United States General Accounting Office

GAO

Report to Congressional Requesters

February 2002

MISSILE DEFENSE

Review of Allegations about an Early National Missile Defense Flight Test

GAO-02-125

Contents

Tables

GAO
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

February 28, 2002

The Honorable Charles E. Grassley
United States Senate

The Honorable Howard L. Berman
House of Representatives

For a number of years, the Department of Defense has been researching and developing defenses against ballistic missile attacks on the United States, its deployed forces, friends, and allies. In 1990, the Department awarded research and development contracts to three contractors to develop and test exoatmospheric kill vehicles.[1] The Department planned to use the best of the three vehicles in a follow-on missile defense program. One of the contractors, Rockwell International, subcontracted a portion of its kill vehicle design work to TRW. TRW was tasked with developing software that could operate on a computer onboard the kill vehicle. The software was to analyze data collected in flight by the kill vehicle's sensor (which collects real-time information about threat objects), enabling the kill vehicle to distinguish an enemy warhead from accompanying decoys.[2]

The three contractors proceeded with development of the kill vehicle designs and built and tested key subsystems (such as the sensor) until 1994. In 1994, the Department of Defense eliminated Martin Marietta from the competition. Both Rockwell—portions of which in December 1996 became Boeing North American—and Hughes—now Raytheon—continued designing and testing their kill vehicles. In 1997 and 1998, the National Missile Defense Joint Program Office[3] conducted tests, in space, of the sensors being developed by the contractors for their competing kill vehicles. Boeing's sensor was tested in June 1997 (Integrated Flight Test 1A) and Raytheon's sensor was tested in January 1998 (Integrated Flight Test 2). Program officials said these tests were not meant to demonstrate

[1] An exoatmospheric kill vehicle is the part of a defensive missile that is designed to hit and destroy an incoming enemy warhead above the earth's atmosphere.

[2] In some instances, the system may also use ground radar data.

[3] The National Missile Defense Joint Program Office reports to the Ballistic Missile Defense Organization within the Department of Defense. The National Missile Defense program is now known as the Ground-based Midcourse Missile Defense Program and the Ballistic Missile Defense Organization is now the Missile Defense Agency.

that the sensor met performance requirements, nor were they intended to be the basis for any contract award decisions. Rather, they were early research and development tests that the program office considered experiments to primarily reduce risk in future flight tests. Specifically, the tests were designed to determine if the sensor could operate in space; to examine the extent to which the sensor could detect small differences in infrared emissions; to determine if the sensor was accurately calibrated; and to collect target signature[4] data for post-mission discrimination analysis.

After the two sensor tests, the program office planned another 19 flight tests from 1999 through 2005 in which the kill vehicle would attempt to intercept a mock warhead. Initially, Boeing's kill vehicle was scheduled for testing in Integrated Flight Test 3 and Raytheon's in Integrated Flight Test 4. However, Boeing became the Lead System Integrator for the National Missile Defense Program in April 1998 and, before the third flight test was conducted, selected Raytheon as the primary kill vehicle developer.[5]

Meanwhile, in September 1995, TRW had hired a senior staff engineer, Dr. Nira Schwartz, to work on various projects, including the company's effort to develop the exoatmospheric kill vehicle's discrimination software. The engineer helped evaluate some facets of a technology known as the Extended Kalman Filter Feature Extractor,[6] which TRW planned to add as an enhancement to its discrimination software. The engineer reported to TRW in February 1996 that tests revealed that the Filter could not extract the key characteristics, or features, from various target objects that an enemy missile might deploy and demanded that the company inform Rockwell and the Department of Defense. TRW fired the engineer in March 1996. In April 1996, the engineer filed a lawsuit under the False Claims Act[7] alleging that TRW[8] falsely reported or hid

[4] A target object's signature is the set of infrared signals emitted by the target.

[5] The Department of Defense continued funding the Boeing kill vehicle at a reduced level as a backup to Raytheon's kill vehicle. In mid-2000, the Department terminated all funding for Boeing's kill vehicle, ending TRW's involvement in development of the kill vehicle's discrimination software.

[6] The Kalman Filter is a mathematical model commonly used in real time data processing to estimate a variable of interest, such as an object's position or velocity. The Extended Kalman Filter Feature Extractor is used to extract features, which are used to perform discrimination.

[7] 31 USC 3729-3733.

information to make the National Missile Defense Joint Program Office believe that the Extended Kalman Filter Feature Extractor met the Department's technical requirements. The engineer has amended the lawsuit several times, including adding allegations that TRW misled the Department of Defense about the ability of its discrimination software to distinguish a warhead from decoys and that TRW's test reports on Integrated Flight Test 1A falsely represented the discrimination software's performance.

The False Claims Act allows a person to bring a lawsuit on behalf of the U.S. government if he or she has knowledge that a person or company has made a false or fraudulent claim against the government. If the suit is successful, the person bringing the lawsuit may share in any money recovered. The Department of Justice reviews all lawsuits filed under the act before deciding whether to join them. If it does, it becomes primarily responsible for prosecuting the case.

To determine whether it should join the engineer's lawsuit against TRW, Justice asked the Defense Criminal Investigative Service, a unit within the Department of Defense Inspector General's office,[9] to examine the allegations. The engineer cooperated with the Investigative Service for more than 2 years. During the course of the Department of Defense's investigation into the allegations of contractor fraud, two groups examined the former employee's specific allegations regarding the performance of TRW's basic discrimination software and performed limited evaluations of the Extended Kalman Filter Feature Extractor. The first was Nichols Research Corporation, a contractor providing technical assistance to the Ground Based Interceptor Project Management Office for its oversight of the exoatmospheric kill vehicle contracts. (This office within the National Missile Defense Joint Program Office is responsible for the exoatmospheric kill vehicle contracts.) Because an investigator for the Defense Criminal Investigative Service was concerned about the ability of Nichols to provide a truly objective assessment, the National Missile Defense Joint Program Office asked an existing advisory group, known as

[8] Rockwell, now Boeing North American, was later added to the lawsuit.

[9] Department of Justice officials told us that they often use other agencies' investigative units to investigate contractor fraud cases.

the Phase One Engineering Team,[10] to undertake another review of the specific allegations of fraud with respect to the software. This group is comprised of scientists from Federally Funded Research and Development Centers who were selected for the review team because of their knowledge of the National Missile Defense system. In addition, both Nichols and the Phase One Engineering Team assessed the feasibility of using the Extended Kalman Filter Feature Extractor to extract additional features from target objects that an enemy missile might deploy.[11]

The Department of Justice and the Defense Criminal Investigative Service investigated the engineer's allegations until March 1999. At that time, the Department of Justice decided not to intervene in the lawsuit. The engineer has continued to pursue her lawsuit without Justice's intervention.

Because you were the principal sponsors of the 1986 amendments to the False Claims Act and are interested in eliminating fraud in federal government programs, you asked that we review the former TRW employee's allegations. As agreed with your offices, we focused our review on Integrated Flight Test 1A. We also examined the evaluations of TRW's discrimination software and Extended Kalman Filter Feature Extractor conducted by Nichols Research Corporation and the Phase One Engineering Team. To answer your question regarding whether the Phase One Engineering Team's evaluation was objective and unbiased, we examined how the National Missile Defense Joint Program Office addressed the team's potential conflicts of interest. We also reviewed the basis for the Department of Justice's decision not to intervene in the

[10] The Phase One Engineering Team, according to its director, was established in 1988 by the Strategic Defense Initiative Organization—later known as the Ballistic Missile Defense Organization—as an umbrella mechanism to obtain technical and engineering support from Federally Funded Research and Development Centers. To ensure that the scientists who work on each review undertaken by the Phase One Engineering Team have the requisite expertise in the subjects they are asked to review, the membership on each review team varies with each assignment. The team assembled to review TRW's software included two individuals from the Massachusetts Institute of Technology's Lincoln Laboratory, two from Lawrence Livermore National Laboratory, and one from the Aerospace Corporation.

[11] In October 1996, TRW removed the Extended Kalman Filter Feature Extractor from its discrimination software. According to company officials, the Filter required computer speed and memory resources that were not available in the kill vehicle's onboard processor. In addition, the officials said that the basic discrimination software would perform adequately even without the Filter.

former TRW employee's lawsuit. Specifically, this report addresses the following questions:

1. Did Boeing and TRW disclose the key results and limitations of the flight test to the National Missile Defense Joint Program Office?

2. What were the methodology, findings, and limitations of the evaluations conducted by Nichols Research Corporation and the Phase One Engineering Team of TRW's discrimination software and the Extended Kalman Filter Feature Extractor?

3. What was the basis for the Department of Justice's decision not to join the lawsuit?

4. How did the National Missile Defense Joint Program Office assure itself that the Phase One Engineering Team could provide an independent and objective review?

Disclosure of Key Results and Limitations

Boeing and TRW disclosed the key results and limitations of Integrated Flight Test 1A in written reports released between August 13, 1997, and April 1, 1998. The contractors explained in a report issued 60 days after the June 1997 test that the test achieved its primary objectives, but that some sensor abnormalities were noted.[12] For example, while the report explained that the sensor detected the deployed targets and collected some usable target signals, the report also stated that some sensor components did not operate as desired and the sensor often detected targets where there were none. In December 1997, the contractors documented other test anomalies. According to briefing charts prepared for a December meeting, the Boeing sensor tested in Integrated Flight Test 1A had a low probability of detection; the sensor's software was not always confident that it had correctly identified some target objects; the software significantly increased the rank of one target object toward the end of the flight; and in-flight calibration of the sensor was inconsistent. Additionally, on April 1, 1998, the contractors submitted an addendum to an earlier report that noted two more problems. In this addendum, the contractors disclosed that their claim that TRW's software successfully distinguished a mock warhead from decoys during a post-flight analysis

[12] Appendix V includes selected requirements that Boeing established before the flight test to evaluate sensor performance and the actual sensor performance characteristics that Boeing and TRW discussed in the report.

was based on tests of the software using about one-third of the target signals collected during Integrated Flight Test 1A. The contractors also noted that TRW reduced the software's reference data[13] so that it would correspond to the collected target signals being analyzed. Project office and Nichols Research officials said that in late August 1997, the contractors orally communicated to them all problems and limitations that were subsequently described in the December 1997 briefing and the April 1998 addendum. However, neither project officials nor contractors could provide us with documentation of these communications.

Although the contractors reported the test's key results and limitations, they described the results using some terms that were not defined. For example, one written report characterized the test as a "success" and the sensor's performance as "excellent." We found that the information in the contractors' reports, in total, enabled officials in the Ground Based Interceptor Project Management Office and Nichols Research to understand the key results and limitations of the test. However, because such terms are qualitative and subjective rather than quantitative and objective, their use increased the likelihood that test results would be interpreted in different ways and might even be misunderstood. As part of our ongoing review of missile defense testing, we are examining the need for improvements in test reporting.

Appendix I provides details on the test and the information disclosed.

Evaluations of TRW's Discrimination Software

Two groups—Nichols Research Corporation and the Phase One Engineering Team— evaluated TRW's basic discrimination software. Nichols evaluated the software by testing it against simulated warheads and decoys similar to those that the contractors were directed to design their software to handle. The evaluation concluded that although the software had some weaknesses, it met performance requirements

[13] Reference data are a collection of predicted characteristics, or features, that target objects are expected to display during flight. The software identifies the warhead from the decoys by comparing the features displayed by the different target objects to the reference data.

established by Boeing in nearly all cases.[14] However, Nichols explained that the software was successful because the simulated threat was relatively simple. Nichols said that TRW's software was highly dependent on prior knowledge about the threat and that the test conditions that Nichols' engineers established for the evaluation included providing perfect knowledge of the features that the simulated warhead and decoys would display during the test.

Nichols' evaluation was limited because it did not test TRW's software using actual flight data from Integrated Flight Test 1A. Nichols told us that it had planned to assess the software's performance using real target signals collected during Integrated Flight Test 1A, but did not do so because resources were limited. Because it did not perform this assessment, Nichols cannot be said to have definitively proved or disproved TRW's claim that its software discriminated the mock warhead from decoys using data collected from Integrated Flight Test 1A.

The Phase One Engineering Team was tasked by the National Missile Defense Joint Program Office to assess the performance of TRW's software and to complete the assessment within 2 months using available data. The team's methodology included determining if TRW's software was based on sound mathematical, engineering, and scientific principles and testing the software's critical modules using data from Integrated Flight Test 1A.

The team reported that although the software had weaknesses, it was well designed and worked properly, with only some changes needed to increase the robustness of the discrimination function. Further, the team reported that the results of its test of the software using Integrated Flight Test 1A data produced essentially the same results as those reported by TRW. Based on its analysis, team members predicted that the software would perform successfully in a future intercept test if target objects deployed as expected.

[14] The Ground Based Interceptor Project Management Office identified the precision (expressed as a probability) with which the exoatmospheric kill vehicle is expected to destroy a warhead with a single shot. To ensure that the kill vehicle would meet this requirement, Boeing established lower-level requirements for each function that affects the kill vehicle's performance, including the discrimination function. Nichols compared the contractor-established software discrimination performance requirement to the software's performance in each simulated scenario.

Because the Phase One Engineering Team did not process the raw data from Integrated Flight Test 1A or develop its own reference data, the team cannot be said to have definitively proved or disproved TRW's claim that its software successfully distinguished the mock warhead from decoys using data collected from Integrated Flight Test 1A. A team member told us its use of Boeing- and TRW-provided data was appropriate because the former TRW employee had not alleged that the contractors tampered with the raw test data or used inappropriate reference data.

In assessing TRW's Extended Kalman Filter Feature Extractor, both Nichols and the Phase One Engineering Team tested whether the Filter could be used to extract an additional feature (key characteristic) from a target object's signal to help identify that object. Nichols tested the Filter's ability against a number of simulated target signals and found that it was generally successful. The Phase One Engineering Team tested the Filter's ability using the signals of one simulated target and one collected during Integrated Flight Test 1A. Both groups concluded that the Filter could feasibly provide additional information about target objects, but neither group's evaluation allowed it to forecast whether the Filter would improve the basic software's discrimination capability.

Appendix II provides additional details on the Nichols and Phase One Engineering Team evaluations.

Justice's Decision Not to Join Suit

The Department of Justice relied primarily on scientific reports, but considered information from two Army legal offices when it determined in March 1999 that it would not intervene in the false claims lawsuit brought by the former TRW employee. The scientific reports were prepared by Nichols Research Corporation and the Phase One Engineering Team. Justice's attorneys said they also considered an opinion of the Army Space and Missile Defense Command's legal office that said it did not consider vouchers submitted by Boeing for work performed by its subcontractor, TRW, as being false claims. In addition, the attorneys said a recommendation from the Army Legal Services Agency that Justice not intervene was a factor in their decision. It is not clear how the Army Legal Services Agency came to that decision as very little documentation is available and agency officials told us that they remember very little about the case.

Appendix III provides additional information on factors that were considered in Justice's decision.

Steps to Assure Independent and Objective Review by the Phase One Engineering Team

When the National Missile Defense Joint Program Office determined that another assessment of TRW's software should be undertaken, it tasked an existing advisory group, known as the Phase One Engineering Team, to conduct this review. Comprised of various Federally Funded Research and Development Centers, this group was established in 1988 by the Strategic Defense Initiative Organization as a mechanism to provide the program office with access to a continuous, independent, and objective source of technical and engineering expertise. Since the Federally Funded Research and Development Centers are authorized, established, and operated for the express purpose of providing the government with independent and objective advice, program officials determined that making use of this existing advisory group would be sufficient to assure an independent and objective review. Program officials said that they relied upon the centers' adherence to requirements contained in both the Federal Acquisition Regulation and their contracts and agreements with their sponsoring federal agencies to assure themselves that the review team could provide an independent, unbiased look at TRW's software.

Appendix IV provides a fuller explanation of the steps taken by the National Missile Defense Joint Program Office to assure itself that the Phase One Engineering Team would provide an independent and objective review.

Agency Comments and Our Evaluation

In commenting on a draft of this report, the Department of Defense and the Department of Justice concurred with our findings. The Department of Defense also suggested technical changes, which we incorporated as appropriate. The Department of Defense's comments are reprinted in appendix VII. The Department of Justice provided its concurrence via e-mail and had no additional comments.

We conducted our review from August 2000 through February 2002 in accordance with generally accepted government auditing standards. Appendix VI provides details on our scope and methodology. The National Missile Defense Joint Program Office's process for releasing documents significantly slowed our work. For example, the program office took approximately 4 months to release key documents, such as Nichols Research Corporation's 1996 and 1998 evaluations of the Extended Kalman Filter Feature Extractor and Nichols' 1997 evaluation of TRW's discrimination software. We requested these and other documents on September 14, 2000, and received them on January 9, 2001.

As arranged with your staff, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days from its issue date. At that time, we plan to provide copies of this report to the Chairmen and Ranking Minority Members of the Senate Committee on Armed Services; the Senate Committee on Appropriations, Subcommittee on Defense; the House Committee on Armed Services; and the House Committee on Appropriations, Subcommittee on Defense; and the Secretary of Defense; the Attorney General; and the Director, Missile Defense Agency. We will make copies available to others upon request.

If you or your staff have any questions concerning this report, please contact Bob Levin, Director, Acquisition and Sourcing Management, on (202) 512-4841; Jack Brock, Managing Director, on (202) 512-4841; or Keith Rhodes, Chief Technologist, on (202) 512-6412. Major contributors to this report are listed in appendix VIII.

Jack L. Brock, Jr.
Managing Director
Acquisition and Sourcing Management

Keith Rhodes
Chief Technologist
Applied Research and Methods

Appendix I: Disclosure of Flight Test's Key Results and Limitations

Boeing and TRW disclosed the key results and limitations of an early sensor flight test, known as Integrated Flight Test 1A, to the Ground Based Interceptor Project Management Office. The contractors included some key results and limitations in written reports submitted soon after the June 1997 test, but others were not included in written reports until December 1997 or April 1998. However, according to project office and Nichols officials, all problems and limitations included in the written reports were communicated orally to the project management office in late August 1997. The deputy project office manager said his office did not report these verbal communications to others within the Program Office or the Department of Defense because the project office was the office within the Department responsible for the Boeing contract.

One problem that was included in initial reports to program officials was a malfunctioning cooling mechanism that did not lower the sensor's temperature to the desired level. Boeing characterized the mechanism's performance as somewhat below expectations but functioning well enough for the sensor's operation. We hired experts to determine the extent to which the problem could affect the sensor's performance. The experts found that the cooling problem degraded the sensor's performance in a number of ways, but would not likely result in extreme performance degradation. The experts studied only how increased noise[1] affected the sensor's performance regarding comparative strengths of the target signals and the noise (signal to noise ratio). The experts did not evaluate discrimination performance, which is dependent on the measurement accuracy of the collected infrared signals. The experts' findings are discussed in more detail later in this appendix.

The Test

Integrated Flight Test 1A, conducted in June 1997, was a test of the Boeing sensor—a highly sensitive, compact, infrared device, consisting of an array of silicon detectors, that is normally mounted on the exoatmospheric kill vehicle. However, in this test, a surrogate launch vehicle carried the sensor above the earth's atmosphere to view a cluster of target objects that included a mock warhead and various decoys. When the sensor detected the target cluster, its silicon detectors began to make precise measurements of the infrared radiation emitted by the target objects. Over the tens of seconds that the target objects were within its field of view, the sensor continuously converted the infrared radiation into an electrical

[1] Noise is undesirable electronic energy from sources other than the target objects.

current, or signal, proportional to the amount of energy collected by the detectors. The sensor then digitized the signal (converted the signals into numerical values), completed a preliminary part of the planned signal processing, and formatted the signal so that it could be transmitted via a data link to a recorder on the ground. After the test, Boeing processed the signals further[2] and formatted them so that TRW could input the signals into its discrimination software to assess its capability to distinguish the mock warhead from decoys. In post-flight ground testing, the software analyzed the processed data and identified the key characteristics, or features, of each signal. The software then compared the features it extracted to the expected features of various types of target objects. Based on this comparison, the software ranked each item according to its likelihood of being the mock warhead. TRW reported that the highest-ranked object was the mock warhead.

The primary objective of Integrated Flight Test 1A was to reduce risk in future flight tests. Specifically, the test was designed to determine if the sensor could operate in space; to examine the extent to which the sensor could detect small differences in infrared emissions; to determine if the sensor was accurately calibrated; and to collect target signature[3] data for post-mission discrimination analysis. In addition, Boeing established quantitative requirements for the test.[4] For example, the sensor was expected to acquire the target objects at a specified distance. According to a Nichols' engineer, Boeing established these requirements to ensure that its exoatmospheric kill vehicle, when fully developed, could destroy a warhead with the single shot precision (expressed as a probability) required by the Ground Based Interceptor Project Management Office. The engineer said that in Integrated Flight Test 1A, Boeing planned to measure its sensor's performance against these lower-level requirements so that Boeing engineers could determine which sensor elements, including the software, required further refinement. However, the engineer told us that because of the various sensor problems, of which the contractor and project office were aware, Boeing determined before the test that it would not use most of these requirements to judge the sensor's performance. (Although Boeing did not judge the performance of its sensor against the

[2] The signal processing that Boeing completed after the test will be completed onboard the exoatmospheric kill vehicle in an operational system.

[3] A target object's signature is the set of infrared signals emitted by the target.

[4] These requirements were established by the contractor and were not imposed by the government.

requirements as it originally planned, Boeing did, in some cases, report the sensor's performance in terms of these requirements. For a summary of selected test requirements and the sensor's performance as reported by Boeing and TRW in their August 22, 1997, report, see app. V.)

Reported Key Results and Limitations

Table 1 provides details on the key results and limitations of Integrated Flight Test 1A that contractors disclosed in various written reports and briefing charts.

Table 1: What and When Key Results and Limitations Were Included in Contractors' Written Reports

August 13, 1997, Report	August 22, 1997, Report	December 11, 1997, Briefing	April 1, 1998, Report
Detected deployed targets	Detected deployed targets	High false alarm rate	Failure of gap-filling module[a]
Target signals collected	Target signals collected	Sensor did not cool to desired temperature	Target signals collected during selected portion of the flight timeline used in assessment of discrimination software
Discrimination software distinguished mock warhead from decoys	Discrimination software distinguished mock warhead from decoys	Software confidence factor remained small for two target objects	Selected reference data used in assessment of discrimination software
	Excellent performance of sensor payload	Sensor had a lower than expected probability of detection	
	Power supply caused noisy target signals	Software significantly increased rank of one target object toward the end of the flight	
	Sensor did not cool to desired temperature	In-flight calibration of sensor was inconsistent	
	High false alarm rate		
	Slow turn-around of launch vehicle caused data loss		

[a]TRW designed a gap-filling module for its discrimination software to replace missing or noisy portions of collected and simulated target signals.

Although the contractors disclosed the key results and limitations of the flight test in written reports and in discussions, the written reports described the results using some terms that were not defined. For example, in their August 22, 1997, report, Boeing and TRW described Integrated Flight Test 1A as a "success" and the performance of the Boeing sensor as "excellent." We asked the contractors to explain their use of these terms. We asked Boeing, for example, why it characterized its sensor's performance as "excellent" when the sensor's silicon detector array did not cool to the desired temperature, the sensor's power supply created excess noise, and the sensor detected numerous false targets. Boeing said that even though the silicon detector array operated at temperatures 20 to 30 percent higher than desired, the sensor produced

useful data. Officials said they knew of no other sensor that would be capable of producing any useful data under those conditions. Boeing officials went on to say that the sensor continuously produced usable, and, much of the time, excellent data in "real-time" during flight. In addition, officials said the sensor component responsible for suppressing background noise in the silicon detector array performed perfectly in space and the silicon detectors collected data in more than one wave band. Boeing concluded that the sensor's performance allowed the test to meet all mission objectives.

Based on our review of the reports and discussions with officials in the Ground Based Interceptor Project Management Office and Nichols Research, we found that the contractors' reports, in total, contained information for those officials to understand the key results and limitations of the test. However, because terms such as "success" and "excellent" are qualitative and subjective rather than quantitative and objective, we believe their use increases the likelihood that test results would be interpreted in different ways and could even be misunderstood. As part of our ongoing review of missile defense testing, we are examining the need for improvements in test reporting.

The August 13 Report

This report, sometimes referred to as the 45-day report, was a series of briefing charts. In it, contractors reported that Integrated Flight Test 1A achieved its principal objectives of reducing risks for subsequent flight tests, demonstrating the performance of the exoatmospheric kill vehicle's sensor, and collecting target signature data. In addition, the report stated that TRW's software successfully distinguished a mock warhead from accompanying decoys.[5]

The August 22 Report

The August 22 report, known as the 60-day report, was a lengthy document that disclosed much more than the August 13 report. As discussed in more detail below, the report explained that some sensor abnormalities were observed during the test, that some signals collected from the target

[5] Boeing and TRW reported that the original test objectives did not include a test of TRW's discrimination software. However, program officials decided immediately prior to the test that it offered an excellent opportunity to assess the software's capability even though post-processing tools needed to assess the software were not yet available and would need rapid development after Integrated Flight Test 1A.

objects were degraded, that the launch vehicle carrying the sensor into space adversely affected the sensor's ability to collect target signals, and that the sensor sometimes detected targets where there were none. These problems were all noted in the body of the report, but the report summary stated that review and analysis subsequent to the test confirmed the "excellent" performance and nominal operation of all sensor subsystems.

Some Sensor Abnormalities Were Observed During the Test

Boeing disclosed in the report that sensor abnormalities were observed during the test and that the sensor experienced a higher than expected false alarm rate. These abnormalities were (1) a cooling mechanism that did not bring the sensor's silicon detectors to the intended operating temperature, (2) a power supply unit[6] that created excess noise, and (3) software that did not function as designed because of the slow turnaround of the surrogate launch vehicle.

In the report's summary, Boeing characterized the cooling mechanism's performance as somewhat below expectations but functioning well enough for the sensor's operation. In the body of the report, Boeing said that the fluctuations in temperature could lead to an apparent decrease in sensor performance. Additionally, Boeing engineers told us that the cooling mechanism's failure to bring the silicon detector array to the required temperature caused the detectors to be noisy. Because the discrimination software identifies objects as a warhead or a decoy by comparing the features of a target's signal with those it expects a warhead or decoy to display, a noisy signal may confuse the software. Boeing and TRW engineers said that they and program office officials were aware that there was a problem with the sensor's cooling mechanism before the test was conducted. However, Boeing believed that the sensor would perform adequately at higher temperatures. According to contractor documents, the sensor did not perform as well as expected, and some target signals were degraded more than anticipated. Boeing disclosed in the report that sensor abnormalities were observed during the test and that the sensor experienced a higher than expected false alarm rate. These abnormalities were (1) a cooling mechanism that did not bring the sensor's silicon detectors to the intended operating temperature, (2) a power supply unit that created excess noise, and (3) software that did not function as designed because of the slow turnaround of the surrogate launch vehicle.

[6] The power supply unit is designed to power the sensor's electronic components.

In the report's summary, Boeing characterized the cooling mechanism's performance as somewhat below expectations but functioning well enough for the sensor's operation. In the body of the report, Boeing said that the fluctuations in temperature could lead to an apparent decrease in sensor performance. Additionally, Boeing engineers told us that the cooling mechanism's failure to bring the silicon detector array to the required temperature caused the detectors to be noisy. Because the discrimination software identifies objects as a warhead or a decoy by comparing the features of a target's signal with those it expects a warhead or decoy to display, a noisy signal may confuse the software. Boeing and TRW engineers said that they and program office officials were aware that there was a problem with the sensor's cooling mechanism before the test was conducted. However, Boeing believed that the sensor would perform adequately at higher temperatures. According to contractor documents, the sensor did not perform as well as expected, and some target signals were degraded more than anticipated.

Power Supply Creates Noise

The report also referred to a problem with the sensor's power supply unit and its effect on target signals. An expert we hired to evaluate the sensor's performance at higher than expected temperatures found that the power supply, rather than the temperature, was the primary cause of excess noise early in the sensor's flight. Boeing engineers told us that they were aware that the power supply was noisy before the test, but, as shown by the test, it was worse than expected.

Payload Launch Vehicle Affected Software's Ability to Remove Background Noise

The report explained that, as expected before the flight, the slow turnaround of the massive launch vehicle on which the sensor was mounted in Integrated Flight Test 1A caused the loss of some target signals. Engineers explained to us that the sensor would eventually be mounted on the lighter, more agile exoatmospheric kill vehicle, which would move back and forth to detect objects that did not initially appear in the sensor's field of view. The engineers said that Boeing designed software that takes into account the kill vehicle's normal motion to remove the background noise, but the software's effectiveness depended on the fast movement of the kill vehicle. Boeing engineers told us that, because of the slow turnaround of the launch vehicle used in the test, the target signals detected during the turnaround were particularly noisy and the software sometimes removed not only the noise but the entire signal as well.

Sensor Sometimes Detected False Targets

The report mentioned that the sensor experienced more false alarms than expected. A false alarm is a detection of a target that is not there. According to the experts we hired, during Integrated Flight Test 1A, the Boeing sensor often mistakenly identified noise produced by the power supply as signals from actual target objects. In a fully automated discrimination software program, a high false alarm rate could overwhelm the tracking software. Because the post-flight processing tools were not fully developed at the time of the August 13 and August 22, 1997, reports, Boeing did not rely upon a fully automated tracking system when it processed the Integrated Flight Test 1A data. Instead, a Boeing engineer manually tracked the target objects. The contractors realized, and reported to the Ground Based Interceptor Project Management Office, that numerous false alarms could cause problems in future flight tests, and they identified software changes to reduce their occurrence.

December 11 Briefing

On December 11, 1997, Boeing and TRW briefed officials from the Ground Based Interceptor Project Management Office and one of its support contractors on various anomalies observed during Integrated Flight Test 1A. The contractors' briefing charts explained the effect the anomalies could have on Integrated Flight Test 3, the first planned intercept test for the Boeing exoatmospheric kill vehicle, identified potential causes of the anomalies, and summarized the solutions to mitigate their effect. While some of the anomalies included in the December 11 briefing charts were referred to in the August 13 and August 22 reports, others were being reported in writing for the first time.

The anomalies referenced in the briefing charts included the sensor's high false alarm rate, the silicon detector array's higher-than-expected temperature, the software's low confidence factor that it had correctly identified two target objects correctly, the sensor's lower than expected probability of detection, and the software's elevation in rank of one target object toward the end of the test. In addition, the charts showed that an in-flight attempt to calibrate the sensor was inconsistent. According to the charts, actions to prevent similar anomalies from occurring or impacting Integrated Flight Test 3 had in most cases already been implemented or were under way.

Contractors Report Further on False Alarms

The contractors again recognized that a large number of false alarms occurred during Integrated Flight Test 1A. According to the briefing charts, false alarms occurred during the slow turnarounds of the surrogate launch vehicle. Additionally, the contractors hypothesized that some false alarms resulted from space-ionizing events. By December 11, engineers

had identified solutions to reduce the number of false alarms in future tests.

Briefing Charts Include Observations on Higher Detector Array Temperature

As they had in the August 22, 1997, report, the contractors recognized that the silicon detector array did not cool properly during Integrated Flight Test 1A. The contractors reported that higher silicon detector array temperatures could cause noisy signals that would adversely impact the detector array's ability to estimate the infrared intensity of observed objects. Efforts to eliminate the impact of the higher temperatures, should they occur in future tests, were on-going at the time of the briefing.

Some Software Confidence Factors Lower Than Expected

Contractors observed that the confidence factor produced by the software was small for two target objects. The software equation that makes a determination as to how confident the software should be to identify a target object correctly, did not work properly for the large balloon or multiple-service launch vehicle. Corrections to the equation had been made by the time of the briefing.

Sensor's Probability of Detection Is Lower Than Expected

The charts state that the Integrated Flight Test 1A sensor had a lower than anticipated probability of detection and a high false alarm rate. Because a part of the tracking, fusion, and discrimination software was designed for a sensor with a high probability of detection and a low false alarm rate, the software did not function optimally and needed revision. Changes to prevent this from happening in future flight tests were under way.

Software Increases the Rank of One Object Near Test's End

The briefing charts showed that TRW's software significantly increased the rank of one target object just before target objects began to leave the sensor's field of view. Although a later Integrated Flight Test 1A report stated the mock warhead was consistently ranked as the most likely target, the charts show that if in Integrated Flight Test 3 the same object's rank began to increase, the software could select the object as the intercept target. In the briefing charts, the contractors reported that TRW made a software change in the model that is used to generate reference data. When reference data was generated with the software change, the importance of the mock warhead was increased, and it was selected as the target. Tests of the software change were in progress as of December 11.

In-Flight Calibration Was Inconsistent

The Boeing sensor measures the infrared emissions of target objects by converting the collected signals into intensity with the help of calibration data obtained from the sensor prior to flight. However, the sensor was not calibrated at the higher temperature range that was experienced during Integrated Flight Test 1A. To remedy the problem, the sensor viewed a

star with known infrared emissions. The measurement of the star's intensity was to have helped fill the gaps in calibration data that was essential to making accurate measurements of the target object signals. Boeing disclosed that the corrections based on the star calibration were inconsistent and did not improve the match of calculated and measured target signatures. Boeing subsequently told us that the star calibration corrections were effective for one of the wavelength bands, but not for another, and that the inconsistency referred to in the briefing charts was in how these bands behaved at temperatures above the intended operating range. Efforts to find and implement solutions were in progress.

April 1, 1998, Report

On April 1, 1998, Boeing submitted a revised addendum to replace an addendum that had accompanied the August 22, 1997, report. This revised addendum was prepared in response to comments and questions submitted by officials from the Ground Based Interceptor Project Management Office, Nichols Research Corporation, and the Defense Criminal Investigative Service concerning the August 22 report. In this addendum, the contractors referred in writing to three problems and limitations that had not been addressed in earlier written test reports or the December 11 briefing. Contractors noted that a gap-filling module, which was designed to replace noisy or missing signals, did not operate as designed. They also disclosed that TRW's analysis of its discrimination software used target signals collected during a selected portion of the flight timeline and used a portion of the Integrated Flight Test 1A reference data that corresponded to this same timeline.

Gap-Filling Software Module Did Not Perform As Designed

The April 1 addendum reported that a gap-filling module that was designed to replace portions of noisy or missing target signals with expected signal values did not operate as designed. TRW officials told us that the module's replacement values were too conservative and resulted in a poor match between collected signals and the signals the software expected the target objects to display.

Assessment Uses Selected Target Signals

The April 1, 1998, addendum also disclosed that the August 13 and August 22 reports, in which TRW conveyed that its software successfully distinguished the mock warhead from decoys, were based on tests of the software using about one-third of the target signals collected during Integrated Flight Test 1A. We talked to TRW officials who told us that Boeing provided several data sets to TRW, including the full data set. The officials said that Boeing provided target signals from the entire timeline

to a TRW office that was developing a prototype version of the exoatmospheric kill vehicle's tracking, fusion, and discrimination software,[7] which was not yet operational. However, TRW representatives said that the test bed version of the software that TRW was using so that it could submit its analysis within 60 days of Integrated Flight Test 1A could not process the full data set. The officials said that shortly before the August 22 report was issued, the prototype version of the tracking, fusion, and discrimination software became functional and engineers were able to use the software to assess the expanded set of target signals. According to the officials, this assessment also resulted in the software's selecting the mock warhead as the most likely target. In our review of the August 22 report, we found no analysis of the expanded set of target signals. The April 1, 1998, report, did include an analysis of a few additional seconds of data collected near the end of Integrated Flight Test 1A, but did not include an analysis of target signals collected at the beginning of the flight.

Most of the signals that were excluded from TRW's discrimination analysis were collected during the early part of the flight, when the sensor's temperature was fluctuating. TRW told us that their software was designed to drop a target object's track if the tracking portion of the software received no data updates for a defined period. This design feature was meant to reduce false tracks that the software might establish if the sensor detected targets where there were none. In Integrated Flight Test 1A, the fluctuation of the sensor's temperature caused the loss of target signals. TRW engineers said that Boeing recognized that this interruption would cause TRW's software to stop tracking all target objects and restart the discrimination process. Therefore, Boeing focused its efforts on

[7] The purpose of TRW's tracking, fusion, and discrimination software, which was being designed to operate on-board Boeing's exoatmospheric kill vehicle, was to record the positions of the target objects as they moved through space, fuse information about the objects collected by ground-based radar with data collected by the kill vehicle's infrared sensor, and discriminate the warhead from decoys. The software's tracking function was not operational when the project office asked the contractors to determine the software's ability to discriminate. As a result, Boeing hand-tracked the target objects so that TRW could use test bed discrimination software, which is almost identical to the discrimination portion of the operational version of the tracking, fusion, and discrimination software, to assess the discrimination capability.

processing those target signals that were collected after the sensor's temperature stabilized and signals were collected continuously.[8]

Some signals collected during the last seconds of the sensor's flight were also excluded. The former TRW employee alleged that these latter signals were excluded because during this time a decoy was selected as the target. The Phase One Engineering Team cited one explanation for the exclusion of the signals. The team said that TRW stopped using data when objects began leaving the sensor's field of view. Our review did not confirm this explanation. We reviewed the target intensities derived from the infrared frames covering that period and found that several seconds of data were excluded before objects began to leave the field of view. Boeing officials gave us another explanation. They said that target signals collected during the last few seconds of the flight were streaking, or blurring, because the sensor was viewing the target objects as it flew by them. Boeing told us that streaking would not occur in an intercept flight because the kill vehicle would have continued to approach the target objects. We could not confirm that the test of TRW's discrimination software, as explained in the August 22, 1997, report, included all target signals that did not streak. We noted that the April 1, 1998, addendum shows that TRW analyzed several more seconds of target signals than is shown in the August 22, 1997, report. It was in these additional seconds that the software began to increase the rank of one decoy as it assessed which target object was most likely the mock warhead. However, the April 1, 1998, addendum also shows that even though the decoy's rank increased the software continued to rank the mock warhead as the most likely target. But, because not all of the Integrated Flight Test 1A timeline was presented in the April 1 addendum, we could not determine whether any portion of the excluded timeline might have been useful data and if there were additional seconds of useful data whether a target object other than the mock warhead might have been ranked as the most likely target.

Corresponding Portions of Reference Data Excluded

The April 1 addendum also documented that portions of the reference data developed for Integrated Flight Test 1A were also excluded from the

[8] When the Ground Based Interceptor Project Management Office asked Boeing to assess the discrimination capability of its sensor's software, TRW's prototype tracking, fusion, and discrimination software was not operational. To perform the requested assessment, TRW used test-bed discrimination software that was almost identical to the discrimination software that TRW engineers designed for the prototype tracking, fusion, and discrimination software. Because the test-bed software did not have the ability to track targets, Boeing performed the tracking function and provided the tracked signals to TRW.

discrimination analysis. Nichols and project office officials told us the software identifies the various target objects by comparing the target signals collected from each object at a given point in their flight to the target signals it expects each object to display at that same point in the flight. Therefore, when target signals collected during a portion of the flight timeline are excluded, reference data developed for the same portion of the timeline must be excluded.

Information Provided Verbally to Project Office

Officials in the National Missile Defense Joint Program Office's Ground Based Interceptor Project Management Office and Nichols Research told us that soon after Integrated Flight Test 1A the contractors orally disclosed all of the problems and limitations cited in the December 11, 1997, briefing and the April 1, 1998, addendum. Contractors made these disclosures to project office and Nichols Research officials during meetings that were held to review Integrated Flight Test 1A results sometime in late August 1997. The project office and contractors could not, however, provide us with documentation of these disclosures.

The current Ground Based Interceptor Project Management Office deputy manager said that the problems that contractors discussed with his office were not specifically communicated to others within the Department of Defense because his office was the office within the Department responsible for the Boeing contract. The project office's assessment was that these problems did not compromise the reported success of the mission, were similar in nature to problems normally found in initial developmental tests, and could be easily corrected.

Effect of Cooling Failure on Sensor's Performance

Because we questioned whether Boeing's sensor could collect any usable target signals if the silicon detector array was not cooled to the desired temperature, we hired sensor experts at Utah State University's Space Dynamics Laboratory to determine the extent to which the sub-optimal cooling degraded the sensor's performance. These experts concluded that the higher temperature of the silicon detectors degraded the sensor's performance in a number of ways, but did not result in extreme degradation. For example, the experts said the higher temperature reduced by approximately 7 percent the distance at which the sensor could detect targets. The experts also said that the rapid temperature fluctuation at the beginning and at the end of data acquisition contributed to the number of times that the sensor detected a false target. However, the experts said the major cause of the false alarms was the power supply

noise that contaminated the electrical signals generated by the sensor in response to the infrared energy. When the sensor signals were processed after Integrated Flight Test 1A, the noise appeared as objects, but they were actually false alarms.

Additionally, the experts said that the precision with which the sensor could estimate the infrared energy emanating from an object based on the electrical signal produced by the energy was especially degraded in one of the sensor's two infrared wave bands. In their report, the experts said that the Massachusetts Institute of Technology's Lincoln Laboratory analyzed the precision with which the Boeing sensor could measure infrared radiation and found large errors in measurement accuracy. The Utah State experts said that their determination that the sensor's measurement capability was degraded in one infrared wave band might partially explain the errors found by Lincoln Laboratory.

Although Boeing's sensor did not cool to the desired temperature during Integrated Flight Test 1A, the experts found that an obstruction in gas flow rather than the sensor's design was at fault. These experts said the sensor's cooling mechanism was properly designed and Boeing's sensor design was sound.

Appendix II: Evaluations of TRW's Software and a Planned Enhancement

Nichols Research Corporation and the Phase One Engineering Team tested TRW's discrimination software and a planned enhancement to that software, known as the Extended Kalman Filter Feature Extractor. Nichols concluded that although it had weaknesses, the discrimination software met performance requirements established by Boeing when it was tested against a simple threat and given near perfect knowledge about the key characteristics, or features, that the target objects would display during flight. The Phase One Engineering Team reported that despite some weaknesses, TRW's discrimination software was well designed and worked properly. Like Nichols, the team found that the software's performance was dependent upon prior knowledge of the target objects. Because Nichols did not test the software's capability using data collected from Integrated Flight Test 1A and the Phase One Engineering Team did not process the raw data from Integrated Flight Test 1A or develop its own reference data, neither group can be said to have definitively proved or disproved TRW's claim that its software successfully identified the mock warhead from decoys using data collected from Integrated Flight Test 1A. From their assessments of TRW's Extended Kalman Filter Feature Extractor, both groups concluded that it was feasible that the Filter could provide additional information about target objects, but neither group determined to what extent the Filter would improve the software's discrimination performance.

Nichols Evaluation of TRW's Discrimination Software

Nichols Research Corporation evaluated TRW's discrimination software to determine if it met performance requirements developed by Boeing. Boeing established discrimination performance requirements to ensure that its exoatmospheric kill vehicle, when fully developed, could destroy a warhead with the single shot precision (expressed as a probability) required by the Ground Based Interceptor Project Management Office.[1] The kill vehicle must perform a number of functions successfully to accurately hit-to-kill its target, such as acquiring the target cluster, discriminating the warhead from other objects, and diverting to hit the warhead. Boeing believed that if it met the performance requirements that it established for each function, including the discrimination function, the exoatmospheric kill vehicle should meet the required single shot probability of kill.

[1] The contractors can prove they have achieved this number only after numerous engagements. Because tests are expensive, the Department expects the contractors to use simulations to demonstrate the required precision. Actual flight tests are used to prove the reliability of the simulations.

Nichols' Methodology

To determine if TRW's software performed as required, Nichols' engineers obtained a copy of TRW's software; verified that the software was based on sound scientific and engineering principles; validated that it operated as designed; and tested its performance in 48 simulated scenarios that included countermeasures, such as decoys, that the system might encounter before 2010.

Nichols validated the software by obtaining a copy of the actual source code from TRW and installing the software in a Nichols computer. Engineers then examined the code line-by-line; verified its logic, data flow, and input and output; and determined that the software accurately reflected TRW's baseline design.

Nichols next verified that the software performed exactly as reported by TRW. Engineers ran 13 TRW-provided test cases through the software and compared the results to those reported by TRW. Nichols reported that their results were generally consistent with TRW's results with only minor performance differences in a few cases.

After analyzing the 13 reference cases, Nichols generated additional test cases by simulating a wide-range of enemy missiles with countermeasures that included decoys. Including the 13 reference cases, Nichols analyzed the software's performance in a total of 48 test scenarios.

Nichols' Key Results

Because the software performed successfully in 45 of 48 simulated test cases, Nichols concluded that the system met the performance requirements established by Boeing. However, Nichols explained that the software met its requirement because it was tested against a simple threat. In addition, Nichols said that the software was given nearly perfect knowledge of the features the simulated warhead and any decoys included in each test would display.

- Nichols found anomalies when it simulated the performance of TRW's software. Nichols' December 2, 1997, report identified anomalies that prevented the software from meeting its performance requirement in 3 of the 48 cases. First, Nichols found that a software module did not work properly. (TRW used this gap-filling software module to replace missing or noisy target signals.) Second, Nichols found that the software's target selection logic did not always work well. As a result, the probability that the software would select the simulated warhead as the target was lower than required in three of the test cases.

- Nichols reported inconsistencies in TRW's software code. Engineers found that in some cases the software did not extract one particular feature from the target signals, and, in other cases, the results improved substantially when this feature was excluded. The Nichols report warned that in cases where this feature was the most important in the discrimination process, the software's performance could be significantly degraded.

- Evaluation Parameters. In its 1997 report, Nichols cautioned that TRW's software met performance requirements because the countermeasures included in the 48 tests were relatively simple. Nichols' testing also assumed perfect knowledge about the warhead and decoys included in the simulations. Engineers told us that all 48 test cases were constructed to test the software against the simple threats that the Department of Defense believed "nations of concern" might deploy before 2010. The engineers said that their evaluation did not include tests of the software against the number and type of decoys deployed in Integrated Flight Test 1A because that threat cluster was more complex than the simple threat that contractors were required to design their software to handle. In addition, Nichols reported that in all 48 test cases perfect reference data was used—that is, the software was told what features the warhead and decoys would display during the simulations. Nichols engineers said TRW's software is sensitive to prior knowledge about the threat and the Ground Based Interceptor Project Management Office was aware of this aspect of TRW's design.

Limitation in Nichols' Evaluation

Nichols' evaluation was limited because it did not test TRW's software using actual flight data from Integrated Flight Test 1A. Nichols told us that in addition to testing TRW's discrimination software using simulated data it had also planned to assess the software's performance using real target signals collected during Integrated Flight Test 1A. Because it did not perform this assessment, Nichols can not be said to have definitively proved or disproved TRW's claim that its software discriminated the mock warhead from decoys using data collected from Integrated Flight Test 1A. Officials said they did not complete this aspect of the evaluation because their resources were limited. However, we noted that Nichols' engineers had already verified TRW's software and obtained the raw target signals collected during Integrated Flight Test 1A. These engineers told us that this assessment could be done within two weeks after Nichols received all required information. (Nichols said it did not have some needed information.)

Phase One Engineering Team's Evaluation

In 1998, the National Missile Defense Joint Program Office asked the Phase One Engineering Team to conduct an assessment, using available data, of TRW's discrimination software, even though Nichols Research Corporation had already concluded that it met the requirements established by Boeing.[2] The program office asked for the second evaluation because the Defense Criminal Investigative Service lead investigator expressed concern about the ability of Nichols to provide a truly objective evaluation.

The Phase One Engineering Team developed a methodology to (1) determine if TRW's software was consistent with scientific, mathematical, and engineering principles; (2) determine whether TRW accurately reported that its software successfully discriminated a mock warhead from decoys using data collected from Integrated Flight Test 1A; and (3) predict the performance of TRW's basic discrimination software against Integrated Flight Test 3 scenarios. The key results of the team's evaluation were that the software was well designed; the contractors accurately reported the results of Integrated Flight Test 1A; and the software would likely perform successfully in Integrated Flight Test 3. The primary limitation was that the team used Boeing- and TRW-processed target data and TRW-developed reference data in determining the accuracy of TRW reports for Integrated Flight Test 1A.

Phase One Engineering Team's Methodology

The team began its work by assuring itself that TRW's discrimination software was based on sound scientific, engineering, and mathematical principles and that those principles had been correctly implemented. It did this primarily by studying technical documents provided by the contractors and the program office. Next, the team began to look at the software's performance using Integrated Flight Test 1A data. The team studied TRW's August 13 and August 22, 1997, test reports to learn more about discrepancies that the Defense Criminal Investigative Service said it found in these reports. Team members also received briefings from the

[2] The Ground Based Interceptor Project Management Office identified the precision (expressed as a probability) with which the exoatmospheric kill vehicle is expected to destroy a warhead with a single shot. To ensure that the kill vehicle would meet this requirement, Boeing established lower-level requirements for each function that affects the kill vehicle's performance, including the discrimination function. Nichols compared the contractor-established software discrimination performance requirement to the software's performance in simulated scenarios.

Defense Criminal Investigative Service, Boeing, TRW, and Nichols Research Corporation.

Team members told us that they did not replicate TRW's software in total. Instead, the team emulated critical functions of TRW's discrimination software and tested those functions using data collected during Integrated Flight Test 1A. To test the ability of TRW's software to extract the features of each target object's signal, the team designed a software routine that mirrored TRW's feature-extraction design. Unlike Nichols, the team did not obtain target signals collected during the test and then process those signals. Rather, the team received Integrated Flight Test 1A target signals that had been processed by Boeing and then further processed by TRW. These signals represented about one-third of the collected signals. Team members input the TRW-supplied target signals into the team's feature-extraction software routine and extracted two features from each target signal. The team then compared the extracted features to TRW's reports on these same features and concluded that TRW's software-extraction process worked as reported by TRW. Next, the team acquired the results of 200 of the 1,000 simulations that TRW had run to determine the features that target objects deployed in Integrated Flight Test 1A would likely display.[3] Using these results, team members developed reference data that the software could compare to the features extracted from Integrated Flight Test 1A target signals. Finally, the team wrote software that ranked the different observed target objects in terms of the probability that each was the mock warhead. The results produced by the team's software were then compared to TRW's reported results.

The team did not perform any additional analysis to predict the performance of the Boeing sensor and its software in Integrated Flight Test 3. Instead, the team used the knowledge that it gained from its assessment of the software's performance using Integrated Flight Test 1A data to estimate the software's performance in the third flight test.

[3] The Phase One Engineering Team reported that TRW ran 1,000 simulations to determine the reference data for Integrated flight Test 1A, but the Team received the results of only 200 simulations. TRW engineers said this was most likely to save time. Also, the engineers said that the only effect of developing reference data from 200 simulations rather than 1,000 simulations is that confidence in the reference data drops from 98 percent to approximately 96 percent.

The Phase One Engineering Team's Key Results

In its report published on January 25, 1999, the Phase One Engineering Team reported that even though it noted some weaknesses, TRW's discrimination software was well designed and worked properly, with only some refinement or redesign needed to increase the robustness of the discrimination function. In addition, the team reported that its test of the software using data from Integrated Flight Test 1A produced essentially the same results as those reported by TRW. The team also predicted that the Boeing sensor and its software would perform well in Integrated Flight Test 3 if target objects deployed as expected.

Weaknesses in TRW's Software

The team's assessment identified some software weaknesses. First, the team reported that TRW's use of a software module to replace missing or noisy target signals was not effective and could actually hurt rather than help the performance of the discrimination software. Second, the Phase One Engineering Team pointed out that while TRW proposed extracting several features from each target-object signal, only a few of the features could be used.

The Phase One Engineering Team also reported that it found TRW's software to be fragile because the software was unlikely to operate effectively if the reference data—or expected target signals—did not closely match the signals that the sensor collected from deployed target objects. The team warned that the software's performance could degrade significantly if incorrect reference data were loaded into the software. Because developing good reference data is dependent upon having the correct information about target characteristics, sensor-to-target geometry, and engagement timelines, unexpected targets might challenge the software. The team suggested that very good knowledge about all of these parameters might not always be available.

Accuracy of Contractors' Integrated Flight Test 1A Reports

The Phase One Engineering Team reported that the results of its evaluation using Integrated Flight Test 1A data supported TRW's claim that in post-flight analysis its software accurately distinguished a mock warhead from decoys. The report stated that TRW explained why there were differences in the discrimination analysis included in the August 13, 1997, Integrated Flight Test 1A test report and that included in the August 22, 1997, report. According to the report, one difference was that TRW mislabeled a chart in the August 22 report. Another difference was that the August 22 discrimination analysis was based on target signals collected over a shorter period of time (see app. I for more information regarding TRW's explanation of report differences). Team members said that they found TRW's explanations reasonable.

Predicted Success in Integrated Flight Test 3

The Phase One Engineering Team predicted that if the targets deployed in Integrated Flight Test 3 performed as expected, TRW's discrimination software would successfully identify the warhead as the target. The team observed that the targets proposed for the flight test had been viewed by Boeing's sensor in Integrated Flight Test 1A and that target-object features collected by the sensor would be extremely useful in constructing reference data for the third flight test. The team concluded that given this prior knowledge, TRW's discrimination software would successfully select the correct target even in the most stressing Integrated Flight Test 3 scenario being considered, if all target objects deployed as expected. However, the team expressed concern about the software's capabilities if objects deployed differently, as had happened in previous flight tests.

Limitations of the Team's Evaluation

The Phase One Engineering Team's conclusion that TRW's software successfully discriminated is based on the assumption that Boeing's and TRW's input data were accurate. The team did not process the raw data collected by the sensor's silicon detector array during Integrated Flight Test 1A or develop their own reference data by running hundreds of simulations. Instead, the team used target signature data extracted by Boeing and TRW and developed reference data from a portion of the simulations that TRW ran for its own post-flight analysis. Because it did not process the raw data from Integrated Flight Test 1A or develop its own reference data, the team cannot be said to have definitively proved or disproved TRW's claim that its software successfully discriminated the mock warhead from decoys using data collected from Integrated Flight Test 1A. A team member told us its use of Boeing- and TRW-provided data was appropriate because the former TRW employee had not alleged that the contractors tampered with the raw test data or used inappropriate reference data.

Evaluations of the Extended Kalman Filter Feature Extractor

Nichols Research Corporation and the Phase One Engineering Team evaluated TRW's Extended Kalman Filter Feature Extractor and determined that it could provide additional information to TRW's discrimination software. However, Nichols Research told us that its evaluation was not an exhaustive analysis of the Filter's capability, but an attempt to determine if a Kalman Filter—which is frequently used to estimate such variables as an object's position or velocity—could extract a feature from an infrared signal. The Phase One Engineering Team reported that because of the limited time available to assess both TRW's discrimination software and the Extended Kalman Filter Feature Extractor, it did not rigorously test the Filter. Its analysis was also aimed

at determining whether the Filter could extract a feature from target objects.

Nichols' Assessments of the Extended Kalman Filter Feature Extractor

Nichols engineers assessed TRW's application of the Kalman Filter in 1996 and again in 1998. For both evaluations, Nichols engineers constructed a stand-alone version of the Filter (the Filter is comprised of mathematical formulas converted into software code) that the engineers believed mirrored TRW's design. However, Nichols designed its 1996 version of the Filter from information extracted and pieced together from multiple documents and without detailed design information from TRW engineers. Nichols Research Corporation and Ground Based Interceptor Project Management Office officials said the Nichols' engineers did not talk with TRW's engineers about the Filter's design because the project office was limiting communication with the contractors in order to prevent disclosure of contractors' proprietary information during the source selection for the exoatmospheric kill vehicle.

Nichols' 1996 Evaluation Methodology and Key Results

In 1996, Nichols engineers tested the Filter's ability to extract the features of simulated signals representative of threat objects. Engineers said that under controlled conditions they attempted to determine from which signals the Filter could extract features successfully and from which signals it could not. Also, because the Filter could not begin to extract features from the target objects unless it had some advance knowledge about the signal, engineers conducted tests to determine how much knowledge about initial conditions the Filter needed.

In its November 1996 report, Nichols concluded that the Filter was unlikely to enhance the capability of TRW's discrimination software. The assessment showed that the Filter could not extract the features of a signal unless the Filter had a great deal of advance knowledge about the signal. It also showed that the Filter was sensitive to "noise" (undesirable energy that degrades the target signal).

Nichols' 1998 Evaluation Methodology and Key Results

By 1998, the competitive phase of the exoatmospheric kill vehicle contracts was over. Based on additional understanding of the Filter's implementation, coupled with its proposed candidacy as an upgrade to the discrimination software, the Ground Based Interceptor Project Management Office asked Nichols to test the Filter again. Nichols engineers were now able to hold discussions with TRW engineers regarding their respective Filter designs. From these discussions, Nichols learned that it had designed two elements of the Filter differently from

TRW. The primary difference was in the number of filters that Nichols and TRW used to preprocess the infrared signals before the feature extraction began. Nichols' design included only one pre-processing filter, while TRW's included several. There was also one less significant difference, which was the difference in a delay time before feature extraction began. Nichols modified its version to address these differences.

In its second assessment, Nichols again examined the feature extraction capability of the Filter. Engineers pointed out that in both assessments the Filter was tested as stand-alone software, not as an integrated part of TRW's discrimination software program.

The new tests showed that the redesigned Filter could perform well against the near-term threat. However, in its report, Nichols expressed reservations that unless the target and specifics of the target's deployment were well defined, the Filter's performance would likely be sub-optimal. Nichols also pointed out that the Filter was unlikely to perform well against targets that exhibited certain characteristics.

Limitations of Nichols' 1998 Assessment

Nichols tested the ability of the Extended Kalman Filter Feature Extractor to extract features over a wide range of object dynamics and characteristics, including elements of the far-term threat.[4] Nichols demonstrated the Filter's ability to extract information (features), but did not assess the Filter's potential impact on the TRW discrimination design.

Because it did not assess the discrimination capability of the Extended Kalman Filter, Nichols could not predict how the Filter would have performed against either the target complex for Integrated Flight Test 1A or the target complex proposed for Integrated Flight Test 3. Target sets for Integrated Flight Test 1A and initially proposed for Integrated Flight Test 3 were more complex than the near-term threat that Nichols tested the Filter against.

In their discussions with us, Nichols' engineers stressed that their assessments should be viewed as an evaluation of a technology concept, not an evaluation of a fully integrated component of the discrimination software. Engineers admitted that their approach to this assessment was less thorough than the evaluation they conducted of TRW's discrimination

[4] The far-term threat consists of missiles and their warheads and decoys that nations of concern are expected to deploy after 2010.

software and that engineers did not fully understand why the additional bank of pre-processing filters improved the Filter's performance.[5] They said a more systematic analysis would be needed to fully evaluate the Filter's performance.

Phase One Engineering Team's Assessment of the Extended Kalman Filter Feature Extractor

The National Missile Defense Joint Program Office did not originally ask the Phase One Engineering Team to evaluate TRW's application of the Kalman Filter. However, the team told us that program officials later asked them to do a quick assessment as an addition to their evaluation of TRW's software. Team members designed an Extended Kalman Filter Feature Extractor similar to TRW's. Like Nichols first design, the Phase One Engineering Team's design was not identical to TRW's Filter. In fact, the team did not include any filters to preprocess the infrared signals before the feature extraction began.[6]

The Phase One Engineering Team tested the capability of its Filter against one simulated target object and one of the objects whose signal was collected during Integrated Flight Test 1A. The team reported that the Filter did stabilize and extract the features of the objects' infrared signals. However, the team added the caveat that the Filter would need good initial knowledge about the target object before it could begin the extraction process.

The team reported that its evaluation of the Filter was limited. It said it did not evaluate the Filter's sensitivity to noise, the information the Filter needed to begin operation, or the extent to which the Filter would improve the performance of the discrimination software.

[5] Nichols reported that additional pre-processing filters resulted in a more stable design for the Extended Kalman Filter Feature Extractor. However, Nichols engineers did not fully understand exactly which mathematical formulas, or parts of formulas, were affected by the addition of the filters.

[6] The team did not find as Nichols had that the absence of the pre-processing filters degraded the performance of the Filter.

Appendix III: Justice's Decision Not to Join Lawsuit

Before deciding in March 1999 not to intervene in the False Claims lawsuit brought by the former TRW employee, the Department of Justice considered scientific reports and information from two Army sources. Specifically, Justice relied upon evaluations of TRW's software conducted by the Nichols Research Corporation and the Phase One Engineering Team (see appendix II for more information on these evaluations), information provided by the Army Space and Missile Defense Command, and a recommendation made by the Army Legal Services Agency. Justice officials told us that the input of the Space and Missile Defense Command carried more weight in the decision-making process than the recommendation by the Army Legal Services Agency because the Command is the contracting agency for the kill vehicle and is therefore more familiar with the contractors involved as well as the technical details of the lawsuit.

Information from the Army Space and Missile Defense Command

The Army Space and Missile Defense Command was brought into this matter in response to an inquiry by the Department of Justice concerning the vouchers that were submitted for cost reimbursement by Boeing for work performed by its subcontractor, TRW. Specifically, Justice asked whether the Army would have paid the contractor's vouchers if Boeing and TRW had misrepresented the capabilities of the software in the vouchers. In a letter to Justice, dated February 24, 1999, the Command stated that the Army did not consider the vouchers submitted by Boeing for TRW's work to be false claims. The letter cited the Nichols' and Phase One Engineering Team's reports as support for its conclusions and noted that a cost-reimbursement research-and-development contract only requires that the contractor exercise its "best efforts."

Recommendation of the Army Legal Services Agency

There is some uncertainty about how the Army Legal Services Agency[1] came to recommend in February 1999 that Justice not intervene in the lawsuit. Army Legal Services had very little documentation to explain the recommendation, and agency officials told us that they remember very little about the case. The agency's letter stated that it was basing its recommendation on conversations with investigators handling the case and on the former TRW engineer's wishes.

[1] The Legal Services Agency serves as the liaison to the Department of Justice for procurement fraud cases involving Army programs.

However, the lead investigator in the case (from the Defense Criminal Investigative Service) stated that he and his team had not recommended to the Army that the case not proceed. The little documentation available shows only that the case attorney's predecessor spoke with the lead investigator shortly after the case was opened. Officials said they could not remember why they cited conversations with case investigators in the letter and agreed that there were no other investigators aside from those in the Defense Criminal Investigative Service. One official stressed that the letter does not explicitly say that the investigators recommended nonintervention.

As for the engineer's wishes, Army Legal Services has no record of direct contacts with the engineer, and agency officials acknowledged that they probably obtained information about the engineer's wishes from Justice. Agency officials also said they could not remember why they cited the engineer's wishes in their letter. The engineer told us that she did tell Justice that if it was not going to help, it should not hinder the case. The engineer also told us that this may have been misinterpreted by the agency as a refusal of any help. Justice officials agreed that the engineer consistently wanted Justice to take up the case. Legal Services Agency officials noted that it would be very unusual for someone not to want help from Justice, especially considering that less than 10 percent of False Claims cases succeed when Justice is not involved.

Army Legal Services Agency officials said that the case was one of several hundred the agency handles at any one time and that their involvement in a case like this one is usually minimal, unless the agency is involved in the prosecution. The officials stated that the Army Space and Missile Defense Command letter likely would have influenced their own letter because the Command's deputy counsel was recognized for his expertise in matters of procurement fraud. They also said that they relied on Justice to provide information about technical details of the case. The case attorney stated that he had not reviewed the Phase One Engineering Team or Nichols studies.

Appendix IV: Steps to Assure Independent and Objective Review by the Phase One Engineering Team

The Defense Criminal Investigative Service, which was investigating the allegations against Boeing and TRW, asked the National Missile Defense Joint Program Office to establish an independent panel to evaluate the capability and performance of TRW's discrimination software. Although Nichols Research Corporation, a support contractor overseeing Boeing's work, had already conducted such an assessment and reported that the software met requirements, the case investigator was concerned about the ability of Nichols to provide a truly objective assessment. In response to the investigator's request, the program office utilized an existing advisory group, known as the Phase One Engineering Team, to conduct the second assessment. Comprised of various Federally Funded Research and Development Centers, this group had been established by the Strategic Defense Initiative Organization in 1988 in order to provide the program office access to a continuous, independent and objective source of technical and engineering expertise. Since Federally Funded Research and Development Centers are expressly authorized, established and operated to provide the government with independent and objective advice, the Joint Program Office officials determined that making use of such a group would be sufficient to assure an independent and objective review. Scientific associations, however, said that there are alternative ways of choosing a panel to review contentious issues. Nonetheless, program officials said that establishing a review team using such methods would likely have increased the time the reviewers needed to complete their work and could have increased the cost of the review.

The Phase One Engineering Team

When the National Missile Defense Joint Program Office determined that it should undertake a review of the TRW discrimination software because of allegations that contractors had misrepresented their work, it turned to the Phase One Engineering Team. The Phase One Engineering Team was established in 1988 by the Strategic Defense Initiative Organization—later known as the Ballistic Missile Defense Organization—as an umbrella mechanism to obtain technical and engineering support from Federally Funded Research and Development Centers. To ensure that the individual scientists who work on each review undertaken by the Phase One Engineering Team have the requisite expertise, membership on each review team varies with each assignment. When asked to advise a program, the director of the Phase One Engineering Team determines which Federally Funded Research and Development Centers have the required expertise. The director then contacts officials at those centers to identify the appropriate scientists for the task. According to the director, the National Missile Defense Joint Program Office does not dictate the individuals who work on a Phase One Engineering Team review. When the

director received the request to conduct a review of TRW's discrimination software, he determined there were three Federally Funded Research and Development Centers best suited to undertake this review. A total of five scientists were then selected from these three centers to comprise the review team: one member from the Aerospace Corporation, sponsored by the U.S. Air Force; two members from the Massachusetts Institute of Technology's Lincoln Laboratory, also sponsored by the U.S. Air Force; and two members from the Lawrence Livermore National Laboratory, sponsored by the Department of Energy.

The federal government established the Federally Funded Research and Development Centers to meet special or long-term research or development needs of the sponsoring federal government agencies that were not being met effectively by existing in-house or contractor resources.[1] The federal government enters into long-term relationships with the Federally Funded Research and Development Centers in order to encourage them to provide the continuity that allows them to attract high quality personnel who will maintain their expertise, retain their objectivity and independence, preserve familiarity with the government's needs, and provide a quick response capability.[2] To achieve these goals, the Federally Funded Research and Development Centers must have access, beyond that required in normal contractual relationships with the government, to government and supplier information, sensitive or proprietary data, and to employees and facilities. Because of this special access, the Federally Funded Research and Development Centers are required by the Federal Acquisition Regulation and agreements with their sponsoring agencies to operate in the public interest with objectivity and independence, to be free from organizational conflicts of interest, and to fully disclose their affairs to the sponsoring agency.[3] To further ensure that they are free from organizational conflicts of interest, Federally Funded Research and Development Centers are operated, managed, and/or administered by a university or consortium of universities; other not-for-profit or non-profit organization; or an industrial firm, as an autonomous organization or as an identifiable separate operating unit of a parent organization.[4]

[1] Federal Acquisition Regulation 35.017(a)(2).

[2] Federal Acquisition Regulation 35.017(a)(4).

[3] Federal Acquisition Regulation 35.017(a)(2).

[4] Federal Acquisition Regulation 35.017(a)(3).

All three of the Federally Funded Research and Development Centers involved in this review had entered into sponsoring agreements and contracts with their respective sponsoring agencies that contain the requirements imposed on such Centers by the Federal Acquisition Regulation. For example, the sponsoring agreement between the Air Force and Lincoln Laboratory requires that Lincoln Laboratory avoid any action that would put its personnel in perceived or actual conflicts of interest regarding either unfair competition or objectivity. Joint Program Office officials said they relied upon adherence to the governing regulations and sponsoring agreements to assure themselves that the members of this review team could provide a fresh, unbiased look at TRW's software.

Officials with whom we spoke expressed confidence in the team's independence. Justice officials said that they had no reason to doubt the objectivity or independence of the review team's members nor the seriousness and thoroughness of their effort. The Phase One Engineering Team director told us that independence is a program goal and that their reviews report the technical truth regardless of what the National Missile Defense Joint Program Office might want to hear. The director noted that the best way to ensure independence is to have the best scientists from different organizations discuss the technical merits of an issue.

Alternative Panel Composition

At your request, we spoke with officials of the National Academy of Sciences and the American Physical Society who told us that there are alternative ways to choose a panel. One method commonly used by these scientific organizations, which frequently conduct studies and evaluate reports or journal articles, is peer review. According to a GAO report[5] that studied federal peer review practices, peer review is a process wherein scientists with knowledge and expertise equal to that of the researchers whose work they review make an independent assessment of the technical or scientific merit of that research. According to the Phase One Engineering Team director, the evaluation performed by the team assigned to review TRW's software was a type of peer review. However, National Academy of Sciences and American Physical Society officials told us that since individuals knowledgeable in a given area often have opinions or biases, an unbiased study team should include members who would, as a group, espouse a broad spectrum of opinions and interests. Such a team

[5] *Federal Research: Peer Review Practices at Federal Science Agencies Vary (GAO/RCED-99-99, March 17, 1999).*

should include both supporters and critics of the issue being studied. These officials told us that it was their opinion that the Phase One Engineering Team members are "insiders" who are unlikely to be overly critical of the National Missile Defense program.

The National Missile Defense Joint Program Office official who requested that the Phase One Engineering Team conduct such a review said that he could have appointed a panel such as that suggested by the National Academy of Sciences and the American Physical Society. But he said that he wanted a panel that was already knowledgeable about warhead discrimination in space and required little additional knowledge to complete its review. The official noted that the team's report was originally intended to be a one-to-two-month effort, even though it eventually took about eight months to complete. Some additional time was required to address further issues raised by the Defense Criminal Investigative Service. A team member said that the statement of work was defined so that the panel could complete the evaluation in a timely manner with the data available. Officials of the National Academy of Sciences and the American Physical Society acknowledged that convening a panel such as the type they suggested would likely have required more time and could have been more costly.

Appendix V: Boeing Integrated Flight Test 1A Requirements and Actual Performance as Reported by Boeing and TRW

The table below includes selected requirements that Boeing established before the flight test to evaluate sensor performance and the actual sensor performance characteristics that Boeing and TRW discussed in the August 22 report.

Table 2: Integrated Flight-Test 1A Requirements Established by Boeing and Actual Performance

Capability Tested[a]	Requirement	Integrated Flight Test 1A performance reported by Boeing/TRW
Acquisition range[b]	The sensor subsystem shall acquire the target objects at a specified distance.	The performance exceeded the requirement.[c]
Probability of detection	The sensor shall detect target objects with a specified precision, which is expressed as a probability.	The performance satisfied the requirement.
False alarm rate	False alarms shall not exceed a specified level.	The performance did not satisfy the requirement. The false alarm rate exceeded Boeing's requirement by more than 200 to 1 because of problems with the power supply and the higher than expected temperature of the sensor.
Infrared radiation measurement precision	The sensor subsystem shall demonstrate a specified measurement precision at a specified range.	The contractor met the requirement in one infrared measurement band, but not in another.
Angular Measurement Precision (AMP)	Given specified conditions, the sensor subsystem shall determine the angular position of the targets with a specified angular measurement precision.	The performance was better than the requirement.
Closely spaced objects resolution	Resolution of closely spaced objects shall be satisfied at a specified range.	The closely spaced objects requirement could not be validated because the targets did not deploy with the required separation.
Silicon detector array cool-down time	The time to cool the silicon detector array to less than a desired temperature shall be less than or equal to a specified length of time.	The performance did not satisfy the requirement because the desired temperature was not reached. Nevertheless, the silicon detector operated as designed at the higher temperatures.
Hold time[d]	With a certain probability, the silicon detector array's temperature shall be held below a desired temperature for a specified minimum length of time.	Even though the detector array's temperature did not reach the desired temperature, the array was cooled to an acceptable operating temperature and held at that temperature for longer than required.

[a]The requirements displayed in the table were established by the contractor and were not imposed by the government. Additionally, because of various sensor problems recognized prior to the test, Boeing waived most of the requirements. Boeing established these requirements to ensure that its exoatmospheric kill vehicle, when fully developed, could destroy a warhead with the single shot precision (expressed as a probability) required by the Ground Based Interceptor Project Management Office.

[b]Boeing's acquisition range specification required that the specified range, detection probability, and false alarm rate be achieved simultaneously. Boeing's Chief Scientist said that because the range and target signals varied with time and the total observation time was sharply limited during Integrated Flight Test 1A, the probability of detection could not be accurately determined. As a result, the test was not a suitable means for assessing whether the sensor can attain the specified acquisition range.

[c]The revised 60-day report states that the sensor did not detect the target until approximately two-thirds of the nominal acquisition range. Boeing engineers told us that while this statement appears to contradict the claim that the target was acquired at 107 percent of the specified range, it does not. Boeing engineers said that the nominal acquisition range refers to the range at which a sensor that is performing as designed would acquire the target, which is a substantially greater range than the specified acquisition range. However, neither Boeing nor TRW could provide documentation of the nominal acquisition range so that we could verify that these statements are not contradictory.

[d]In the main body of the August 22 report, the contractor discussed "hold time." However, it is not mentioned in the appendix to the August 22 report that lists the performance characteristics against which Boeing planned to evaluate its sensor's performance. Rather, the appendix refers to a "minimum target object viewing" time, which has the same requirement as the hold time. Boeing reported that its sensor collected target signals over approximately 54 seconds.

Appendix VI: Scope and Methodology

We determined whether Boeing and TRW disclosed key results and limitations of Integrated Flight Test 1A to the National Missile Defense Joint Program Office by examining test reports submitted to the program office on August 13, 1997, August 22, 1997, and April 1, 1998, and by examining the December 11, 1997, briefing charts. We also held discussions with and examined various reports and documents prepared by Boeing North American, Anaheim, California; TRW Inc., Redondo Beach, California; the Raytheon Company, Tucson, Arizona; Nichols Research Corporation, Huntsville, Alabama; the Phase One Engineering Team, Washington, D.C.; the Massachusetts Institute of Technology/Lincoln Laboratory, Lexington, Massachusetts; the National Missile Defense Joint Program Office, Arlington, Virginia, and Huntsville, Alabama; the Office of the Director, Operational Test and Evaluation, Washington D.C.; the U.S. Army Space and Missile Defense Command, Huntsville, Alabama; the Defense Criminal Investigative Service, Mission Viejo, California, and Arlington, Virginia; and the Institute for Defense Analyses, Alexandria, Virginia.

We held discussions with and examined documents prepared by Dr. Theodore Postol, Massachusetts Institute of Technology, Cambridge, Massachusetts; Dr. Nira Schwartz, Torrance, California; and Mr. Roy Danchick, Santa Monica, California.

In addition, we hired the Utah State University Space Dynamics Laboratory, Logan, Utah, to examine the performance of the Boeing sensor because we needed to determine the effect the higher operating temperature had on the sensor's performance. As agreed with your offices, we did not replicate TRW's assessment of its software using target signals that the Boeing sensor collected during the test. This would have required us to make engineers and computers available to verify TRW's software, format raw target signals for input into the software, develop reference data, and run the data through the software. We did not have these resources available, and we, therefore, cannot attest to the accuracy of TRW's discrimination claims.

We examined the methodology, key results, and limitations of evaluations completed by Nichols Research Corporation and the Phase One Engineering Team by analyzing Nichols' report on TRW's discrimination software dated December 1997; Nichols' reports on the Extended Kalman Filter dated November 1996 and November 1998; and the Phase One Engineering Team's "Independent Review of TRW Discrimination Techniques" dated January 1999. In addition, we held discussions with the Nichols engineers and Phase One Engineering Team members that

conducted the assessments and with officials from the National Missile Defense Joint Program Office. We did not replicate the evaluations conducted by Nichols and the Phase One Engineering Team and cannot attest to the accuracy of their reports.

We examined the basis for the Department of Justice's decision not to intervene in the False Claims lawsuit by holding discussions with and examining documents prepared by the Department of Justice, Washington, D.C. We also held discussions with and reviewed documents at the U.S. Army Legal Services Agency, Arlington, Virginia, and the U.S. Army Space and Missile Defense Command, Huntsville, Alabama.

We reviewed the National Missile Defense Joint Program Office's efforts to address potential conflicts of interest that an expert panel might have in reviewing the results of Integrated Flight Test 1A by holding discussions with National Missile Defense Joint Program Office officials and with members of the expert panel, known as the Phase One Engineering Team. We also examined the federal regulations and support agreements agreed to by the Federally Funded Research and Development Centers and national laboratory that employed the panel members. Last, as you requested, we discussed alternative methods of establishing an expert panel with the American Physical Society, Ridge, New York; and the National Academy of Sciences' National Research Council, Washington, D.C.

Our work was conducted from August 2000 through February 2002 according to generally accepted government auditing standards. The length of time the National Missile Defense Joint Program Office required to release documents to us significantly slowed our review. For example, the program office required approximately 4 months to release key documents such as Nichols 1997 evaluation of TRW's discrimination software and Nichols 1996 and 1998 evaluations of the Extended Kalman Filter Feature Extractor. We requested these and other documents on September 14, 2000, and received them on January 9, 2001.

Appendix VII: Comments from the Department of Defense

OFFICE OF THE UNDER SECRETARY OF DEFENSE
3000 DEFENSE PENTAGON
WASHINGTON, DC 20301-3000

ACQUISITION,
TECHNOLOGY
AND LOGISTICS

December 20, 2001

Mr. Jack L. Brock
Managing Director, Acquisition and Sourcing Management
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Brock:

This is the Department of Defense (DoD) response to the General Accounting Office (GAO) draft report to Congressional Requestors, GAO-02-125, "MISSILE DEFENSE: Review of Allegations About an Early National Missile Defense Flight Test," received November 1, 2001 (GAO Code 707558). The Department appreciates the opportunity to comment on the draft report.

The Department concurs with the comments contained in the draft report (GAO did not have any recommendations) and recommends minor changes to the draft which will downgrade the report classification from Secret to Unclassified.

Sincerely,

George R. Schneiter
Director
Strategic and Tactical Systems

Appendix VIII: Major Contributors

Acquisition and Sourcing Management	Bob Levin, Director Barbara Haynes, Assistant Director Cristina Chaplain, Assistant Director, Communications David Hand, Analyst-in-charge Subrata Ghoshroy, Technical Advisor William Petrick, Analyst
Applied Research and Methods	Nabajyoti Barkakati, Senior Level Technologist Hai Tran, Senior Level Technologist
General Counsel	Stephanie May, Assistant General Counsel

GAO's Mission

The General Accounting Office, the investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO e-mail this list to you every afternoon, go to www.gao.gov and select "Subscribe to daily e-mail alert for newly released products" under the GAO Reports heading.

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office
P.O. Box 37050
Washington, D.C. 20013

To order by Phone:
Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

Visit GAO's Document Distribution Center

GAO Building
Room 1100, 700 4th Street, NW (corner of 4th and G Streets, NW)
Washington, D.C. 20013

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm,
E-mail: fraudnet@gao.gov, or
1-800-424-5454 or (202) 512-7470 (automated answering system).

Public Affairs

Jeff Nelligan, Managing Director, NelliganJ@gao.gov (202) 512-4800
U.S. General Accounting Office, 441 G. Street NW, Room 7149,
Washington, D.C. 20548