FOR OFFICIAL USE ONLY
UNTIL RELEASED BY THE
HOUSE ARMED SERVICES COMMITTEE
JOINT SEAPOWER AND EXPEDITIONARY
FORCES AND AIR AND LAND FORCES
SUBCOMMITTEES

TESTIMONY OF

DR. J. MICHAEL GILMORE

DIRECTOR, OPERATIONAL TEST AND EVALUATION

OFFICE OF THE SECRETARY OF DEFENSE

BEFORE THE UNITED STATES HOUSE
ARMED SERVICES COMMITTEE

JOINT SEAPOWER AND EXPEDITIONARY FORCES
AND AIR AND LAND FORCES
SUBCOMMITTEES

March 24, 2010

FOR OFFICIAL USE ONLY
UNTIL RELEASED BY THE
HOUSE ARMED SERVICES COMMITTEE
JOINT SEAPOWER AND EXPEDITIONARY
FORCES AND AIR AND LAND FORCES
SUBCOMMITTEES

Joint Strike Fighter
Dr. J. Michael Gilmore
Director, Operational Test and Evaluation (DOT&E)
Office of the Secretary of Defense

Good morning Chairman Taylor, Chairman Smith, Congressman Akin, Congressman Bartlett, and distinguished Members of the Subcommittees. I am here at your request to discuss the Joint Strike Fighter (JSF) program.

In my view, the primary issues in the Joint Strike Fighter program have been late delivery of test aircraft and the failure to adjust to that reality by building and resourcing realistic system development and test plans, as well as plans for producing and delivering aircraft. These problems have increased concurrency between testing and production beyond what was originally expected and beyond historical precedent. The resultant delays relative to unrealistic plans and the associated increase in costs to complete development created the need to restructure the program, which is in progress. In my FY 2009 Annual Report, I assessed that completion of Initial Operational Test and Evaluation (IOT&E) of the most capable combat capability now formally planned (the so-called Block 3 aircraft) could occur in early to mid-2016, provided certain changes are made to specific aspects of the program. Key changes needed include providing sufficient flight test aircraft, providing the resources and time needed to develop, deliver, and test effective software, accounting realistically for the inevitable discovery of problems during flight testing, and providing the engineering and other resources needed

to maintain an adequate pace of testing. I would like to review the status of these issues as I understand them today.

- Sufficient flight test aircraft. In the past fiscal year the program failed to meet the planned goals for testing, primarily due to the late delivery of test aircraft. As of today, three of the twelve previously planned flight test aircraft operate at one of the government test centers. Expectations at this time last year were that ten flight test aircraft would have begun productive flight test activity by now, with the final two following in the next 90 days. The program office now projects that all twelve of the previously planned developmental flight test aircraft will ferry to test centers by February, 2011. More test aircraft, generated from production lots, are needed to complete Block 3 development. I agree with the assessment of the Joint Estimating Team that two C-model aircraft, one A-model aircraft, and at least one B-model aircraft are needed in addition to the twelve previously planned developmental test aircraft to complete developmental testing in March 2015. Using production aircraft as developmental flight test assets, however, needs to be carefully managed to assure the original purposes for those aircraft, including operational test and evaluation, can still be met, either by returning the borrowed aircraft, or replacing them with other production aircraft.

- Software development and test. The delivery schedules for the remaining mission systems software Blocks 1, 2, and 3 have been extended by about one year compared to the plans existing in August 2008. The late delivery of test aircraft has, so far, masked the effect of delays in software development. Extending the use of the Cooperative Avionics Test Bed through the end of developmental testing was a good decision, as the test bed provides opportunities to discover problems with integration before software is loaded for use on an actual flight test aircraft. I understand the contractor has also proposed creating a new, additional software integration and test line. Although a lack of software integration and test resources was not identified previously as a problem by program management, the new test line will be very useful, provided the contractor has the manpower to operate it and simultaneously accomplish multiple integration activities. However, the reality is that flight test of the essential warfighting capabilities has yet to start. Mission systems flight test in F-35 aircraft begins when aircraft BF-4, the first of four previously planned mission systems test aircraft, ferries to a test center. This is currently planned to occur in May of this year. Only one of the remaining three previously planned mission systems test aircraft is expected to ferry before the end of 2010, with the final two delivering early in 2011. The initial mission systems testing will involve only very limited Block 1 capabilities. By

mid-2011, flight test is planned to transition from Block 1 to Block 2 capabilities—this will be an important point where the program will deal with the realities of the performance of software providing the first significant combat capability. Throughout this testing, the program needs to assure software is released to flight test only when it is ready. The JSF program must also prepare to cope with the many software and mission systems integration problems that will be discovered during flight testing, as has been the case with all complex software-intensive programs of this kind.

- Realistic Schedules and Sufficient Resources. The program's ability to maintain an adequate pace of testing is dependent on how the government and contractors manage several aspects of the planned strategy for verifying the performance of the aircraft using flight testing and modeling.
    - Integration of multiple test venues. The fundamental test strategy is to integrate multiple test venues, including contractor labs and models, as well as the Cooperative Avionics Test Bed, using F-35 flight test as a "capstone" event. Effective orchestration of these venues and this build-up process is critical to assure efficient use of flight test sorties. We have yet to see how the process being put in place will cope with multiple events for three different variants operating at two flight test centers. Ultimately, those responsible for

issuing, or rescinding, flight clearances will need time and resources to examine the data and, if necessary, request and receive additional information. This has already been the case with the Short Takeoff/Vertical Landing (STOVL) aircraft testing being conducted at Patuxent River. The contractor predicted early last December that the first vertical landing test event would be achieved by the end of 2009. This event was achieved March 18, 2010; in 2008 it was projected to occur in July 2009. Restrictions on flight associated with vertical landings will likely persist for some time due to discoveries made during recent flight testing. The decision cycle for understanding flight test results and achieving flight clearance will be under considerable pressure in the coming months, and will require continual supervision in order to meet test goals.

- Accreditation of models. The testing strategy puts a high premium on accreditation of the labs and models that the program plans to use in the build-up to flight test. As of November 2009, about 40 percent of the currently-planned model accreditation activities are planned to complete in 2013 or thereafter. While this gives the program time to incorporate performance data from flight tests in the accreditation process, it also highlights the limited margin available if the models and labs cannot play the intended role of limiting flight

test sorties to the amount currently planned. Flight test hours were reduced 26 percent in 2007 when modeling and simulation was assumed to be an acceptable substitute for flight testing. If those models cannot be accredited as planned, that assumption may not be realized, and flight test sorties will have to be added. I have recommended that an independent (of both the program office and contractor) review be conducted of the accreditation of the labs and models planned for use as test venues.

- Resources at the flight test centers. Adequacy of flight test center resources is also a concern. Adequate spare parts, trained personnel (including engineers), and training/mission rehearsal tools are essential to reaching the eventual pace of flight test events totaling over 140 per month. Early results in the area of spare parts usage and availability for the three aircraft now flying at Patuxent River are encouraging. However, managing adequate resources at two flight test centers after the remaining test aircraft are delivered requires considerable focus and early response as issues arise; again, this has been the case with other programs of this complexity. A high fidelity mission simulator located at or near the primary flight testing center, which I understand is being considered by the

program office, will also be key to sustaining an adequate pace of testing.

- Margin for discoveries. While it is difficult to determine definitively what level of additional schedule is needed to account for issues yet to be identified, it is important to acknowledge that this is the reality of testing—we will discover problems and need to make adjustments. Engine performance in ground tests, deficiencies in the flight control surface actuators, and slow progress towards the first vertical landing are examples that have already occurred. A more recent example is the need to modify the keel beam of C-model test aircraft and change the design of the keel beam used in production aircraft. I understand the program is now including short periods of down-time in its revised test plans to modify test aircraft to incorporate pre-planned configuration updates. Combined with planning for a realistic number of refly and regression test sorties, these constitute the margin for discovery. I note that, while the planned flight testing refly and regression rates were recently increased, they remain below historical experience.

In my FY2009 Annual Report, I also stated that the mission capability of Low Rate Initial Production aircraft is unclear. The program has identified the aircraft flight

envelope, as well as the functions that mission systems and the autonomic logistics systems should be capable of performing, that it hopes will be available when decisions are made to proceed with each lot of aircraft produced. However, it is not yet clear that the program's flight test plans provide reasonable assurance this performance will actually be demonstrated before each lot is delivered. Synchronization of testing with production deliveries is a key issue because during the past three years, delivery of production aircraft has been delayed up to nine months, while accomplishments in flight testing have been delayed two years. This synchronization is needed in order to plan for conducting all test and evaluation and will assist the Services as they make plans for fielding and supporting the early production aircraft. Planning in each case requires knowledge of the combat capability expected to be delivered to the government with each lot and variant of production aircraft.

I want to briefly highlight a system vulnerability issue included in my Annual Report. The program office is executing a comprehensive, robust, and fully funded Live Fire test plan. However, the program's removal of shutoff fuses for engine fueldraulics lines, coupled with the prior removal of dry bay fire extinguishers, has increased the likelihood of aircraft combat losses from ballistic threat induced fires. F-35 live fire testing to date has shown that threat impact into fuel tanks results in sustained fires. In addition, the F-35 will be more vulnerable to typical non-combat fires caused by fuel leaks and other system failures without the fire-suppression systems. At present, only the

Integrated Power Plant (IPP) bay has a fire suppression system. Though the configuration control process has approved the program office's request to remove these safety systems as an acceptable system trade to balance weight, cost, and risk, I remain concerned regarding the aircraft's vulnerability to threat-induced and safety-related fires.

In the remainder of this testimony, I will address the specific topics requested in my written invitation to speak at this hearing. I have already described what I consider to be the primary schedule and performance risks that existed in the JSF program and the steps taken by Dr. Carter to address those risks. The steps being taken to restructure the program reduce substantially --- but by no means eliminate --- the risk that key deficiencies in combat capability will be discovered during operational testing. The restructured program continues to have an unprecedented level of concurrency between production and testing. Along with the services, we recently received the program's latest version of its proposed new flight testing plans. In conjunction with developmental and operational test experts in the military Services, we will conduct an integrated test review to analyze the effects of the adjustments made in the revised plan relative to the program's prior plans, and assure those adjustments are consistent with the direction provided by Dr. Carter. Key issues we will examine include the following: the types of additional test aircraft to be used, as well as when and how long they will be used; whether the number of sorties to be flown by each test aircraft each month is consistent with historical experience and the availability of support assets (such as tankers) at flight

test centers; whether sufficient contractor and government flight test engineers will be available at the test centers to support the planned pace of testing; whether projected deliveries of flight test aircraft are consistent with reasonable extrapolation of experience-to-date in assembling and delivering aircraft to test centers; whether developmental testing of the mission data load (the data loaded in the aircraft's mission computers describing key characteristics of both friendly and enemy aircraft, air-to-air weapons, and surface-to-air weapons) will be adequate to assure readiness for IOT&E; whether the high-fidelity mission simulation needed to conduct IOT&E (and which would be very useful during developmental testing) will be robust and accredited; whether adequate time and resources are allocated to train operational test pilots; and whether sufficient fully production-representative aircraft will be available to conduct IOT&E.

My understanding of the current status of three programmatic risks contained in my FY 2009 Annual Report that are not discussed above is as follows:

- STOVL Testing. As of March 18, 2010, developmental test aircraft BF-1 had completed 23 flights at Patuxent River and had conducted the first vertical landing. During these flights the test team discovered damage to the linkages for the aircraft's auxiliary air inlet doors after unanticipated stresses experienced during side-slip. This resulted in the imposition of operating limitations on flight test and production aircraft until changes to the linkages

are made. The team also discovered greater than expected loads on the 3-bearing swivel duct doors, which allow the engine's nozzle to operate in STOVL mode. Subsequent testing revealed the doors have greater strength than predicted, and previously tested, permitting flight testing to continue. The test team continues to analyze the root cause of the increased loads. The contractor and test team are also working to determine how to instrument the STOVL aircraft's drive shaft (which connects the engine to the lift fan) to measure how it moves and deforms when the aircraft performs short take-offs and vertical landings. These measurements are needed to determine how the shaft must be re-designed to assure its durability. A new shaft is planned to be incorporated in production aircraft beginning with Lot 4. Additionally, the contractor and test team have determined that the STOVL clutch (which engages to transfer power from the engine to the lift fan through the drive shaft) heats excessively during conventional "up and away" flight. An overheated clutch could fail when engaged for vertical landing. Changes to both hardware and software will be needed to correct this problem; work began in February to define the needed changes. These discoveries highlight the challenges that will continue to be experienced in testing this complex set of aircraft.

- F135 engine fan trailing edge tip failure. The root cause of the failure experienced during previous testing of the F135 has been identified. The fan blade was incapable of withstanding the actual internal stresses experienced, which were greater than predicted. High-cycle fatigue testing of a new-design "clipped blade trailing edge" confirms the re-designed fan blade can withstand actual stresses. Initial service release engines are being retrofitted with the re-designed fan blade. Engines produced for use in aircraft after March 2010 will incorporate the re-designed fan blade.

- Inadequate Simulation Environment for Operational Testing. The verification simulation is a man-in-the-loop simulation planned for both developmental testing of integrated systems performance and operational testing of Block 2 and Block 3 capability. The contractor reduced the content of the simulation environment below that needed to conduct adequate operational testing—for example, numbers and types of both threat and friendly forces incorporated in the simulation are inadequate to support realistic operational testing. The operational test team has documented the specific changes needed. In response, the contractor and program office have begun to develop plans to improve the simulation. However, progress is very slow and I am concerned about readiness for testing. When accredited, this simulation, as was the case

with a similar asset in the F-22 program, will be a key venue for integrated testing.

C-model Keel Beam. The carrier variant design requires a modification to a section of its keel beam to withstand the stress of carrier catapult launches. I first became aware of this issue January 20, 2010, after submitting my annual report to Congress. I received a briefing describing the details of the problem on February 12, 2010. That briefing was presented by members of the program office and Naval Air Systems Command. The airframe design process is both concurrent and iterative; initial designs and models are updated as test data become available. During this process, an error was made in assigning data to the Finite Element Model of the aircraft's structure in the area of the aircraft's keel. When the correct data were used, the Finite Element Models indicated loads experienced in a particular part of the aircraft's keel beam during catapult would exceed design limits by 37 percent, indicating the keel would likely fail during catapult. This deficiency must be corrected by modifying the structure of the existing test aircraft before carrier trials can be conducted; it must also be corrected in ground test articles used for static structural and durability testing. The first C-model production aircraft will be delivered in Lot 4 and will use a new-design keel beam. The impact to the overall test program is still to be determined; however, at least one C-model test aircraft will be unavailable for up to four months as it undergoes modification and repair. Previous schedules for conducting static and durability ground testing of the C-model aircraft have been extended about six months; that testing is now expected to complete in

mid-2011. Effects on flight testing may be significant if more than one C-model test aircraft must be repaired to support carrier suitability testing. CF-3, the first mission systems carrier-variant aircraft, is the primary carrier trial aircraft and is scheduled to undergo the repair during its assembly before ferrying later this calendar year. Aircraft CF-1 is planned to receive the repair after first flight. CF-5, the developmental test aircraft added to production Lot 4, will receive the new design for the affected area of the keel. Flight sciences testing will begin on the C-model aircraft with delivery of CF-1 to Patuxent River later this calendar year (currently planned to occur in October 2010). Full scale static testing of the C-model is also currently scheduled to start in October 2010 and full scale durability testing in February 2012. Until the data from all these tests are analyzed and determined to be consistent with the predictions of the Finite Element Model, we will not know whether additional structural modifications to C-model aircraft will be necessary to achieve the combat capability currently required by the Navy.

Risk in the F-35 Program. The F-35 flight test plan originally spanned 65 months using 14 aircraft; it was reduced by the program office and the contractor in 2007 to a 52-month plan using 12 aircraft. The restructured plan directed by Dr. Carter allows 65 months to complete developmental flight testing and provides up to 16 aircraft to conduct that testing. Absent the additional four test aircraft, the Joint Estimating Team determined that completing developmental testing would require 82 months. The test plan schedule we just received from the program office for review allocates up to 65

months to complete developmental testing using up to four additional aircraft, but anticipates that testing might be completed more quickly---in about 57 months. The program office plan also appears to adopt more conservative assumptions than had been the case previously regarding the number of sorties each test aircraft will actually be able to fly and incorporates additional flight test hours. Thus, the revised plan appears to address many of the issues identified by the Joint Estimating Team. However, I still have the following concerns regarding the revised plan: availability of test aircraft for flight appears to be assumed to be about 80 percent, which is well above the 50 percent availability experienced in the F-22 program; the test aircraft being added are late-to-need and do not include sufficient numbers of C-models; although flight test hours have been added, the number of hours to be flown remains below the levels anticipated by the Joint Estimating Team; the plan for accrediting models does not explicitly incorporate independent review that I am recommending and remains aggressive and optimistic; and it remains unclear whether deficiencies in the verification simulation will be fully corrected.

B-Model and C-Model Shipboard Operations. The program office, in conjunction with the Services, tracks 47 individual issues for resolution towards successful B-model integration and STOVL operations from large-deck amphibious ships. Approximately one-half of these issues involve aircraft-ship integration, such as coping with main engine and integrated power pack exhaust. The remaining one-half deal with training and

manpower issues, such as security protocols and acceptance of training devices. To resolve these issues, several ship changes are being considered, as are potential modifications to production aircraft such as re-directing the Integrated Power Pack exhaust to limit effects on takeoff surfaces on the amphibious ships. The issues being reviewed do not appear to prohibit the ability to conduct initial flight testing aboard an L-class ship, which is now planned to begin March 2011. The Navy has also begun analyzing the issues associated with operating C-model JSFs from its aircraft carriers. This planning is not as mature as that for use of the B-model aboard amphibious ships. Challenges to complete C-model integration include: sufficiency of the jet-blast deflectors, effects of engine exhaust on surfaces and personnel, and hangar space requirements.

Engines. Testing of the F135 engine is slightly less than two years behind the schedule planned as recently as 2005. The F135 has encountered a number of technical problems. In late 2007 and early 2008, for example, failures occurring in ground testing of the F135 in STOVL modes revealed the need for design and manufacturing changes of turbine blades. Currently, the program is pursuing design changes in the drive shaft and clutch assembly of the STOVL propulsion system. For all variants, the program is dealing with a phenomenon known as afterburner screech, unstable pressures created in the engine's afterburner section that will limit flight testing at high Mach number and altitude until resolved. F135 ground test hours now total 13,223 and flight test hours total

over 200. Production of F135 engines is approximately one year behind the schedule planned in 2005.

The F136 "first engine to test" was delivered five months later than planned in the 2005 schedule. F136 ground test hours total 639. Three F136 engines in the most recent configuration are currently in test and have accumulated 135 hours of ground testing. Ground testing was recently restarted in January 2010 after testing in late 2009 revealed a failure in the combustor area, which necessitated modifications to a bearing and the combustor.

Increases in Aircraft Production. The Independent Manufacturing Review Team (IMRT) chartered by Dr. Carter provided a list of actions the contractor should undertake to prepare for award of the contract for production Lot 5. The IMRT recommended execution of a 20-item plan which, if followed, would greatly improve the quality and efficiency of aircraft production. The program office indicates it is executing this plan.

In conclusion, establishing realistic plans and adjusting to new realities revealed through flight test is essential as we move forward in the JSF program. Restructuring the test program and funding development consistent with the Joint Estimating Team's analysis are essential steps being taken now. In my view, the program needs to adjust continually to balance the pressure to complete testing on schedule and the need to

demonstrate that the combat performance needed by the Navy, Marines, and Air Force has been achieved. The demonstrated performance of the aircraft should have the greatest influence on the decisions and adjustments that need to be made as the program progresses.