute to the intelligence picture.

2. Unexecuted Tasks

During the course of testing, tasks will be assigned to the UAV unit that, for one reason or another, may not be executed. At the top level, there will be tasks that the requesting unit will send that will never be received by the UAV unit, possibly because of a communication breakdown between the two units. At the lowest level, tasks could remain unexecuted due to ETOS constraints, reliability failures, or AV losses. For this reason, it is incumbent upon the test agency to document each and every task destined for the UAV unit.[5]

One example in which all the tasks were not collected during the data management process was the TUAV IOT. During this test, data were collected for all reports issued by the UAV platoon. After the fact, the data collectors attempted to correlate each report to a task issued by the test controllers; however, there was no corresponding list of tasks that were not executed. Based on experience during previous OTs tests it is highly unlikely that each and every task assigned to the UAV unit will be executed.

Other reasons a task might be unexecuted by the UAV unit include the following:

- Weather over the target area precludes observing the tasked location

---

4 Using the RST format described previously, a timeliness score of two or higher is considered a passing score, an accuracy score of two or higher is considered passing, and a contribution score of one or higher is considered passing.

5 The total number of tasks assigned to the UAV unit must be known in order to implement some of the analytical methods described later.

- A task was unsuccessfully prosecuted (operator error, searching the wrong location, reporting incorrect coordinates)
- The target area was not imaged due to the end of the AV on station time
- Air defense assets precluded the UAV from imaging the assigned location
- Mission was aborted due to reliability issues.

The root cause of all unsuccessful tasks should be explored. During the data reduction process, a set of codes should be developed to easily identify unexecuted missions as well as the casual factor for the failure.

3. Successful Task Completion Rate

The most important metric to measure UAV effectiveness is the task completion rate. This metric provides information regarding the expected number of tasks that may be successfully completed by the UAV.

The use of waterfall charts has proven most effective in summarizing the mission effectiveness of UAV systems. The reader is instantly able to get a feel for the success of the UAV, as well as the major causes of unsuccessful missions. Figure V-1 and Figure V-2 present waterfall charts from previous UAV test programs. Note that in Figure V-2 the number of tasks is known; in Figure V-1, the total number of tasks is unknown. Since this number is unknown, the task success rate based upon data from this test is a best-case estimate. In actuality, the actual task success rate may be significantly lower.

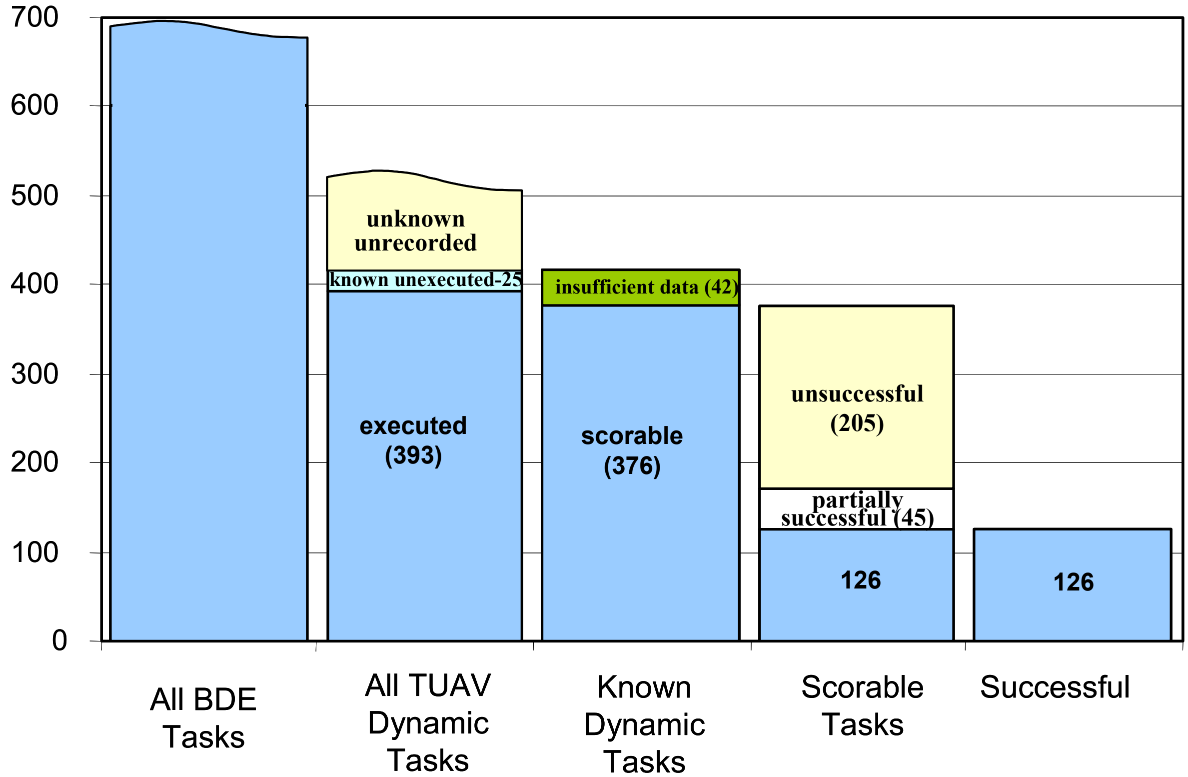

Figure V-1. Shadow TUAV IOT Waterfall Chart

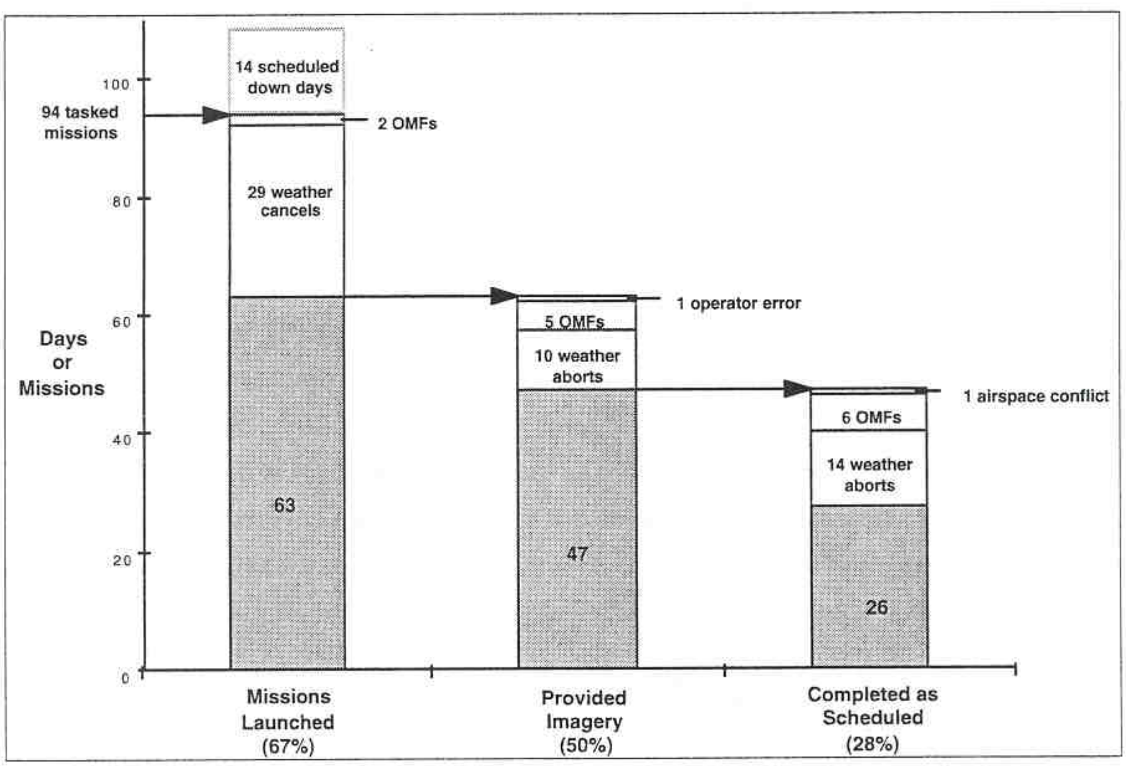

Figure V-2. Predator ACTD Operational Assessment Waterfall Chart

REFERENCES

1. GAO Report, *Unmanned Aerial Vehicle: Outrider Demonstration Will Be Inadequate To Justify Further Production*, 23 September 1997

2. Major Stephen L. Barrett, *Unmanned Aerial Vehicles: The Gulf to Bosnia and Beyond*, Air Command and Staff College, Maxwell Air Force Base, Montgomery Alabama, March 1997

3. GAO Report to the Chairman, Legislation and National Security Subcommittee, Committee on Government Operations, House of Representatives, *Unmanned Aerial Vehicles, Performance of Short-Range System Still in Question*, December 1993

4. IDA Paper P-3732, *Assessment of the Tactical Unmanned Aerial Vehicle (TUAV) Initial Operational Test*, 2 January 2003

5. IDA Paper P-2759, *Assessment of the Results of the Short-Range Unmanned Aerial Vehicle Limited User Test (LUT)-I*, 9 September 1992

6. IDA Document D-842, *Unmanned Air Vehicle Survivability in Test and Evaluation and Combat*, 21 December 1990 (SECRET/NOFORN)

7. IDA Paper P-3732, *Assessment of the Tactical Unmanned Aerial Vehicle (TUAV) Initial Operational Test*, *Annex – Survivability Assessment (U)*, 2 January 2003 (SECRET)

8. IDA Document D-2462, *Unmanned Aerial Vehicle Operations in Support of Operation Allied Force*, 8 September 2000 (SECRET/NATO)

9. IDA Memorandum Report M-345, *Report on the Monitoring of RPV Aquila Testing, Volume 1 Chapter III: RPV Survivability*, 7 October 1987 (SECRET)

10. Major Stephen L. Barrett, *Unmanned Aerial Vehicles: The Gulf to Bosnia and Beyond*, Air Command and Staff College, Maxwell Air Force Base, Montgomery Alabama, March 1997

11. IDA Memorandum Report M-345, *Report on the Monitoring of RPV AQUILA Testing*, *Volume 1: Summary and Discussion and Analysis*, 29 October 1987

12. Commander, Operational Test and Evaluation Force, *Initial Operational Test and Evaluation (OT-IIA) (Operational Assessment) of the Vertical Takeoff and Landing Tactical Unmanned Aerial Vehicle (VTUAV) Program*, 1 April 2001

13. IDA Paper P-3207, *"Predator" Medium Altitude Endurance (MAE) Unmanned Aerial Vehicle (UAV) Early Operational Assessment*, 9 October 1996

14. Air Force Operational Test and Evaluation Command, *Early Operational Assessment for RQ-4 (Global Hawk) System*, October 2000

15. Director, Operational Test and Evaluation, Info Memo, *Predator Medium Altitude Unmanned Aerial Vehicle Operational Test and Evaluation Report*, 3 October 2001

16. Institute for Defense Analyses Memorandum, *Review of Army Outrider Military Utility Assessment (MUA)*, 31 July 1998

APPENDIX A

ACRONYMS

APPENDIX A
ACRONYMS

ACTD	Advanced Concept Technology Demonstration
AFOTEC	Air Force Operational Test and Evaluation Command
AO	Action Officer
ATO	Air Tasking Order
AV	Air Vehicle
AVGAS	Aviation Gas
AVT	Air Vehicle Transport
BDA	Bomb Damage Assessment
B-LRIP	Beyond Low Rate Initial Production
CAMS	Core Automated Maintenance System
CDL	Common Data Link
CEP	Circular Error Probable
CGS	Common Ground Station
COI	Critical Operational Issues
COMOPTEVFOR	Commander, Operational Test and Evaluation Force
CONOPS	Concept of Operations
CR-UAV	Close Range Unmanned Aerial Vehicle
CU	Cruise Missiles and Unmanned Aerial Vehicles
C2	Command and Control
C3I	Command, Control, Communications, and Intelligence
DA	Density Altitude
DARPA	Defense Advanced Research Programs Office
DIA	Defense Intelligence Agency
DME	Displacement, Movement, and Emplacement
DMMF	Division Mobile Maintenance Facility
DODD	Department of Defense Directive
DODI	Department of Defense Instruction
DOT&E	Director, Operational Test and Evaluation
DT	Developmental Test
EMI	Electro-magnetic Interference
EO	Electro-Optic
ETOS	Expected Time On-Station

EUCOM	United States European Command
FFE	Fire For Effect
GCS	Ground Control Station
GFAC	Ground Forward Air Controller
GPS	Global Positioning Satellite
HFE	Heavy Fuel Engine
HMMWV	High-Mobility Multipurpose Wheeled Vehicle
IES	Imagery Exploitation System
IFF	Identification, Friend or Foe
IOT	Initial Operational Test (Army)
IR	Infrared
IRs	Intelligence Requirements
ISR	Intelligence, Surveillance and Reconnaissance
JITC	Joint Interoperability Test Command
JRMET	Joint Reliability and Maintainability Evaluation Team
JROC	Joint Requirements Oversight Council
JROCM	Joint Requirements Oversight Council
JSTARS	Joint Surveillance and Target Acquisition Radar System
JTCG/AS	Joint Technical Coordinating Group on Aircraft Survivability
JTUAV	Joint Tactical Unmanned Aerial Vehicle
KPP	Key Performance Parameter
LIC	Low Intensity Conflict
LOS	Line-of-Sight
LRE	Launch and Recovery Element
LRIP	Low Rate Initial Production
LTIOV	Last Time Information of Value
LUT	Limited User Test
MAE	Medium Altitude Endurance
MAT	Minimum Analytical Threshold
MCE	Mission Control Element
MESL	Mission Essential Subsystems List
MIC	Mid-Intensity Conflict
MPI	Mean Point of Impact
MPO	Mission Payload Operator
MST	Mission Support Team
MTBMAF	Mean Time Between Mission Affecting Failures

MTBOMF	Mean Time Between Operational Mission Failures
MTBSMA	Mean Time Between Scheduled Maintenance Action
MTBUMA	Mean Time Between Unscheduled Maintenance Actions
MTTR	Mean Time To Repair
MUA	Military Utility Assessment
NAI	Named Area of Interest
NIIRS	National Imagery Interpretability Rating Scale
NIR	Near Infrared
OA	Operational Assessment
OCS	Observer/Controller
OMS/MP	Operational Mode Summary/Mission Profile (Army)
OPEVAL	Operational Evaluation (Navy)
OPFOR	Opposing Force
OPTEMPO	Operational Tempo
ORD	Operational Requirements Document
OT	Operational Test
OTA	Operational Test Agencies
OT&E	Operational Test and Evaluation
PEO	Program Executive Office (Navy)
PIR	Priority Intelligence Requirement
PM	Program Manager
Pr(R/D)	Probability of Recognition, given detection
RAM	Reliability, Availability, and Maintainability
RCS	Radar Cross Section
RGA	Recovery Guidance Aid
RPV	Remotely Piloted Vehicle
RST	Report Success Template
RSTA	Reconnaissance, Surveillance, and Target Acquisition
RTCA	Real Time Casualty Assessment
RVT	Remote Video Terminal
SAR	Side Apertures Radar
SCD	Systems Capability Demonstration
SEP	Spherical Error Probable
SOF	Special Operations Forces
SR-UAV	Short Range Unmanned Aerial Vehicle
TALS	TUAV Automatic Landing System
TCS	Tactical Control System
TEMP	Test and Evaluation Master Plan

TLE	Target Location Error
TOC	Tactical Operations Center
TOO	Targets of Opportunity
TTP	Tactics, Techniques and Procedures
TUAV	Tactical Unmanned Aerial Vehicle
UAV	Unmanned Aerial Vehicle
UCAR	Unmanned Common Automatic Recovery System
USACOM	US Atlantic Command
VMU-1	Marine Unmanned Vehicle Squadron 1
VMU-2	Marine Unmanned Vehicle Squadron 2
VTOL	Vertical Takeoff and Landing
VTUAV	Vertical Takeoff Unmanned Aerial Vehicle

APPENDIX B

REPORT SUCCESS TEMPLATES AND EVALUATING MISSION EFFECTIVENESS

APPENDIX B
SUMMARY OF UAV OPERATIONAL TESTING

A. INTRODUCTION

This appendix reviews the OT conducted on UAVs that were considered formal acquisition programs. Unlike ACTD prgrams, discussed in the next chapter, these UAV systems have undergone formal OT periods as they progressed towards full-rate production.

1. Limited User Test

A LUT is usually conducted early in the life of a formal acquisition program in order to reduce risk prior to entering IOT.[1] A LUT addresses limited operational issues and is used to accomplish the following objectives:

- Testing necessary to supplement DT before a decision to purchase long-lead items or Low Rate Initial Production (LRIP) release decision for IOT.
- Testing necessary to verify a fix to a problem discovered in IOT that must be verified before the production decision (for example, the problem is of such magnitude that verification of a fix cannot be deferred to follow-on testing).
- As needed to support NDI or modifications that may not require a dedicated phase of IOT before a production decision.
- A LUT will not be used to circumvent requirements for IOT before a production approval decision as prescribed by statute, and DOD directives.
- A LUT will not be used to piecemeal IOT through a series of limited objective tests.

By its very nature, a LUT allows for a very limited scope of evaluation. System maturity at the time of the LUT probably represents a fraction of the capability desired by the customer. In addition, operator training and proficiency may be at such a low level that overall system performance could be adversely affected. Normally, a LUT is used to assess the *potential* of a system to meet the users' requirements, so great care must be taken to ensure that a system's capabilities are not overstated based upon performance during a LUT.

While a LUT is not supposed to be the only OT conducted on a system prior to fielding, there have been programs where this has been the case. Specifically, the Hunter

---

1 LUT is an Army term; the Navy and Air Force label similar events "Operational Assessments."

TUAV system was fielded to, and deployed by, operational units without additional OT beyond a single LUT.

2. Initial Operational Test

Prior to proceeding to full-rate production, UAV systems are required to undergo a formal test and evaluation period.[2] The goal of an IOT is to determine the operational effectiveness and operational suitability of the production representative system. During this test, fully trained service members operate the system under operationally realistic conditions as envisioned by the concept of operations.

The data collected during the IOT is used by DOT&E to publish the B-LRIP report to be submitted to Congress. A favorable B-LRIP report normally forms the basis for Congressional funding of full-rate production and fielding of the system.

3. Advanced Concept Technology Demonstrations

Note that Predator was an ACTD that transitioned to a formal acquisition program. Thus, the Predator IOT is reviewed below while the ACTD portion of Predator is reviewed in the next chapter. Global Hawk, which started as an ACTD, has also transitioned into a formal acquisition program, but there has been no formal Global Hawk OT period. Although Global Hawk is omitted from this appendix, the Global Hawk ACTD is reviewed in the Appendix C.

B. AQUILA

The modern history of the Army's battlefield drone efforts began in 1973, when the Defense Advanced Research Projects Agency (DARPA) began a program called PRAIRIE, which tested a UAV with a TV camera and a laser target designator. PRAIRIE was able to target a truck and direct a laser-guided bomb onto it.

Since DARPA does not have a charter to build operational systems, it passed the concept on to the Army, which proceeded with the next phase of development. Ford Aerospace had implemented PRAIRIE, but the Army put the follow-on effort, named "Aquila," up for bid, and Lockheed was the low bidder (Figure B-1, Table B-1).

---

2 The Army refers to this as the "Initial Operational Test"; the Navy refers to this as the "Operational Evaluation."

Aquila was a remotely-piloted air vehicle (RPV) system designed to perform reconnaissance, target acquisition, artillery fire adjustment, and target designation for laser-guided munitions such as Copperhead artillery rounds and HELLFIRE missiles. The concept of operations required that the RPV penetrate enemy territory 20 to 30 kilometers away during which time the RPV might be acquired and engaged, or countered, by enemy systems, such as air defense units or enemy radio-frequency jammers (Ref. 11).

Figure B-1. Aquila Air Vehicle

The Artillery branch was the Aquila program proponent within the U.S. Army. As planned, RPV equipment and personnel were organized into batteries. A full battery comprised 13 AV, battery headquarters section, two Central Launch and Recovery Sections, three Forward Control Sections, one support maintenance section, and one RPV maintenance section. An Aquila battery was to be a Corps asset, usually assigned to support a division.

The Aquila AV was a tailless aircraft, driven by a 24-horsepower piston engine with a pusher propeller. A truck mounted, hydraulic launch subsystem catapulted the AV into the air. A near infrared (NIR) source was mounted on the nose of the AV and was used in conjunction with the Recovery Guidance Aid (RGA) to facilitate auto recovery. Using the NIR source and the RGA, the AV is automatically guided to a truck-mounted, vertical-net recovery subsystem. A backup parachute recovery system was provided for use in emergency recovery.

During flight, the AV was capable of operating in six guidance modes and a jinking mode. Each guidance mode, except recovery, can be augmented with the jinking mode. These modes were manual, waypoint, circle, racetrack, figure eight, and final approach. The attitude reference assembly contained two gyros and three accelerometers.

When integrated with the flight control electronics package, the system provides the ability to compute AV velocity, geographic position, and AV attitude.[3]

Table B-1. Aquila AV Characteristics

	System Composition	
Air Vehicle	Wingspan	12.6 feet
	Length	6.4 feet
	Max Takeoff Weight	258 pounds
	Empty Weight	171 lbs
	Engines(s)	One 26hp pusher
	Fuel Capacity	28 pounds (gasoline)
Performance Characteristics	Flight Duration	3 hours
	Max operating range	27 nm radius
	Max Speed	112 mph
	Stall Speed	52 mph
	Climb Speed	80 mph
	Service Ceiling	12,000 (mission altitude 4,900 feet)
Payload	Payload	Electro-Optical[a] Laser rangefinder/designator
	Data Link	Encrypted, LOS
	Payload Weight	59.5 pounds

a Development work on the Aquila program was terminated prior to successful development of an IR sensor.

The AV was designed for minimum Radar Cross Section (RCS) and IR signature. The design features incorporated into AV design include airframe shaping and metalization for minimum RCS. The air inlet ducts were also screened to minimize RCS. The engine IR signature was reduced by the design and placement of the air exhaust ducts with engine exhaust ducted through the propeller.

1. Operational Testing

The U. S. Army completed tests in 1986 and 1987 to assess Aquila's mission performance and survivability. The first was an OT at Ft. Hood, Texas, conducted

3 Note that Aquila development occurred prior to the widespread use of the GPS.

between November 1986 and March 1987. The second test was live firing against Aquila at White Sands Missile Range during May 1987 (Ref. 9).

The OT was conducted as a series of nine exercise periods, each of which lasted 3 to 7 days. During this test the First Cavalry Division acted as the supported division, issuing mission orders and firing the artillery. The division was represented by the division tactical command post, one brigade command post, and a division artillery command post, with one artillery battalion operations center and one firing battery. During each exercise period, the RPV battery was ordered by the division to occupy positions and conduct missions against a simulated threat array. Missions included a search of designated areas, route reconnaissance, and a search for cued targets (with the suspected location provided), conventional artillery adjust fire missions, and target designation for Copperhead artillery rounds. The target arrays, with up to approximately 50 vehicles, were scripted to position, moved, and hidden according to Soviet Doctrine.

The OT included:

- 105 air vehicle flights in 36 days of active field trial testing
- Approximately 4,400 presentation of targets under test control with 750 detections of those targets (along with 2,781 detections of potential targets not under test control)
- 100 conventional artillery live fire missions
- 20 live Copperhead firings (with 14 hits).

Live fire at the Aquila was conducted at White Sands Missile Range. It included the following:

- 2,600 rounds of antiaircraft artillery
- One surface-to-air IR missile firing in five trials.

2. Major Test Results

Approximately 16 percent of all test-controlled targets were detected, even though the air vehicle flew within 1 km of 60 percent of the remaining (i.e., undetected) targets. In areas actually searched by the RPV, 39 percent of the presented targets were found.[4] The RPV was able to locate targets to within a 50 meter accuracy approximately 10 percent of the time in its passive mode, that is, without the laser.

---

4 The search technique employed by the test unit was partially to blame for the poor detection rate. It was postulated that improved search techniques could increase the percentage of detected targets.

In calling for and adjusting conventional artillery, 17 of 100 engagements were for a first round fire for effect. The remaining 83 missions were for an adjustment round. During testing, 45 percent of the first adjustment rounds were within 50 meters of the target and 85 percent were within 100 meters. Of the 17 engagements of first volley for effect, two were declared "no test" due to range control issues. Of the remaining 15, seven (46 percent) failed to come within 50 meters of the target. However, the ability to adjust conventional artillery fire with Aquila exceeded that of ground-based observers using lasers ranging for fire adjustment based upon results of previous OTs. Overall, Aquila did not meet the user criterion of 85 percent of mean points of impact within 50 meters of the target.

Even with contractor involvement, repair times were significantly longer than required due to inadequate fault diagnostics, manuals, and procedures. At the time of testing, spare parts requirements had yet to be determined. It was estimated that sufficient spares would not be available until two years after system deployment.

Due to the high workload, four people were required to control the air vehicle instead of three, as envisioned by the user representatives. Many of the errors associated with AV operations were induced by factors related to system design.

During live fire survivability testing, Aquila was not hit by either the antiaircraft artillery or the IR missile.[5,6]

Analysts felt that test results overstated Aquila capabilities due to contractor maintenance, test range terrain, and no flights in adverse weather. During survivability testing, no radar-directed surface-to-air missile threats were employed.

3. Test Adequacy Issues

The 1986 test represented the first time the U.S. Army employed an RPV to perform and sustain all the functions of emplacement, launch schedule compliance, and target acquisition and engagement in support of a tactical force. Prior to this, there was no testing to develop tactics and techniques to optimize performance. The analysts were challenged to separate, when possible, employment techniques from system capabilities.

---

5 The lone IR missile firing was declared a "no test" due to a suspected malfunction in the missile.

6 For a detailed analysis of Aquila survivability see IDA Memorandum Report M-345 (Ref. 9).

a. Test Design

The area in which the test was conducted contained not only targets controlled and known to the test force but also uncontrolled targets, including military vehicles passing through or by the area. These targets of opportunity could be considered valid targets by the RPV battery. However, these latter targets of opportunity were not only distractions to the RPV but also proved confusing to those attempting to analyze the data. Of the 3,531 detections reported by RPV operators, only 21 percent (750) were of instrumented targets.

A full RPV battery was not employed during the test. Only one of two Central Launch and Recovery Sections and one of three Forward Control Sections participated. Heavy contractor involvement invalidated all suitability data.

b. Requirements Definition

The criteria used to measure system success for probability of detection and artillery missions were incomplete and confusing. In the case of detection probabilities, the two criteria statements (50 percent detections of moving target arrays, and 30 percent detections of stationary target arrays) present a major problem: neither the area nor the time of search were specified. Intuitively, one would expect a higher detection rate given a smaller search area and a longer search time. Further, stationary targets were either out in the open, under hasty camouflage or fully camouflaged, but the percent of targets to be detected in each situation was not specified.

The Army criterion that 85 percent of the MPI for Fire for Effect (FFE) fall within 50 meters of the target was very stringent compared to results obtained from fire support means prior to this test. For example, during the Battery Computer System Follow-on Evaluation, less than 10 percent of the FFE engagements resulted in a MPI of less than 50 meters. For the artillery mission, valuable information was missing from the provided data.

4. Program Termination

Despite the continued DT of the Aquila system, it was determined that the system was too big, too expensive, and required too many support vehicles, a major cost-driver in the system. A single battery contained 13 air vehicles and 300 tons of equipment that required seven C-5 transports to move.

Of additional concern was a Government Accounting Office (GAO) report (Ref. 1) that stated that of the 10 most important operational considerations, Aquila failed five: launch; detection, recognition and location; survivability; reliability and maintainability; and human engineering. Artillery adjustment and growth potential were listed as "minor problems." The GAO recommended that the Aquila not be cleared for production until all test deficiencies were completed.

Eventually, the cost of the program, both monetarily and politically, grew too large, and the development effort was concluded without ever entering its full-scale production phase.

C. PIONEER

In 1985, Israeli Aircraft Industries formed a new company to develop the next generation of remotely-piloted vehicles. This effort was primarily aimed at the U.S. Navy, who was soliciting proposals at the time. Pioneer, the system developed by this new company, was selected as an interim solution for the Navy in 1986.

Pioneer was procured to provide imagery intelligence for tactical commanders on land and at sea. It flew over 300 combat reconnaissance missions during Persian Gulf operations in 1990-1991, and has flown in contingency operations over Bosnia, Haiti, and Somalia since 1994.

Pioneer is a twin-tailboom monoplane with a pusher propeller (Figure B-2, Table B-2). A typical system consists of up to eight air vehicles, a Ground Control Station, a Tracking Control Unit, a Portable Control Station, four Remote Receiving Stations, Pneumatic or Rocket-Assisted Launchers, and Shipboard-Net or Land-Based Runway-Arrestment Recovery Systems.

For launch and recovery, Pioneer employs the Unmanned Common Automatic Recovery System (UCARS).[7] UCARS is designed to provide automatic take-off and landing in all kinds of conditions. It operates on both shipboard and land-based systems, during day or night, and in all weather conditions. During land-based operations, Pioneer can be launched from a wheeled take-off, from a pneumatic catapult or by a jet-assisted take-off booster rocket. Land-based recoveries are conducted via wheeled landings using a tail hook and arresting cables.

---

7 UCARS was only used for shipboard operations.

During sea-based operations, the AV is launched via jet assisted take-off booster rockets. Recovery aboard ship is accomplished by flying into a net attached to an energy-absorbing system. The recovery net can also be used for land-based recoveries.

Table B-2. Pioneer AV Characteristics

System Composition	5 AVs with payloads, 1 GCS, 1 PCS, 1-4 RRS	
Air Vehicle	Wingspan	16.9 feet
	Length	14 feet
	Max Takeoff Weight	450lbs
	Empty Weight	304 pounds
	Engines(s)	1 x Sachs&Fitchel SF350; 2 cylinder, 2-stroke, 26hp AVGAS (100 Octane)
	Fuel Capacity	11 gallons (AVGAS)
	Avionics	GPS navigation; Mode IIIC IFF
	Launch/Recovery	RATO, pneumatic rail, runway/flight deck w/net
Performance Characteristics	Flight Duration	6 hours
	Max operating range	100 nm radius
	Max Speed	110kts
	Stall Speed	52kts
	Loiter/Cruise Speed	65 kts
	Service Ceiling	15,000
Payload	Payload	EO (Tamam MKD-200) IR (Tamam MKD-400) Chemical detection[a] Mine detection[a] Radio Relay[a] COMINT[a]
	Data Link	C-Band/UHF
	Payload Weight	70lbs

a Demonstrated capability

1. Operational Testing

Pioneer skipped the development stage of the formal acquisition process. Instead, nine systems were acquired (eight AVs apiece) and immediately fielded aboard the USS *Iowa* (BB61) in November 1986. Shortly thereafter, Pioneer experienced numerous problems, which resulted in the loss of several vehicles, and recovery aboard ship proved to be problematic. Electro-magnetic interference (EMI) from the host ships systems, as well as nearby ships, were to blame for many of these failures. Ultimately, a $50 million research and development effort was required to upgrade the nine Pioneer systems to a "minimum essential capability." The cost of the original nine systems was $87.7 million, so the upgrade program cost over half as much as the original systems. Even cursory OT would probably have prevented most, if not all, of the early mishaps, saving the Navy a significant amount of money.

2. Test Adequacy Issues

While no OT was conducted prior to fielding, the early fielding efforts of the Pioneer system provide some valuable lessons. First and foremost, the UAV systems must undergo EMI testing prior to fielding. This is especially critical of sea-based systems that are required to operate in the complex electromagnetic spectrum around Naval ships.

Pioneer, originally procured as an interim system, has never met the objective requirements set forth at the time it was fielded. Even through extensive upgrades, the system is still lacking in many respects. While the system has performed well during real world operational deployments, many problems still exist.

Regardless of scope, OT could have been used to identify shortcomings in system reliability and maintainability, and would also have placed additional focus on the training of operators and maintainers.

The lack of OT led the Navy to costly and time-consuming trial-and-error tests while trying to adapt the system for shipboard use. Pioneer is a prime example of the difficulties encountered when attempting to field a nondevelopmental item to the operating forces. Planning for future UAV systems should program for these difficulties to ensure that sufficient OT is conducted prior to moving into a low-rate initial production stage.

Figure B-2. Pioneer Air Vehicle

D. HUNTER

Between 1988 and 1990, the Joint Requirements Oversight Council (JROC) validated Mission Need Statements for four categories of UAV capabilities: Close-Range, Short-Range, Medium-Range, and Endurance. The short-range and close-range UAVs were to provide near real-time imagery intelligence for Army, Marine Corps and Navy tactical commanders. A Joint Tactical UAV (JT UAV) program was defined as a single system that would comprise short-range (200 km), close-range (50 km), and marinized-air vehicles. The medium-range and endurance-air vehicles were not part of the JT UAV program.

The initial JT UAV system was the Hunter short-range UAV (SR-UAV) built by Israeli Aircraft Industries and TRW (Figure B-3, Table B-3). The Army awarded an LRIP contract for seven Hunter systems in 1993.

Hunter was developed to provide both ground and maritime forces with near-real-time imagery intelligence within a 125 km direct radius of action, extensible to 200 km by using another Hunter as an airborne relay. The mission of the Hunter was to provide a day and night reconnaissance, intelligence, surveillance, and target acquisition asset to Corps and MAGTF Commanders. This program originated in 1988.

A subsequent shipboard version for the U.S. Navy was proposed. Hunter was fielded to both the U.S. Army and the U.S. Marine Corps. Shipboard compatibility was demonstrated aboard the USS Essex in 1993, but the Navy version was later abandoned.

While there were differences in the proposed organization of UAV units between the two Services, these mostly involved specific pieces of equipment assigned to each unit. This discussion will focus on the organization and operational concept of the U.S. Army SR-UAV company.

Table B-3. Hunter AV Characteristics

Air Vehicle	Wingspan	29 feet
	Length	22.6 feet
	Max Take-off Weight	1,600 pounds
	Empty Weight	1,190 pounds
	Engines(s)	Two 45 horsepower engines (one pusher and one puller)
	Fuel Capacity	190 liters
Performance Characteristics	Flight Duration	8-10 hours
	Max operating range	125km 200km with aerial relay
	Max Speed	110kts
	Stall Speed	53kts
	Service Ceiling	16,000 feet
Payload	Payload	Day/Night Imagery Daylight Air data relay Chemical threat detection system[a] Laser designator (for Hellfire missiles)[a] VHF/UHF comm. Relay package[a]
	Data Link	C-band unencrypted (two uplinks, two downlinks)
	Payload Weight	200 lbs

a Capability demonstrated since program termination

The Hunter AV is launched and recovered much like a conventional aircraft utilizing a rolling take-off and landing to a hard surface. The AV was designed to operate from a paved or unpaved road with a minimum width of 14 meters and at least 300 meters in length at sea level. Detailed site preparation would likely be necessary unless the site is an airfield or other suitable location such as a highway. As a planned alternative, a zero-length, rocket-assisted take-off from an open area of 250 meters may be used to launch the AV (Ref. 5).

Unlike Aquila, with its automatic launch and recovery systems, Hunter utilized an external flight control box that was used by the external pilot during launches and recoveries. Under normal circumstances, the AV is returned to its launch site and recovered by the external pilot. An arresting hook, which extends down from the rear of the fuselage, is used to engage arresting cables strung across the landing strip to bring the

AV to a halt. A backup parachute recovery system is provided for use in emergency recovery. The air vehicle used GPS in order to compute AV geographic position.

Figure B-3. Hunter Air Vehicle

It was anticipated that one UAV–SR company would be integrated into the Airborne Exploitation Battalion of the Military Intelligence Brigade that supports each Corps and Echelon Above Corps. In contrast to Aquila, which had the Artillery branch as a proponent, the Military Intelligence branch was the Hunter proponent for the U.S. Army. The basic components of the Army UAV–SR company were:

- A company headquarters responsible for the command and control of the organization and its logistical support
- A mission and flight control element, which includes a mission planning station, GCS, and launch and recovery capability
- Ground data terminals that link the GCS with in-flight UAVs
- Aerial vehicles
- Modular mission payloads, that include day-only sensors, day/night sensors, and airborne data relays
- Remote video capability
- A ground support equipment element for supply and maintenance, capable of operating at two locations.

1. Limited User Test I

There was only one OT period conducted on the Hunter system. The Limited User Test – I (LUT I) was conducted at Fort Huachuca, Arizona, from 31 May through 3 July 1992 (Ref. 5). During each week, a total of three missions were usually scheduled, typically one on Monday and two on Wednesday. Tuesday and Thursday were generally set aside for system maintenance and crew rest. Friday was set aside for makeup flights as required.

The purpose of LUT I was to examine the potential operational effectiveness and suitability of the Hunter system. Initially, two candidate SR-UAV systems were to be tested during LUT I. One candidate, the McDonnell-Douglas/Development Sciences Corporation Sky Owl UAV system was not certified as ready to execute the requirements of LUT I, and the system was not allowed to participate in the test. The results of LUT I were used to support a LRIP decision to buy seven Hunter systems.

Targets sets employed during LUT I were developed to represent enemy activity typical of the MIC and LIC environments being considered. The first three weeks of missions were flown against MIC target sets, which were based on a Southwest Asia scenario. Each of the three weeks of the MIC environment represented different phases of the war. The different phases reflected varying assumptions concerning factors such as the use of camouflage, and the ratio of moving to stationary targets. Missions flown during the last two weeks were against LIC targets, and were similar to those that might be encountered during a conflict with guerilla, or insurgent forces.

The Hunter program contained an additional LUT and an IOT prior to a full-rate production decision originally scheduled for September 1994. However, the program was terminated prior to these tests being executed.

2. Major Test Results

Thirty-six air vehicles were launched during the course of LUT I; of the 28 scheduled launch times, in nine instances launches were delayed greater than 30 minutes.

During the MIC portion of the test, the Hunter system attempted 79 out of 118 planned tasks in 10 missions.[8] During the LIC phase of the test, the system attempted 66 out of 70 planned missions. Approximately one-third of the missions scheduled that required a radio relay (to meet maximum-range operations) resulted in a failure. Aerial relay was cited as one of the most critical issues to address during the LUT.

The data indicated that the Hunter UAV has a 92 percent probability of detecting and recognizing *target sets* in a MIC scenario given good cueing (less than 300 meters

8 These attempts represent the frequency with which the system could arrive in the vicinity of the target, but not whether the system was successful or unsuccessful in accomplishing the assigned mission.

from cued location to target location) and no threat. Detection of individual target elements was much more difficult. Of the 737 elements presented to the UAV, 44 percent (325) were detected. The most difficult conditions involved camouflaged targets, both day and night, where typically only 30 to 40 percent of the target elements were detected. For target elements that were detected, the recognition rate was about 80 percent (260 out of 325).

The data for the LIC environment target coverage are more problematic to sort out. The detection rates observed are dominated by moving targets, during both the day and night, and hot targets at night. These categories account for 73 percent of the targets attempted during the LIC phase, during which 84 percent of the targets attempted were detected. In other categories, which accounted for the remaining 27 percent of target sets attempted, only 50 percent of the target sets were detected. Against cold, stationary targets at night, a target one would expect to encounter frequently in a LIC environment but that appeared only 22 percent of the time, a detection rate of only 4 out of 10 was observed.

3. Test Adequacy Issues

All supply and maintenance support was provided by the contractor; little useful information concerning operational availability and associated issues, such as the logistics supportability of the system, its maintainability by typical soldiers, and the accuracy of built-in test was available to support the assessment of these issues. A logistics demonstration conducted after the procurement of seven LRIP systems revealed that the system was not user-sustainable. Neither a survivability assessment, nor an interoperability assessment were conducted.

Because of the availability and maturity of the equipment evaluated, lack of a completely organized and trained UAV company to operate the equipment, and limited experience of the soldier operators, many of the evaluation issues could not be resolved adequately. Most importantly, issues pertaining to company operations – such as the effectiveness of the company, and its ability to support operations from dispersed sites; and to support sustained operations, unit level command and control, unit transportability, and unit displacements, moves, and emplacements – could not adequately be addressed.

In terms of test execution, the size and geometry of the test site did not match the employment concept of the SR-UAV system. In addition, there was perfect cueing for static targets, and lack of reaction to the realistic air threat portrayed during the LUT.

It must be stated that a LUT, by its very nature, is defined by a very limited scope. Therefore, the limited training and experience available to system operators may explain many of the results observed during this LUT. It was anticipated that many of these issues would have to be addressed in subsequent OT (LUT II and IOT), when many of the identified limitations would be eliminated. However, the Hunter system underwent no further OT prior to limited fielding to Army units. While it may be hypothesized that an experienced Hunter unit would perform better (in terms of target detection) than the LUT I unit, no data has been generated to support this contention.

4. System Termination

The Hunter UAV system was a major acquisition program with costs in the billions of dollars; however, the program was plagued by poor performance (including over 20 crashes) and delays, and in October 1995, the JROC recommended termination of the program (JROCM 126-95). Termination was recommended mainly on the grounds that sufficient funding would not be available for both the Hunter and the close range TUAV.

The first two acquired prototype systems, which included 16 air vehicles, were delivered in September 1994 to the U.S. Army's 304th Military Intelligence Battalion. A total of seven LRIP systems (62 air vehicles) had been delivered by September 1995, which have been used extensively for training exercises and in support of real world contingency operations.

E. SHADOW

The Close Range UAV (CR-UAV) variant was to provide near real-time imagery intelligence for lower-level ground force combat tactical units. The air vehicle was to have a 50-kilometer range to support artillery targeting. A series of air vehicle and payload demonstrations were carried out in 1992; six air vehicle contractors and three payload vendors conducted demonstrations at the Yuma Proving Ground and Redstone Arsenal. Additionally, a Cost and Operational Effectiveness Analysis, completed in 1994, substantiated the concept of mixing short- and close-range UAVs to meet the ground force requirements. The first CR-UAV ORD was signed and released on 12 September 1995.

Around October 1995, the JROC endorsed an ACTD approach to develop the CR-UAV (JROCM 125-95). Subsequently, JROCM 150-95 summarized performance and cost goals and directed the Army, Navy, and Marine Corps to assess the feasibility of a single UAV system to meet the Services' needs. The requirements of the 1995 CR-UAV ORD were incorporated into the JROC memo along with the requirements of the other participating services. These performance parameters were not stated as requirements, but as goals the contractor were to come "as close as possible" to achieving. In 1996, after a paper source selection process, a two-year, $52 million contract was awarded to Alliant Techsystems for the Outrider tactical UAV ACTD.

After the two-year ACTD, the Army concluded that the Outrider TUAV had the potential to meet its tactical UAV requirements while the Navy and Marine Corps concluded the Outrider did not meet Naval tactical UAV requirements. JROCM 109-98 then recommended that the Services pursue separate air vehicle solutions to meet their requirements as previously described in JROCM 150-95. The resulting Army TUAV ORD was validated by the JROC on 11 March 1999 (JROCM 030-99). This JROCM also encouraged the defense acquisition executive to pursue a path that obtains a 200-kilometer range objective and permits a single TUAV system to meet Army requirements.

On 12 March 1999, a Defense Acquisition Board decision designated the Army TUAV as an ACAT II program. In April 1999, the Army acquisition executive approved the TUAV competitive “best value” acquisition strategy. This acquisition strategy included a Systems Capabilities Demonstration (SCD), or fly-off, as input to a source selection evaluation board decision of a TUAV system. The SCD took place at Fort Huachuca between 4 October and 24 November 1999. During the SCD, four contractors demonstrated ground and flight operations of their off-the-shelf UAVs, of which Alliant Techsystem's Outrider was one of the competitors.

On 21 December 1999, the Army acquisition executive approved the Milestone II for the TUAV program, and on 27 December 1999, the Army awarded an LRIP contract to AAI Corporation for four Shadow 200 systems. The Military Intelligence branch of the U.S. Army was chosen as the Service proponent for TUAV.[9]

The Shadow 200 is a small, lightweight, tactical UAV system. It will be employed as a ground maneuver commander's primary day/night reconnaissance, surveillance, target acquisition, and battle damage assessment system. The TUAV is intended to provide the commander with a number of benefits:

---

9 This may change in the near future as the Aviation branch is to assume proponency for the TUAV within the U.S. Army.

- Enhanced situational awareness
- Target acquisition capability
- Battle damage assessment
- Enhanced battle management capabilities such as friendly situation and battlefield visualization.

The TUAV system consists of four main components: the GCS, air vehicles, modular mission payloads, and communications equipment. The system has two GCSs, two ground data terminals, one portable ground control station, one portable ground data terminal, and four remote video terminals.

The system can carry enough supplies and spares for an initial 72 hours of operations, and is transportable in two high-mobility multipurpose-wheeled vehicles (HMMWVs) with shelters, with one additional HMMWV with trailer as a personnel and equipment carrier, and one HMMWV as the AV transport (AVT) with the launcher on a trailer. The maintenance section consists of one HMMWV with a shelter and trail; which transports spares and provides maintenance support to the flight platoon.

A single TUAV system includes three Shadow 200 air vehicles with a fourth AV as part of the issued equipment of the maintenance section (Figure B-4, Table B-4). The AV is launched using a pneumatic launcher, mounted on a trailer that is pulled by the AV transport HMMWV. Required system capabilities include launch and recovery of the AV on an unprepared surface normally available in the brigade area of operations. The AV uses the tactical automatic landing system for recovery, without pilot intervention, and is stopped with an arresting hook and cable system.

In case of an emergency, the AV can be recovered in a small area by deploying the parachute at low altitude. When deployed, the parachute flips the AV upside down to prevent damage to the payload, but the parachute is not sufficient to prevent serious damage to the AV.

1. Operational Testing

The initial operational test and evaluation (IOT) began in April 2001 – only 15 months after the original contract award. Reliability problems surfaced during the first days of the test and ultimately four air vehicle crashes occurred in 35.5 flight hours. This IOT was downgraded to a LUT. All flight operations of the Shadow 200 were halted until an external review of the program was completed. The review found no systemic problems with the system and attributed many of the problems to poor quality control during manufacturing. Many fixes were implemented and one year later a new IOT was conducted using the original test plan.

Table B-4. Shadow AV Characteristics

System Composition	4 AVs with payloads, 2 GCS/GDT, PGCS/PGDT, TALS, 2 transport HMMWVs, 4 RVTs (6 total HMMWVs, 4 shelters and 2 trailers)	
Air Vehicle	Wingspan	12.6 feet
	Length	11.3 feet
	Max Take-off Weight	330 pounds
	Empty Weight	261lbs
	Engines(s)	One 38 hp pusher
	Fuel Capacity	35 liters (MOGAS)
	Avionics	GPS; BAE systems Navigation Sensor Unit[a]
	Launch and Recovery	Pneumatic rail, runway/autonomous runway recovery with arresting cable and tail hook
Performance Characteristics	Flight Duration	5+ hours
	Max operating range	75nm
	Max Speed	97 kts
	Loiter Speed	65-70 kts
	Service Ceiling	14,000ft
Payload	Payload	EO/IR
	Data Link	UHF, L-Band, C-Band
	Payload Weight	60lbs

a The Navigation Sensing Unit combines GPS, inertial, and air data sensors to provide navigation solutions.

Figure B-4. Shadow 200 Air Vehicle and Launcher

The TUAV IOT was conducted from 23 April to 6 May 2002 at Fort Hood, Texas. The IOT was to be conducted in two phases, with each phase lasting five days (Ref. 4). A TUAV ground control station was integrated into the 1st Brigade, 4th Infantry Division (Mechanized) TOC. The TUAV launch and recovery elements were set up at a tactical airstrip within Fort Hood ranges. Phase I was conducted in accordance with the OMS/MP while Phase II was to be conducted in a free-play exercise environment; 227 flight hours and 170 on-station hours took place during the test.

The scenario used for this exercise employed a Kosovo-like peacekeeping environment. The tactical situation ranged from transition to conflict and support and stability operations to mid-intensity conflict. The target set consisted of approximately 30 tracked and 60-wheeled vehicles, and threat simulators to replicate surface-to-air missile sites.

2. Major Test Results

System operational effectiveness encompasses three phases of a UAV mission: planning, execution, and product. There was no quantitative data collected in the area of mission planning. However, testing revealed system capability problems in the areas of mission execution and mission product.

Mission execution included: emplacement, flight performance, survivability, and AV recovery. The system met its requirements for emplacement time and flight performance. The ORD requires that a single UAV have an endurance of 5 hours; 18 of the 53 flights (34 percent) during the IOT met that criterion, with the longest flight lasting 5.62 hours.

The system has a known limitation in the area of AV recovery. The ORD requires that the TUAV be able to launch and recover from an unprepared area the size of a soccer field. While the system did successfully demonstrate two landings on an unprepared surface, it was not as treacherous as the ORD requirements, and no data were collected on damage to the AV. The small number of landings and lack of data collection make it impossible to determine the long-term effects on the system of such operations. System limitations requiring an optimized landing site could negatively impact a maneuver commander's employment of this system.

AV susceptibility to detection was high; it was seen and heard within the effective range of many threat systems, and unsophisticated threats can easily detect and locate the AV and ground segment using electronic support measures. Also, electromagnetic environmental testing revealed significant vulnerabilities (Ref. 7).

Mission product evaluation included target location error, image quality, artillery adjustment, and interoperability. The payload exceeded its requirement for image quality during technical testing, but the system demonstrated deficiencies in the areas of target location accuracy and support to artillery adjustment.

Median target location errors were determined to be in excess of 200 meters (with a threshold requirement of 80 meters). The system was unable to perform artillery adjustment. Observed procedures to support accurate and timely artillery adjustment were inadequate for second round FFE missions prescribed by the ORD. The artillery adjustment test was the only data collected during IOT on the systems ability to support target acquisition; however, the large target location error demonstrated by this system makes it not effective for target acquisition missions.[10]

JITC did not fully certify the TUAV system in accordance with DoDD 4630.4 and DoDI 4630.8 because the testing was conducted using software versions that had not been fielded. A specified interface certification was granted for only the configurations used during the OT. Interoperability certification with the fielded software versions would be necessary for compatibility with the majority of the Army in case of contingency operations.

---

[10] Compare this to the Aquila program which was able to put 45 percent of the first adjustment rounds within 50 meters of the target.

During the IOT, the TUAV platoon provided 544 intelligence reports to the brigade. The success of a particular report in answering a task was evaluated with user-developed success templates, which were used to score a report for its timeliness, accuracy or completeness, and contribution. Using the template methodology, 284 of the 544 reports could be scored.[11] The overall success rate was 57 percent, which exceeded the Army's 50 percent requirement.

Another method used to evaluate operational effectiveness was to examine the percentage of successfully accomplished tasks. Of the 376 tasks examined, 133 (35 percent) were not answered because clouds obscured the ground. An additional 15 (four percent) were not answered because air defense assets were reported in the vicinity of the tasked location; 57 tasks (15 percent) were unsuccessful due to operator error. The success rate for the scored taskings was 34 percent (126 tasks).[12] These results need to be qualified due to the unrealistic threat portrayal and the unknown number of unanswered tasks. Consequently, observed performance levels for the TUAV reconnaissance and surveillance missions are only valid under fair weather conditions, in the absence of an air threat, and require accurate (<100 m) cueing.

The OMS/MP requires that the TUAV platoon be capable of providing on-station time for five consecutive days of 12, 18, 18, 18, and 8 hours. During the first 5 days of the IOT, the platoon attempted to fly according to the OMS/MP and demonstrated the capability on 4 out of 5 days. Also, the ability to complete the OMS/MP has been demonstrated in earlier testing.

Primary suitability measures examined in the IOT assessment were reliability, maintainability, and availability. Fifteen of the 19 system aborts observed during the IOT were associated with the air vehicle; of these, engine and flight control problems caused the majority of failures. The system did not meet its requirements for reliability or maintainability, but inherent redundancy in the system allowed for acceptable operational availability. Table B-5 summarizes TUAV suitability metrics.

---

11 Reports in which clouds obscured the tasked target area and reports for which there were uninstrumented targets were not scored.

12 45 tasks were partially successful, i.e., they were either timely or accurate and complete, and therefore, cannot be considered successful.

Table B-5. TUAV Suitability Metrics

Metric	Value	ORD Requirement
Mean Time Between Systems Aborts	12 hours	20 hours
Mean Time to Repair	0.5 hours	1.5 hours
Availability	0.85	0.94

3. Test Adequacy Issues

The scenario was not well matched with the combat capabilities of the test brigade. The test brigade was organized as a heavy force equipped with tank and infantry fighting vehicles. All of the targets injected into the brigade's area represented roadblock, criminal, or benign activities. The targets were not sufficiently threatening to force the brigade to act on information reported by the TUAV platoon. Also, the unrealistic threat portrayal included unthreatened access to targets and consistently accurate cueing. Consequently, the contribution of the TUAV to the brigade commander could not be evaluated.

Data collected did not include all of the missions and taskings assigned to the TUAV platoon. Tasks not completed by the TUAV platoon were not collected. Without this knowledge, it was impossible to completely evaluate the contribution of the TUAV platoon to the brigade commander's requirement for timely and accurate information. Neither were the on-station mission requirements for the TUAV collected. The scoring of subsystem failures depended on knowledge of the brigade commander's mission requirements.

Testers authorized the TUAV to fly over threat territory even though threat air defenses were able to detect the TUAV AVs. This unrestricted ability of the AVs to fly where desired eliminated the operational requirement for them to observe targets from realistic slant ranges, and improved their opportunity to loiter over targets. Reports submitted under these conditions are probably optimistic, as are artillery-targeting results.

The size of the test area at Fort Hood restricted TUAV flight operations to a 30-by-30 kilometer area rather than the 50 by 50 kilometer area prescribed in the ORD. The platoon compensated by loitering the AV after takeoff and prior to landing to simulate the travel time between the L/R site and the on-station location. The platoon did not compensate for travel time between two observation points while the AV was on-station.

The AV landed on prepared surfaces rather than the unimproved surfaces specified in the ORD. After completion of the IOT, the platoon demonstrated its capability by moving the launch and recovery section to an unimproved location and successfully landing an AV twice.

The wheeled and tracked vehicles used as targets did not provide observation of the variety of combat, combat support, and combat service support vehicles TUAV operators might expect to encounter. The simulation of brigade subordinate elements meant that no brigade equipment was located throughout the brigade's area. Consequently, the test provided no opportunity to either demonstrate that the TUAV system could recognize vehicles or discern friendly vehicles from enemy vehicles.

Actions by testers or test support personnel limited operational realism. As a result, the success achieved by the TUAV system in detecting threat vehicles was probably optimistic for the following reasons:

- Targets were located close to prominent features, usually a road or intersection.
- Threat personnel made no attempt to cover or conceal threat vehicles. They also made no attempt to decoy or deceive TUAV operators.
- Division white cell staff cued the brigade to targets with accurate precise grid locations (six-digit coordinates). Target locations should have reflected the accuracy of the sensor that initially detected the target.

F. FIRE SCOUT

In November 1998, the JROC directed the Navy and the Army to pursue separate AV solutions to satisfy their tactical UAV requirements, and the Navy submitted its operational requirement for a vertical takeoff and landing tactical UAV (VTUAV). The JROC subsequently validated the Navy's VTUAV ORD in January 1999, with the following Key Performance Parameters (KPPs): ability to conduct Vertical Takeoff and Landing (VTOL) operations from a land-based site and all air-capable ships; ability to maintain a steady state hover; automatic launch and recovery capability; 200-pound payload capability; deck-restraining capability; ability to transfer control of the AV from one ground-control station to another; and ability to use either JP-5 or JP-8 heavy fuel. An interoperability KPP was added to the ORD that gives nine interfaces with which the VTUAV must be interoperable.

The VTUAV system is required to provide a RSTA and communications relay capability in support of littoral operations for the Navy and Marine Corps. The purpose of the VTUAV system is to collect and pass on information utilizing an airborne sensor platform that will provide the commander with extended and enhanced battle space situational awareness. The VTUAV, which will incorporate an EO/IR/Laser designator payload, should deliver timely, accurate, and complete information about the Commander's area of interest in near real time.

The VTUAV system is based on the Fire Scout AV (Figure B-5, Table B-6), which is itself based on the Schweizer Aircraft Corporation Model 330 manned turbine helicopter. The most significant change from the manned version is the replacement of the cockpit with a redundant flight control system including actuators, avionics, and software to support unmanned flight and payload operations. An existing Allison Rolls Royce gas turbine engine powers the AV. The ground control element will use the Tactical Control System architecture to support system functionality and intelligence product dissemination to other C4I nodes.

1. Operational Testing

Commander, Operational Test and Evaluation Force (COMOPTEVFOR) and the Marine Corps OT&E Activity completed an OA, OT-IIA, during April 2001 (Ref. 12). OT-IIA consisted of analyses of limited flight data, manned air vehicle data, and the developer's proposal. Subject matter experts from Marine Unmanned Aerial Vehicle Squadrons ONE (VMU-1) and TWO (VMU-2), Navy Fleet Composite Squadron Six (VC-6), and engineers from Ryan Aeronautical Center and Naval Air Systems Command were consulted.

2. Major Test Results

COMOPTEVFOR's OA rated the VTUAV system as both potentially operationally effective and potentially operationally suitable. The primary risk to the program is its dependence on the tactical control system (TCS) for its ground control element; without it, the VTUAV cannot be effective.

Final recommendations from the OA were to continue program development; however, fielding was not recommended until the current concept of employment with respect to forward deployed or stateside basing is reviewed. Other findings of the OA include:

- The VTUAV requirement to "be capable of conducting VTOL operations at 4,000 feet density altitude (DA) from all air-capable ships, and be capable of conducting VTOL operations at 4,000 feet DA from an unprepared land-based site (threshold/KPP)" was deemed operationally restrictive

- The Fire Scout AV's inability to be delivered via vertical replenishment may unduly restrict the system's capability to operate from remote and austere locations
- The VTUAV system will likely exceed set-up and pack-up timelines by several fold
- Although the VTUAV will likely meet the 25 meter spherical error probable/target location error requirement, it is inadequate for precision weapons guidance.

Figure B-5. Fire Scout Air Vehicle

3. Test Adequacy Issues

Since the OT conducted to date consisted of an operational assessment, there were no test adequacy issues. However, the value of early involvement of operational testers cannot be overstated because it not only provides valuable feedback to the program manager, but also helps the OTA design future test periods. By identifying system capabilities and limitations early, the OTA is better able to focus testing on high risk system requirements and reduce test time and assets by limiting the amount of testing in low risk areas.

Table B-6. Fire Scout AV Characteristics

Air Vehicle	Rotor Diameter	27.5 feet
	Length	23 feet
	Max Takeoff Weight	2,550 pounds
	Empty Weight	1,457 pounds
	Engines(s)	One Allison Rolls Royce 250-C20W
	Fuel Capacity	793 lbs
Performance Characteristics	Flight Duration	3 hours @ 110nm; 6 hours total
	Max operating range	150 nm
	Max Speed	125 kts
	Cruise Speed	110 kts
	Service Ceiling	20,000 feet
Payload	Payload	IAI Tamam EO/IR/Laser designator/voice relay
	Data Link	L3 Comm TCDL
	Payload Weight	200 pounds

G. PREDATOR

The ground-controlled Predator UAV system is a theater-level asset intended to provide a cued and non-cued reconnaissance, surveillance, and target acquisition capability. It carries electro-optical, infrared, and synthetic aperture radar sensor payloads. A Predator system includes four AVs and a ground control segment.

Since the Predator started as an ACTD program, there are no formal requirements for the Predator system. The ACTD program was initiated in 1994 to fulfill an urgent need identified by the Joint Staff to provide "continuous all-weather coverage of world-wide targets" and intelligence information on mobile targets for the in-theater commander that the current national, theater, and tactical intelligence collection assets could not provide. The long dwell capability was intended to provide the theater commander with continuous 24-hour coverage of any area of interest.

The Predator air segment consists of four full-composite AVs powered by a turbo-charged Rotax 914 engine fueled by aviation gasoline (Figure B-6, Table B-7). It has retractable, steerable tricycle-landing gear, made primarily of composites. Weight growth, including a more powerful engine with higher fuel consumption, has resulted in a decrease in air vehicle endurance from the ACTD configuration; the current endurance is approximately 24 hours, while the ACTD AV demonstrated endurances of up to 40 hours (with no payload).

Launch and recovery occur via normal hard-surface runway operations, much like a manned aircraft. The AV can operate autonomously or under continuous manual control, although take-off and landing must be manually controlled.

The air vehicle can carry EO, IR, and SAR payloads simultaneously. The EO/IR sensors are in a gyro-stabilized platform capable of rotating for a 360-degree field of regard. The EO subsystem consists of two identical daylight video cameras; one provides a spotter lens, and the other has a continuous zoom lens. Both cameras provide color video. The IR subsystem has three fields of view available and a doubler for a total of six discrete fields of view. The EO and IR payloads were designed to provide imagery of Level 6 quality on the NIIRS at 15,000 feet slant range. NIIRS 6 corresponds to a ground-resolvable distance of between 40 and 75 centimeters (16 and 30 inches).

A de-icing system comprising ice detectors and glycol weeping wings (also called "wet wings") provides the capability to transit through moderate icing conditions at any time during the flight. The Predator is not certified by the Air Force to operate in instrument meteorological conditions under instrument flight rules. Even under visual meteorological conditions there are limitations that constrain launch and recovery under visual flight rules (Table B-8). For example, the Predator cannot be launched in adverse weather, including any visible moisture such as rain, snow, ice, frost, or fog. Crosswind limitations for takeoff and landing are 17 knots.

1. Operational Testing

The IOT, conducted from 16 to 26 October 2000, comprised three phases (Ref. 15). The first phase involved collecting and scoring reliability data on the production representative system during training flights for several months prior to the IOT flight operations phases. The second phase consisted of two dedicated sorties designed to examine specific mission areas, and the third phase covered 7 days of continuous operations. The test included one system comprising four AVs with EO/IR/SAR payloads, a ground control station, a Predator Primary Satellite Link, and 57 personnel (a planned deployed increment for a Predator squadron). Two sets of pre-production wet wings were provided.

Figure B-6. Predator AV

Table B-7. Predator AV Characteristics

Air Vehicle	Wingspan	48.7 feet
	Length	27 feet
	Max Takeoff Weight	2,250lbs
	Empty Weight	1,200lbs
	Engines(s)	Rotax 914; 100hp
	Fuel Capacity	660 lbs (AVGAS)
	Avionics	INS/GPS; Mode III, IV IFF
Performance Characteristics	Flight Duration	12 hours @ 400nm
	Max Speed	120 kts
	Service Ceiling	25,000 feet MSL
Payload	Payload	Electro-Optic/Infrared (EO/IR) Synthetic Aperture Radar (SAR) 2 x Hellfire Missiles Laser Rangefinder/Designator
	Data Link	C-Band (line-of-sight) Ku-Band (over-the-horizon communications)
	Payload Weight	450 pounds

Table B-8. Predator's Operating Guidelines

Minimums (Operational Flights)	800 feet ceiling, 2 mile visibility or 500 feet and 1 mile above minimum for deployed operating base (whichever is greater)
Minimums (Functional Check Flights)	2,500 feet ceiling and 3 mile visibility in daylight hours only required after scheduled and unscheduled maintenance
Cloud separation	Launch: do not launch with visible moisture present In flight: avoid visible moisture
Winds	17 knots crosswinds 30 knots total winds

The first two flight days of the IOT consisted of two sorties to demonstrate Predator's capability to perform strike support, search and rescue, and mortar adjustment. Both of these sorties were conducted on the Nellis Ranges and lasted about 8 hours each. A concept of operations for each of these missions was developed specifically for this test. Because of the small sample size and demonstration nature of these events, only the Predator's potential ability to contribute to these mission areas was evaluated. Data supporting sensor performance and reliability and maintainability were also collected during these flights.

Two strike support scenarios, unopposed by surface or air threats, were conducted. The first scenario consisted of fighters (F-16s and A-10s) on airborne alert being directed to a Predator-located target by a ground forward air controller (GFAC) situated in Predator's GCS. The GFAC communicated with the fighters through the Predator's onboard ARC-210 radio. The fighters dropped live ordnance on the targets and Predator provided battle damage assessment (BDA) video for analysis.

The second strike support scenario involved the Predator supporting mortar adjustment. The Silver Flag airbase defense unit at Nellis fired mortars at a target location provided by the Predator. A mortar spotter located in the GCS provided miss distances to the gunners who adjusted and fired more rounds, and the Predator provided BDA video for analysis.

Two combat search and rescue scenarios were conducted. In the first, a pilot, acting as a surviving aircrew member, was located on the ground. This survivor used standard signaling devices such as mirrors, smoke, and parachute panels to see how long it would take the Predator to find him and how far away these devices could be seen with the Predator's sensors. The Predator operators were given geo-coordinates and visual ground references to begin the search. The second scenario simulated a higher threat environment in which the pilot, in an evading status, hid and attempted to conceal his location. The supporting OA-10 pilots were also used to help locate the ground survivor and direct the Predator to his location.

The final phase of the OT&E began with a simulated deployment to a designated area at Indian Springs Air Force Auxiliary Field, Nevada. A deployment package for spare parts was used and a timed teardown and buildup exercise was conducted at the conclusion of the flying phase.

The 7-day, continuous operations phase focused on examining the ability of the system to maintain a continuous presence over the battlefield. An exercise Air Operations Center issued a daily air tasking order (ATO) and target collection deck to the Predator operations cell at Indian Springs. Imagery was ultimately provided to a simulated exploitation cell where imagery analysts assessed image quality using NIIRS and also provided a more subjective imagery adequacy evaluation. The C4I architecture used for the tasking and dissemination during the test represented a deployed situation in which there is access to local telephone landlines. It did not include satellite transmission of the EO/IR video from the GCS to other command and control nodes.

Supplemental evaluations and studies provided additional data for the evaluation of Predator's operational effectiveness and suitability, including an imagery degradation study, survivability modeling and flights, target location accuracy flights, and a JITC evaluation of interoperability, which are described below.

2. Major Test Results

The Predator ORD defines presence as the ability to maintain "continuous (with on-station relief) 24-hour intelligence coverage of any target in the operating area within the parameters outlined in the ORD." The ORD further clarifies that the Predator's "primary operating area is from the forward line of troops (100-150 NM from the Operating Base) to the rear of the enemy second echelon (i.e., up to 400 NM radius from the Operating Base)." The ORD defines an ETOS measure, (which is the fraction of tasked time an aircraft is on-station and mission capable) as 75 percent.

The resulting point estimate for ETOS during the IOT was 68 percent. Note that this ETOS estimate applies to those conditions observed during the test. In particular, it only applies to the distances flown during the test and the number of AVs available.[13]

Throughout the first two lead-in sorties, air-to-ground communications through Predator were operationally ineffective. The ARC-210 radio could not be used for communicating with a ground FAC, mortar spotter, or other personnel on the ground. Back-up communications links had to be used during the strike support missions.

The Predator carries four sensors: day TV spotter, day TV continuous zoom, IR, and SAR; only the day TV spotter camera demonstrated the capability to recognize targets at a 30,000 foot slant range with a probability of recognition of 0.69. The IR camera could detect (something versus nothing) targets, but could classify (tracked versus wheeled) only 21 percent of the time and recognize (T-72 versus M1A1) only 5 percent of the time. The day TV camera could detect targets, but not classify or recognize them. The sensor results are summarized in Table B-9. (Note that the threshold requirement does not give a threshold for the probability of search, detect, or recognize; the objective requirement is for a probability of recognition greater than 90 percent.) During darkness or inclement weather, the sensor suite can detect but cannot classify or recognize targets at 30,000 feet slant range.

The program office conceded prior to the test that the SAR would not be able to meet ORD classification and recognition requirements. Therefore, the SAR's capabilities to search, detect, and recognize at a 30,000-foot slant range was not assessed during the test.

Table B-9. Sensor Performance at 30,000 Feet Slant Range

Sensor	Probability of Classification (tracked/wheeled)	Probability of Recognition (T-72/M1A1)	Number of Targets
Day TV Spotter	1.00	0.69	13
Day TV Zoom	0.00	0.00	4
IR	0.21	0.05	19

13 Because of the limited range area, flying racetrack patterns for a specified period of time simulated ingress and egress from the target area. The first sortie simulated coverage of targets 5 hours away from the operating base. Subsequent sorties were scheduled to simulate targets 2 hours away.

During the IOT, 92 preplanned and 21 ad hoc targets were tasked during the 7-day continuous operations phase, for a total of 113 tasked targets. Of these, 58 targets were excluded because of weather and range restrictions. For the remaining 55 targets, 33 were covered and 22 remained uncovered. Target coverage rate was 60 percent when weather and range constraints are considered; however, when all targets are considered, only 29 percent of the targets tasked for the 7 days were imaged.

The ORD requires a minimum NIIRS rating of 5.5, which corresponds to a ground resolvable distance of about 1 meter. As an example, this allows one to determine whether a radar set is vehicle- or trailer-mounted. Table B-10 summarizes the NIIRS ratings per sensor.

The Predator imagery was also evaluated for mission adequacy by mission type. For missions where high resolution is an obvious necessity, such as BDA or reconnaissance missions where target recognition and classification are required, the Predator must fly closer to the target to provide adequate imagery. On the other hand, surveillance missions do not require as high a resolution level and adequate imagery can be obtained at longer ranges. For surveillance missions, the IR sensor provided adequate imagery for all ranges collected, out to 38,000 feet, and the day TV provided adequate imagery between 10,000 and 22,500 feet.

For site survey missions, the analysts rated imagery inadequate if the scene width was too small (as was the case when the spotter mode was used). The employment of wide fields of view provided by the IR and day TV cameras, showed that adequate site survey missions could be conducted.

Table B-10. NIIRS Ratings per Sensor at 30,000 Feet

Sensor	No. Images Rated	NIIRS (at 30,000 feet)	Slant Range for Recognition (feet)
Day TV	22	2.7	6,100
Spotter	34	6.2	42,500
IR	30	4.6	11,500
SAR	10	5.9	–*

* Not assessed during the IOT

A measure available for comparison with historical deployment results is mission reliability, which is defined as the probability of completing a launched sortie without aborting because of a mission-affecting failure. Mission reliability is a reflection of both the reliability of the system and the duration of the missions. For the IOT, there were 10 sorties with 3 air aborts, resulting in an observed mission reliability of 70 percent. The IOT mission reliability for the Predator compared to combat deployments is shown in Table B-11. Performance during IOT was consistent with that observed in previous deployments.

Table B-11. Predator IOT Mission Reliability Comparison With Operational Deployments

Event	Dates	Flight Hours	Air Aborts	Mean Time Between Air Abort (hrs)	Average Completed Sortie Length (hrs)	Mission Reliability Observed/ (Calculated)
IOT	10/00	116.5	3	38.8	19	0.70 (0.61)
Kosovo (Op. Allied Force)	3/99-6/99	1026	26	39.5	13.2	0.69 (0.72)
Bosnia (Op. Joint Endeavor)	3/96-12/97	2814	66	42.6	10.9	0.74 (0.77)
Bosnia (Op. Provide Promise)	7/95-10/95	731	24	30.5	10.6	0.65 (0.71)

An effort was made to expand the available data on the baseline system by including the Joint Reliability and Maintainability Evaluation Team (JRMET)-scored CAMS data for the IOT baseline system collected from May through October 2000, but this data suffers from numerous inconsistencies that prevent it from providing the operational context necessary to determine the number of mission-affecting failures. For example, data available for the IOT AVs indicated only a single code 3 incident during IOT when there were in fact four. CAMS debrief data only showed two air aborts for the entire 6-month period despite three air aborts occurring in 1 week of IOT alone, while CAMS maintenance data indicated 18 air aborts during 687.4 operating hours. This latter data gives a mean time between air abort of 38.2 hours, which is consistent with the IOT and deployment data. The internal inconsistencies of CAMS data made it unreliable and therefore posttest analyses relied on the observed failure rate in IOT to calculate reliability values.

3. Test Adequacy Issues

Because several issues regarding the conduct of IOT remained unresolved, in 1998, the DOT&E conditionally approved the Predator TEMP. Some of the issues ultimately resulted in a marginally adequate test. For example, the envisioned and TEMP-compliant IOT test scope was reduced by Air Force Operational Test Command (AFOTEC) 4 months prior to the scheduled start date based on the belief that deployed operations could provide additional information on effectiveness and suitability; however, only limited data from any of the deployments with the 11^{th} or 15^{th} RS became available to AFOTEC. Both planned deployment exercises and the theater-representative satellite dissemination were eliminated. DOT&E approved this test plan only after reviewing additional information on the supporting test events and the proposed flight schedule for the continuous operations phase.

Continuous operations were interrupted by both weather and range restrictions. Limited airspace and the necessity to deconflict with other flight operations on the Nellis ranges further restricted operations, offering little tactical uncertainty to Predator operators. Range size, even including the surrounding military operations areas, did not allow for operations and relief-on-station at ranges up to 400 nautical miles, as required in the concept of operations.

Although the wish to include more consecutive flights with the de-icing wet wings was incorporated in the test design, the wet wings were subsequently grounded and not flown at all during the test. The lack of wet wings, combined with other weather and range constraints, prevented coverage of targets in the northern range area for most of the test.

Poor weather grounded the system on three separate occasions, and the loss of range time for 24 hours on another occasion reduced the flight hours to nearly half of the original plan.

Assessments of the quality of the imagery product were limited to the NIIRS ratings; feedback from other imagery users, such as strike mission aircrew or planners, was not solicited. Most of the ranges and target areas were well known to the Predator crew who train in the same areas, making search and detection of targets relatively easy.

Assessment of the onboard UHF radio was limited to the first two dedicated sorties. Evaluation of the transponder for Mode IV identification friend-or-foe (IFF) was not accomplished because no airborne or ground command and control (C2) platforms capable of Mode IV interrogation were utilized during the test. A limited theater-representative C4I architecture was in place, thus limiting the evaluation of tasking, dissemination, timeliness, and image quality.

A backup GCS manned by contractors was allowed to monitor the AVs via a line-of-sight (LOS) data link during the relief-on-station procedure. This backup procedure necessitated that relief-on-station procedures occur very near the launch site. Because multiple air and ground aborts increased the number of relief attempts, this test limitation adversely impacted the test when it required the on-station AV to disrupt current tasks and return to the relief-on-station point more frequently than expected, reducing the amount of time spent over the target area prosecuting the assigned target list.

The number of AVs effectively available also changed during the test. Initially, the test began with two AVs outfitted with the de-icing wet wings and two with conventional wings. The wet wings, however, were grounded during the first OPTEMPO sortie. This became important since maintenance problems downed the two dry wing aircraft, leaving none available for 21 hours until they were repaired.

The disparity between the apparently successful combat operations of the Predator and the system that did not perform well in the IOT is largely attributable to the fact that the deployed system is tasked and operated well within known limitations such as ETOS, weather restrictions, expected threats, and expected accuracy and dissemination abilities. Additionally, the operators in the field have developed workarounds, which are somewhat effective but often cumbersome, for many system deficiencies.

APPENDIX C

SUMMARY OF UAV OPERATIONAL TESTING

APPENDIX C
SUMMARY OF UAV ADVANCED CONCEPT TECHNOLOGY DEMONSTRATIONS

A. INTRODUCTION

The ACTD process was developed by the DoD acquisition community in 1994 to facilitate the integration and demonstration of new military capabilities based on mature advanced technologies. The ACTD process was designed to allow the user to evaluate the military utility of a new technology before committing to a major acquisition effort, to develop concepts of operation, and to retain a low-cost residual operational capability. At the end of the ACTD period (30 months in the case of the Predator), the user will conduct a Military Utility Assessment (MUA) of the system. If the system is assessed to have military utility and there is a continuing requirement, the system transitions into a formal acquisition program; however, if only a few systems are needed, the ACTD assets can be modified or retrofitted and fielded in the limited quantity available. But if the system has little military utility, the program ends and no more are acquired.

Under the formal acquisition process, OT is a statutory requirement to proceed beyond LRIP, but no requirement exists for formal OT prior to an ACTD system entering low-rate production. Nevertheless, many ACTD systems that entered low-rate production have resulted in the procurment of substantial inventories of unsatisfactory weapons requiring costly modifications to achieve satisfactory performance for the end user (Ref. 1).

The primary lesson learned from UAV ACTD programs is that once an ACTD program has demonstrated “military utility” and is deemed successful, the asset should not be deployed until the post-ACTD preplanned product improvement program has been completed and the system is deemed supportable. While it is understandable that system with proven military utility would be in demand, immediately deploying assets upon ACTD acceptance makes it extremely difficult to correct or implement all identified equipment modifications and upgrades.

The user-sponsor plays a key role in any ACTD program since the goal of the ACTD is to demonstrate the military utility of the system. Two UAV programs that started as ACTDs, the Predator and the Global Hawk, successfully crossed over from an

ACTD to the formal acquisition process. Two other programs, the Dark Star and the Outrider, were unable to transition from ACTD status to a formal acquisition program.

B. PREDATOR

In July 1993, the Joint Staff identified "an urgent need for the capability of an Endurance Unmanned Aerial Vehicle system." In response to the Joint Staff request and at the direction of the Under Secretary of Defense for Acquisition, the Navy's Program Executive Office for Cruise Missiles and Unmanned Aerial Vehicles, PEO(CU), issued a 30-month contract beginning January 1994 for a Medium-Altitude Endurance (MAE) UAV. The contract called for a demonstration of the UAV system within 6 months and a fieldable prototype of that system comprising three AVs and one GCS within 12 months. Within 24 months, a total of 10 AVs and three GCSs were to be available for fielding and deployment. A working model of the MAE UAV was developed within 6 months, and in June 1994, the MAE UAV was designated as an ACTD.

The U.S. Atlantic Command (USACOM) was selected to be the operational user representative for the Predator ACTD.[1] USACOM's responsibilities included providing the force participants, equipment, and scenarios for exercises and deployments. Additionally, USACOM continuously updated the Concept of Operations document and assessed the overall military utility of the system at the end of the ACTD.

DOT&E suggested that the Service OTAs perform an independent OA of the Predator system to support the acquisition decision. In response, the Deputy Under Secretary of Defense for Advanced Technology asked the OTAs to characterize the system and not assess it, since there were no firm requirements to assess performance against. AFOTEC was the lead Service OTA and produced a report within 60 days, which was coordinated with the Army’s OT&E Command and COMOPTEVFOR. The AFOTEC report was eventually updated with data from the second European Command (EUCOM) deployment prior to the program review.

1. Operational Assessment

One exercise and two operational deployments formed the basis of the OA requested by DOT&E:

---

1 For a detailed description of the Predator system, refer to Appendix B.

- Roving Sands – Optic Cobra Joint Training Exercise – 15 April to 8 May 1995 at Fort Sumner, New Mexico. Three AVs equipped with EO/IR payloads and C-band line-of-sight data links, one GCS, and one Trojan Spirit II conducted 19 missions.
- Operation Nomad Vigil – 15 July through 25 October 1995 deployment to Gjader, Albania, in support of Operation Provide Promise. Four AVs, one GCS, and two Trojan Spirit IIs conducted 69 operational missions. The UHF and Ku-band satellite data links were used for the first time.
- Operation Nomad Endeavor – 17 March through 3 July 1996 deployment to Taszar, Hungary, in support of Operation Joint Endeavor. Three AVs conducted 63 operational missions.

Predator taskings during Roving Sands fell into two general categories: support to Special Operations Forces (SOF) and detection of opposing force (OPFOR) assets. Taskings in support of SOF consisted of helicopter landing zone reconnaissance and surveys of the Pecos River. OPFOR assets included tactical ballistic missile locations, activities, and other targets of opportunity, such as integrated air defense sites. The Predator received cued and ad hoc taskings and also performed area and route reconnaissance.

During Nomad Vigil, Predator missions included monitoring threats to friendly forces, establishing safe areas, and in particular, identifying violations of temporary exclusion zones. The Predator was tasked to provide information to assist in planning withdrawal and search and rescue operations through preplanned and ad hoc missions. A hostile air defense environment existed throughout Nomad Vigil. As a result, Predator was tasked to perform battle damage assessments in areas too risky for manned platforms.

Throughout the course of Nomad Vigil, the Predator was forced to operate from a remote airfield due to international airspace restrictions. This remote basing location and operational factors meant nearly half of each mission was spent in transit to and from the target area.

The nature of the Predator tasking during Nomad Endeavor was somewhat different. After the peace accord was signed, the hostile air defense system abated. Preplanned target requests often involved imaging of suspected mass grave sights at night to look for changes from the previous day. Ad hoc targets included diversions to monitor small groups of people or vehicles crossing the zone of separation. (Because it was feared that riots could break out as people tried to return to their homeland.)

The military utility assessment of the Predator UAV system focused on mission accomplishment, availability, and reliability.

2. Major Assessment Results

The mission accomplishment rate was defined as the percentage of tasked missions completed without an abort. As shown in Table C-1, it was significantly affected by the weather. At the time of the assessment, it was felt that the number of weather-related mission aborts would decrease as a result of planned upgrades; however, this has not been born out by subsequent employment of the system.

Table C-1. Summary of Mission Accomplishment Rates

Exercise/Deployment	Excluding Weather	Overall
Nomad Endeavor	63%	28%
Nomad Vigil	54%	38%
Roving Sands	69%	38%

Target coverage rates were used to determine the percentage of the tasked targets for which the collection objectives were satisfied. These rates are shown in Table C-2.

Table C-2. Target Coverage Rates

Exercise/Deployment	Excluding Weather	Overall
Nomad Endeavor	51%	22%
Nomad Vigil	22%	17%
Roving Sands	NA	32%

The mission availability of the Predator system during the assessment events was quite good; even though the unit was equipped with at least three AVs, and was typically tasked to fly only six missions per week, many of which were cancelled due to weather. The maintenance team was therefore not severely stressed. Table C-3 presents a summary of availability and reliability metrics from the military utility assessment.

Table C-3. Summary of Mission Availability and Mission Reliability

Exercise/Deployment	Mission Availability	MTBMAF (hr)	Avg. Mission Duration (hr)	Mission Reliability
Nomad Endeavor	98%	25	7.0	76%
Nomad Vigil	99%	29	10.6	65%
Roving Sands	96%	54	8.5	84%

During the assessment several reliability and maintainability metrics were evaluated. However, it must be remembered that these metrics were evaluated using contractor maintenance and, in some cases, incomplete data sets. Table C-4 presents point estimates for each of these metrics.

Table C-4. Predator Reliability and Maintainability Summary

	Roving Sands	Nomad Vigil	Nomad Endeavor
Oper. Flight Hours	162	731	468
MTBMAF (hr)	54	29	26
MTBUMA (hr)	5.8	6.8	3.6
MTTR (hr)	N/A	1.7	1.9
MTBSMA	20.3	10.6	15.9

MTBMAF = Mean Time Between Mission Affecting Failures
MTBUMA = Mean Time Between Unscheduled Maintenance Actions
MTTR = Mean Time To Repair
MTBSMA = Mean Time Between Scheduled Maintenance Action

It was determined that additional evaluation was required in several areas due to the lack of quantifiable data. In particular, system survivability, supportability, target geolocation accuracy, training, and manpower requirements need to be examined before additional systems were procured or fielded (Ref. 13).

3. MUA Adequacy Issues

Since the Predator started as an ACTD program, there were no approved operational requirements, which meant there was no "yardstick" by which to judge the Predator's performance. Additionally, the data used to evaluate the Predator performance was collected outside a formal OT period. While these operations provided excellent opportunities to collect qualitative and anecdotal information, problems were encountered collecting quantitative data.

Of significant note was the lack of ground truth data. While the system may have detected numerous target sets through the course of each mission, it was not possible to determine either the number of undetected targets or the accuracy of the submitted reports. These are key performance characteristics that must be explored throughout a programs test cycle.

Evolution of the Predator system through the course of the events that comprised the assessment impacted the analytical effort. Initially, the system was not equipped with the Ku-band data link and the SAR sensor, both of which have since become standard equipment. Additionally, the manner in which the system was employed also evolved throughout these events. Both factors had some effect on data collection and analyses.

With the substantial amount of contractor operations and maintenance involved, it was difficult to evaluate military personnel capabilities to operate and maintain the system. Finally, many desired system capabilities were not demonstrated because the system was never tasked to perform them during the course of the exercise and two operational deployments.

C. OUTRIDER

The Outrider ACTD grew out of the Joint Tactical UAV program. With the cancellation of the Hunter, the Outrider ACTD was initiated to determine if one UAV system was capable of meeting JROC requirements for both the short-range and tactical UAVs. The ACTD involved the development, production, and deployment of an unmanned AV system to provide reconnaissance, surveillance, and target acquisition information to Marine Corps, Army, and Navy units. Requirements for the system were set forth in JROC Memorandum 150-95 with Navy managing the ACTD (Ref. 1).

A paper source selection was conducted, and in May 1996 the Joint Program Office awarded a two-year ACTD contract to deliver six complete Outrider systems with spares by March 1998, but the program experienced many setbacks and delays. For example, the proposed heavy fuel engine (HFE) did not develop as expected, and a gasoline alternative to the HFE did not demonstrate the required performance. Weight increases in the AV to meet mission requirements led to insufficient wing and fuselage length. Eventually, the wing and fuselage were made longer, and a rotary gasoline engine was substituted.

DoD planned to acquire Outrider UAV (Table C-5, Figure C-1) in a common configuration to meet the joint requirements. Mission requirements for the Outrider UAV system included:

- Four hours on-station time at a range of 200km
- Use of GPS for navigation and target reporting
- Launch from unprepared ground strips and Navy ships
- EO/IR sensors to provide 24-hour surveillance
- Complete system transportable by two high mobility multiwheeled vehicles
- Transportable by a single C-130

Included in the Outrider program was one mobile maintenance facility for every three Outrider systems.

Table C-5. Outrider AV Characteristics

System Composition	4 AVs, 4 MMPs, 2 GCS, 1 RVT, Launch/Recovery GSE	
Air Vehicle	Wingspan	11.1 feet
	Length	9.9 feet
	Max Takeoff Weight	385 lbs
	Engines(s)	Pusher
	Fuel Capacity	8.5 gallons (Heavy fuel)
Performance Characteristics	Flight Duration	4.9 hours @ 200km
	Max operating range	124 miles
	Maximum Speed	110 kts
	Stall Speed	34 kts
	Service Ceiling	15,000 ft
Payload	Payload	Color CCD FLIR
	Data Link	C-Band

Figure C-1. Outrider Air Vehicle

1. Military Utility Assessment

Because of this prolonged development and testing, most of the scheduled demonstrations and exercises for the ACTD did not occur. Army testing scheduled to begin during May 1997 did not commence until 29 April 1998, and lasted until 30 June 1998. This was a combined test with personnel from the Army (15th MI Battalion) and Marine Corps (VMU-1) participating.

The MUA was to examine two primary measures of military utility. The first was to measure the ability to conduct reconnaissance and answer intelligence requirements in support of a Korean-based brigade commander's mission. The second measure was to demonstrate sufficient range and endurance to cover a brigade commander's area of interest. In addition to these measures, the MUA also examined system sustainability.

The Outrider ACTD was severely limited because of extensive redesign, development, and late systems deliveries. The user was not able to complete individual crew training and certification (Ref. 16).

2. MUA Results

During the ACTD, the Outrider showed military utility to conduct and answer the Brigade's intelligence requirements. The system also demonstrated the range and endurance to cover the brigade's area of interest. However, there was not enough information to assess flight operations in an operational environment and during displacement, movement, and emplacement operations.

Other significant observations and recommendations made by the test unit included:

- The use of AVGAS was seen as a problem[2]
- Safety issues were identified with engine start-up procedures and the engine noise level
- There was no planned external pilot capability for the Outrider unit and no backup for the auto-recovery system other than a parachute landing[3]
- Outrider coverage should extend to at least 70 to 80 kilometers with more on-station time available.

The test unit also recommended that an additional 300 to 400 flight hours were required to adequately evaluate the system. The recommendations further stated that adjust fire operations, operations with more mature software, and operations in terrain that is more mountainous than Fort Hood should be conducted.

3. MUA Adequacy Issues

This assessment was missing the following critical elements:

- No tactical movement or training with a maneuver unit in an operational environment
- Limited use of the auto take-off and landing capability
- No operations from unimproved surfaces
- No continuous operations for assessment of OPTEMPO capability
- No night operations.

The Outrider program repeated many of the mistakes made during previous UAV acquisition programs, most notably the Pioneer and Hunter systems. While started as an ACTD program, DoD awarded further low-rate initial production contracts without conducting OT or demonstrating that the system is user-supportable. Like these earlier systems, the time and effort required to integrate non-developmental items was severely underestimated.

Integrating components necessary to satisfy the naval requirements, such as interference shielding and stronger landing gear delayed the initial flight by 4 months. Due the limited time allotted to ACTDs (2 years maximum) the first flight delay resulted

---

2 The use of AVGAS was a major lesson learned from UAV operations during Desert Shield/Desert Storm. At that time, AVGAS was in short supply with a single source available within 200 miles of the front lines. [Ref 15].

3 The TUAV, which took the place of the Outrider, does not have an external pilot capability. This was not raised as an issue during the TUAV IOT final report.

in less time for users to assess the system military utility. As a result of these and other modifications, the weight of a fueled AV ballooned from a projected weight of 385 pounds at the time of contract award to an actual weight of 578 pounds. This weight increase created significant sea-based launch and recovery issues that were never resolved.

The Outrider data link was not compatible with the common data link (CDL) standard. The analog data link used by Outrider was in the same band widely used by European and Korean televisions stations. The analog data link, versus the digital CDL, left little room for future upgrades. While a digital CDL could be developed for Outrider, it was felt that the associated costs would exceed Outriders post-ACTD cost limit.

The ACTD demonstration schedule failed to adequately evaluate many key requirements. The ACTD did not evaluate supportability, a critical shortcoming of the Hunter system. The user performed only basic maintenance during operational demonstrations, and the contractor performed all other maintenance. There were no logistics demonstrations to show that the system is user-supportable without contractor assistance.

4. ACTD Results

During the Outrider ACTD, there was evidence that the land-based Outrider UAV provided some degree of military utility. However, a fully joint program could not be accomplished. Consequently, joint requirements were modified to permit the use of more than one type of AV to meet the needs of all of the Services. The Army effort resulted in the land-based Shadow 200 TUAV. The Navy effort has been focused on a VTUAV for use on ships with small landing areas and in urban areas ashore.

D. GLOBAL HAWK

The Global Hawk High Altitude, Long-Endurance UAV (Figure C-2, Table C-6) started as an ACTD designed to satisfy the Defense Airborne Reconnaissance Office goal of providing extended reconnaissance capability to the Joint Forces Commander. Originally managed by DARPA, Global Hawk has since transitioned to the U.S. Air Force.

Figure C-2. Global Hawk AV

Originally, DARPA awarded five contracts for a 6-month initial design effort Phase I of the program. The program was then slated to fund two of the initial contractors to move into Phase II, an advanced development phase, but due to budget cuts, only one contractor was selected in April 1995. In Phase II, Northrup Grumman built two advanced concept air vehicles and a GCS. The first flight, scheduled to occur prior to the end of 1997, did not take place until 28 February 1998.

Phase III, an operational evaluation phase, was to include eight AVs and two ground control stations. This was later reduced to three AVs in order to remain within funding constraints.

A Global Hawk system is composed of three elements, a set of AVs, a common ground segment, and a common support segment. The AV is optimized for reconnaissance and surveillance missions in low-to-moderate threat areas where range and endurance are paramount. The imaging range of the AV payload is 20 to 270 kilometers. The AV is designed for fully autonomous operations from start-up, to taxi, take-off, mission execution, and recovery. Fully automated take-offs and landings are conducted to paved runways.

Table C-6. Global Hawk AV Characteristics

Air Vehicle	Wingspan	116.2 feet
	Length	44.4 feet
	Max Takeoff Weight	25,600 pounds
	Empty Weight	9,200 pounds
	Engines(s)	Allison Rolls-Royce AE3007H 7,600 pound turbofan
	Fuel Capacity	14,190 pounds
Performance Characteristics	Flight Duration	24 hours @ 1,200 nautical miles
	Max operating range	Unlimited
	Max Speed	400 knots
	Service Ceiling	65,000 feet
Payload	Payload	SAR/MTI EO/IR
	Data Link	Ku-Band SATCOM
	Payload Weight	2,000 pounds

The common ground segment, providing command and control to the AV, is composed of the Launch and Recovery Element (LRE) and the Mission Control Element (MCE). The LRE contains a differential GPS required for ground operations and launch of the AV. The LRE does have a mission planning capability, making it fully redundant to the MCE.

The MCE provides AV mission control and image processing and dissemination capabilities. There are four operator workstations within the MCE, one each for mission planning, command and control, image quality control, and communications management. The design of the LRE and MCE allows for operations in geographically separate locations, so the MCE can be deployed with the supported commands primary exploitation site.

1. Military Utility Assessment

The MUA consisted of 11 Service and Joint exercises conducted between June 1999 and June 2000 (see Table C-7). Since this was an ACTD program, the Global Hawk MUA was conducted in a crawl, walk, and then run manner in order to reflect the level of the systems integration into the conduct of the exercise (Ref. 14).

During the crawl stage, Global Hawk demonstrated basic capabilities. During the walk stage, greater emphasis was placed on the OPTEMPO, collection plan, and sortie duration. The final run stage was designed to demonstrate Global Hawk's contribution to the conduct of the battle.

Table C-7. Global Hawk MUA Events

Stage	Exercise	Date	Flight Hours	Sorties (planned/actual)
Crawl	Roving Sands	19-36 Jun 1999	32.0	3/3
	Extended Range 1	15-28 Jul 1999	36.9	4/2
	ER2/JEFX/CAX	30-31 Aug 1999	25.0	1/1
	CAX 10-99	9-10 Sep 1999	18.8	1/1
Walk	Extended Range 3	4-8 Oct 1999	50.1	2/2
	Extended Range 4	19-25 Oct 1999	50.3	2/2
	JFTEX-W	9-21 Nov 1999	46.4	5/3
	CJTF-6	2-6 Dec 1999	31.9	3/2
Run	Deployment/JTF-6	20 Apr 2000	10.5	1/1
	Linked Seas	8-11 May 2000	42.1	3/2
	JTFEX	17-19 May 2000	37.2	3/2
	Re-deployment	19 Jun 2000	8.4	4/1
TOTAL			389.6	32/22

2. MUA Results

The Global Hawk MUA mission accomplishment rate is presented in Figure C-3. As shown, 40 percent of the tasked missions were successfully completed and provided imagery. The percentage of tasked images that arrived to the end user varied from 44 to 79 percent. OPTEMPO influenced this percentage; an increase in it resulted in a decreased percentage of delivered tasked images.

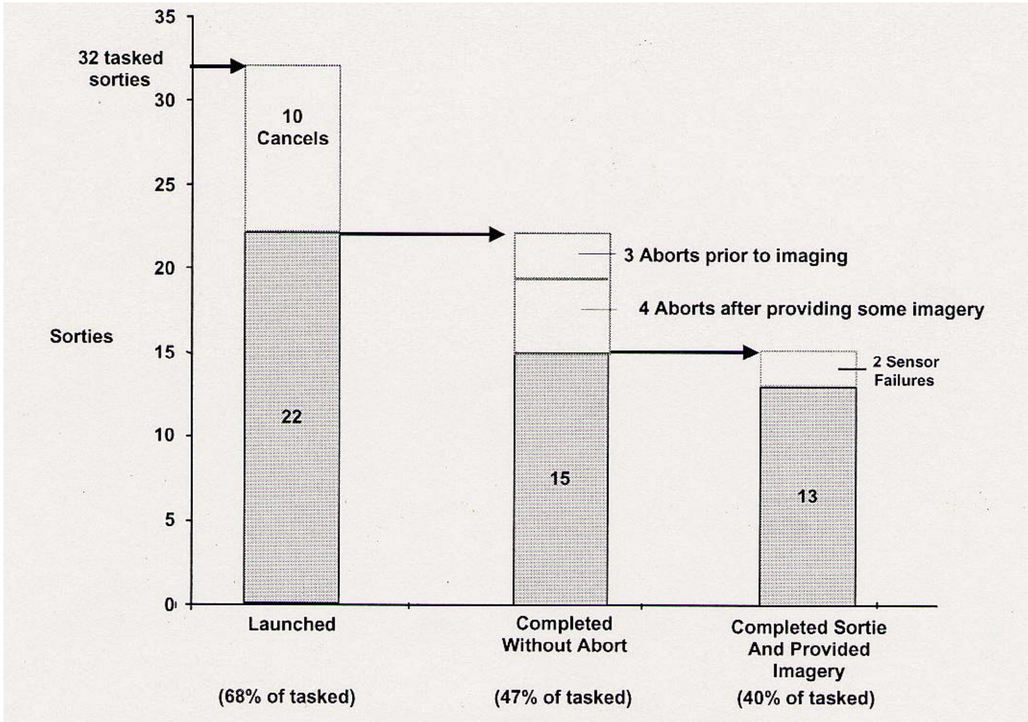

Figure C-3. Global Hawk Mission Accomplishment

The original goal of the ACTD was to collect 79 images per hour. During the ACTD exercise period, the average number of images per hour ranged from 1 to 13 images. Again, fewer images per hour were taken during the higher OPTEMPO exercises.

Another goal of the ACTD was to have a NIIRS rating of 5 (SAR) to 7 (spot mode). The average score for 1,448 images was 3.6, with 64 percent of the images scored between 3 and 4. Image quality was not degraded until the payload to target slant range exceeded 140 kilometers.

Timeliness of the Global Hawk system can be measured from the time the system begins to image a target until that image is disseminated to either the MCE or an Imagery Exploitation System (IES). The time needed to disseminate SAR imagery to an IES depends on the type of connection between the MCE and the IES. Several methods were used to make this link during the various exercises. A direct downlink provided the fastest connection, about three minutes. A 100Mbps-network connection between the MCE and the IES averaged about 15 minutes, although 10 percent of the scenes took more than one hour to complete. Dissemination to an IES via a remote connection (Global Broadcast System) averaged 4 hours.

The process required to plan a mission was cumbersome and error-prone. At the beginning of the ACTD, it took 11 weeks to plan a mission; by the end of the ACTD period, missions took only 50 days to plan. In fact, excessive mission planning time caused cancellation of two of four sorties in one exercise. The ORD requirement is for 12 hours. Associated with mission planning is the fact that re-tasking procedures were not well defined and had a detrimental effect on the entire collection plan.

The longest sortie during the ACTD period was 28 hours. Mission availability during this time was 72 percent, based upon 23 successful launches in 32 attempts. Of the 22 missions launched, seven were aborted in flight while two more experienced sensor failures. This equates to a mission reliability (the ability of the system to complete a mission once launched) of 59 percent. Mean Time Between Operational Mission Failures (MTBOMF) was 22 hours, based upon 375.3 flight hours and 17 operational mission failures.

Overall, the Global Hawk MUA declared that the system had military utility in its current configuration. However, the potential operational effectiveness of the Global Hawk was diminished by deficiencies in the areas of mission planning, imagery dissemination, and scene accountability. Furthermore, survivability and EO/IR performance was unknown.

The potential operational suitability was demonstrated with adequate mission availability and mission reliability for the level of tasking experienced during the scheduled exercises. However, long sortie duration and a more stressing OPTEMPO were not exercised. Other factors that could affect suitability were lack of a training plan, spare parts, a logistics infrastructure, a maintenance concept, and a designated main operating base.

3. MUA Adequacy Issues

There were several limitations to the MUA as planned and executed. First, due to two AV mishaps, participation in 4 of the 15 planned exercises was cancelled. The grounding of the AV after these mishaps also reduced the amount of time available for flight testing, thus the system did not mature as rapidly as expected.

The original assessment plan expected it to fly between 1,200 and 1,700 flight hours during 77 fly days. At the end of the one-year MUA period, the aircraft had only flown 54 sorties to accumulate approximately 673.2 flight hours. Of these sorties and hours, 22 sorties and 389.6 flight hours occurred during the course of a scheduled exercise. The small number of flight hours affected both the level of integration in joint and combined exercises (full integration was never achieved) and reduced the confidence in the observed suitability parameters.

During the MUA Global Hawk did not demonstrate a sustained OPTEMPO. The maximum OPTEMPO demonstrated during the MUA was four launches in a 12-day period. During this same period, two missions were cancelled due to incomplete mission plans, and two missions were aborted prior to mission completion.

Throughout the ACTD, system maintenance and operations depended upon contractor personnel. Additionally, operations relied heavily on contractor-provided equipment that is not part of the planned configuration.

At the time, the system was not cleared for operations in instrument flight rules conditions. Therefore, the ability to operate in adverse weather was not formally assessed.

Due to the previously mentioned mishaps, there were no EO/IR payloads available during the MUA. The scope of the MUA was further constrained by the functionality of the SAR/Ground Moving Target Indicator not being fully mature during the MUA. As a result, only a limited SAR sensor was assessed. At the completion of the MUA much still remained to be assessed with regards to the EO/IR sensors, including cueing, accuracy, and reliability.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
12/1/2003	Final	1986 - 2002

4. TITLE AND SUBTITLE	
Unmanned Aerial Vehicle Operational Test and Evaluation Lessons Learned	**5a. CONTRACT NUMBER** DASW01-04-C-0003
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Lee Carr, Kristen Lambrecht, Scott Shaw, Greg Whittier, Catherine Warner	**5e. TASK NUMBER** BD-9-2299
	5f. WORK UNIT NUMBER 229910
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 4850 Mark Center Drive Alexandria, Virginia 22311-1882	**8. PERFORMING ORGANIZATION REPORT NUMBER** P-3821
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Director, Operational Test and Evaluation (DOT&E) The Pentagon 1700 Defense Washington, DC 20301	**10. SPONSOR/MONITOR'S ACRONYM(S)** DOT&E
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.
Directorate for Freedom of Information and Security Review, 15 April 2004.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This paper highlight areas of UAV testing that have proven problematic from DOT&E's perspective in past UAV operational tests. Armed with this knowledge, DOT&E action officers should be better able to positively influence the scope and conduct of future UAV operational testing. Separate chapters focus on lessons learned during each stage of UAV operational testing to include requirements development, design of the operational test, test execution, and methodologies used to assess the mission effectiveness of UAVs. Two appendices at the end of the paper describe operational testing on UAVs conducted since 1986. UAV systems included in this review include Aquila, Pioneer, Hunter, Shadow, Outrider, Fire Scout, Predator, and Global Hawk.

15. SUBJECT TERMS
UAV, Unmanned Aerial Vehicles, Outrider, Predator, Hunter, Global Hawk, Shadow, Pioneer, Aquila, Fire Scout

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Mr. David Duma, DOT&E
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unlimited	142	**19b. TELEPHONE NUMBER** *(include area code)* (703) 697-4818